



ELIAS MANENSA SABE

**MODELOS ADITIVOS GENERALIZADOS PARA LOCAÇÃO,
ESCALA E FORMA: ANÁLISE HISTÓRICA E A PROPOSIÇÃO
DE NOVAS DISTRIBUIÇÕES ZERO-AJUSTADAS**

**LAVRAS – MG
2025**

ELIAS MANENSA SABE

**MODELOS ADITIVOS GENERALIZADOS PARA LOCAÇÃO, ESCALA E FORMA:
ANÁLISE HISTÓRICA E A PROPOSIÇÃO DE NOVAS DISTRIBUIÇÕES
ZERO-AJUSTADAS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para obtenção do título de Mestre.

Prof. Dr. Luiz Ricardo Nakamura
Orientador

Prof^a. Dr^a Andréa Cristina Konrath
Coorientador

**LAVRAS – MG
2025**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Sabe, Elias Manensa

Modelos aditivos generalizados para locação, escala e forma: análise histórica e a proposição de novas distribuições zero-ajustadas / Elias Manensa Sabe. – 2025.

145 p. : il.

Dissertação (mestrado acadêmico)–Universidade Federal de Lavras, 2025.

Orientador: Prof. Dr. João Domingos Scalon.

Coorientador: Prof^ª. Dr^ª Andréa Cristina Konrath.

Bibliografia.

1. Distribuição Zero-Ajustada. 2. zBCCG e zBCPE. 3. GAMLSS. I. Nakamura, Luiz Ricardo. II. Konrath, Andréa Cristina.

ELIAS MANENSA SABE

**MODELOS ADITIVOS GENERALIZADOS PARA LOCAÇÃO, ESCALA E FORMA:
ANÁLISE HISTÓRICA E A PROPOSIÇÃO DE NOVAS DISTRIBUIÇÕES
ZERO-AJUSTADAS**

**GENERALIZED ADDITIVE MODELS FOR LOCATION, SCALE, AND SHAPE:
HISTORICAL ANALYSIS AND THE PROPOSAL OF NEW ZERO-ADJUSTED
DISTRIBUTIONS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para obtenção do título de Mestre.

APROVADA em 18 de fevereiro de 2025.

Prof^a. Dr^a Gislene Araújo Pereira UNIFAL
Prof. Dr. Pedro Henrique Ramos Cerqueira UEL

Prof. Dr. Luiz Ricardo Nakamura
Orientador

Prof^a. Dr^a Andréa Cristina Konrath
Co-Orientador

**LAVRAS – MG
2025**

*Dedico este trabalho a todos os benefícios que advêm do esforço da minha família e, in memoriam, ao meu irmão **Matheus M. Sabe Vinho** (1980–2024), pelo carinho dedicado.
Obrigado e descanse em paz!*

AGRADECIMENTOS

A realização e elaboração de uma dissertação constituem um caminho repleto de desafios, incertezas, decepções, ansiedades, convicções e conquistas. Esta jornada é percorrida na companhia direta e/ou indireta de várias personalidades. Na impossibilidade de mencionar a todos e reconhecendo que alguns sacrifícios precisaram ser feitos, gostaria de realçar aqueles que mais se destacaram ao longo dessa caminhada.

Nesse sentido, meu primeiro agradecimento se estende a Deus, Todo-Poderoso, pelo dom da vida e por permitir que esta dissertação chegasse ao seu término, pois Ele é o maior "Estatístico" que existe. Agradeço ao Instituto de Ciências Exatas e Tecnológicas (ICET), em particular ao Departamento de Ciências Exatas (DES) da Universidade Federal de Lavras, e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Quero expressar minha sincera admiração pela disponibilidade, escuta, orientação, empenho, encorajamento e apoio dos meus orientadores, Prof. Dr. Luiz Ricardo Nakamura e Prof^a. Dr^a Andréa Cristina Konrath, pela paciência e sabedoria que demonstraram.

Aos meus irmãos e minhas irmãs, bem como aos demais familiares, agradeço pela amizade e pelo apoio, seja declarado ou silencioso, ao longo deste processo. Ao Prof. Dr. Geraldo Luís Macalane, diretor da Faculdade de Letras da Universidade Rovuma, pelo incentivo na conclusão do processo e pela acessibilidade durante as dificuldades enfrentadas ao longo do curso.

Gostaria de expressar minha mais sincera gratidão aos meus colegas que, durante esses dois anos, tornaram-se como irmãos. Às irmãs que ganhei ao longo deste processo acadêmico: Viviane, Vanessa, Neila e Nilza, pelos momentos que compartilhamos. Agradeço também aos amigos que incansavelmente estiveram ao meu lado, tanto nos momentos bons quanto nos desafiadores. Embora seus nomes não estejam registrados nesta simples folha, permanecem gravados na memória como co-participantes desta realização acadêmica.

De forma muito especial, gostaria de direcionar meus agradecimentos aos meus heróis: pai, mãe, amigos e conselheiros, carinhosamente conhecidos como vô JJ (João José Chauque) e vô Zuzu (Zulfa Amade Valigi Chauque), pilares fundamentais na minha vida. A minha mãe acadêmica, Yohana Taise Hoffmann pela recepção calorosa. E a todos aqueles que não foram mencionados, mas que, direta ou indiretamente, estiveram presentes na minha vida, contribuindo com apoio moral para que esta formação chegasse ao fim sem sobressaltos, muito obrigado.

A política serve a um momento no presente, mas uma equação é eterna.

(Albert Einstein)

RESUMO

Do ponto de vista da classe dos modelos de regressão remontam ao século XIX e visa compreender como um conjunto de variáveis preditoras influencia ou explica uma ou mais variáveis respostas. Dentro dessa classe de modelos de regressão univariados, podemos construir uma linha histórica que começa com os modelos lineares (LM), seguidos pelos modelos lineares generalizados (GLM), modelos aditivos generalizados (GAM) e, finalmente, os modelos aditivos generalizados para localização, escala e forma (GAMLSS), também conhecidos atualmente como modelos de regressão distribucional. Os GAMLSS podem ser considerados os modelos de regressão mais flexíveis disponíveis na literatura, dada sua versatilidade para modelar respostas com diferentes características, como, por exemplo, forte assimetria, diferentes graus de curtose ou excesso de zeros. Este estudo teve objetivo construir distribuições contínuas com probabilidade extra de ocorrência de zeros (zero-ajustadas) utilizando os modelos aditivos generalizados para localização, escala e forma (GAMLSS). Para alcançar o objetivo geral, foram elencadas duas etapas. A primeira etapa, é apresentado um artigo que versa sobre a evolução histórica dos GAMLSS, desde seus primórdios, destacando suas principais contribuições e inovações em diversas áreas do conhecimento. Já à segunda etapa, é apresentado um artigo que busca contribuir para o avanço da modelagem de dados com os GAMLSS, focando em respostas contínuas positivas com excesso de zeros, desenvolvendo novas distribuições zero-ajustadas. Foram propostas duas distribuições pertencentes à família Box-Cox: Box-Cox Cole e Green Zero-Ajustada (zBCCG) e Box-Cox Exponencial Potência Zero-Ajustada (zBCPE). Os estudos de simulação mostraram que os estimadores de máxima verossimilhança para os parâmetros das distribuições zBCCG e zBCPE fornecem resultados consistentes para diferentes tamanhos amostrais e cenários, incluindo distribuições simétricas, assimétricas positivas e negativas, bem como cenários platicúrticos e leptocúrticos. Para o estudo de caso analisado a distribuição zBCPE apresentou maior flexibilidade e melhor desempenho estatístico em comparação a zBCCG. Assim pode-se concluir que, os modelos GAMLSS têm um impacto significativo em diversas áreas do conhecimento, como evidenciado pela crescente produção científica, além de sua reconhecida flexibilidade na modelagem de dados com características complexas.

Palavras-chave: modelos de regressão; GAMLSS; distribuições zero-ajustadas; modelagem estatística; análise bibliométrica; seguro agrícola.

ABSTRACT

From the perspective of the class of regression models, which dates back to the 19th century, the aim is to understand how a set of predictor variables influences or explains one or more response variables. Within this class of univariate regression models, we can trace a historical line that begins with linear models (LM), followed by generalized linear models (GLM), generalized additive models (GAM), and finally, generalized additive models for location, scale, and shape (GAMLSS), also currently known as distributional regression models. GAMLSS can be considered the most flexible regression models available in the literature, given their versatility in modeling responses with different characteristics, such as strong asymmetry, varying degrees of kurtosis, or excess zeros. This study aimed to construct continuous distributions with an extra probability of zero occurrence (zero-adjusted) using generalized additive models for location, scale, and shape (GAMLSS). To achieve the main objective, two stages were outlined. The first stage presents an article that discusses the historical evolution of GAMLSS, from their origins, highlighting their main contributions and innovations in various fields of knowledge. The second stage presents an article that seeks to contribute to the advancement of data modeling with GAMLSS, focusing on positive continuous responses with excess zeros, developing new zero-adjusted distributions. Two distributions belonging to the Box-Cox family were proposed: the Zero-Adjusted Box-Cox Cole and Green (zBCCG) and the Zero-Adjusted Box-Cox Power Exponential (zBCPE). Simulation studies showed that the maximum likelihood estimators for the parameters of the zBCCG and zBCPE distributions provide consistent results for different sample sizes and scenarios, including symmetric, positively and negatively skewed distributions, as well as platykurtic and leptokurtic scenarios. For the analyzed case study, the zBCPE distribution demonstrated greater flexibility and better statistical performance compared to the zBCCG. Thus, it can be concluded that GAMLSS models have a significant impact on various fields of knowledge, as evidenced by the growing scientific production, in addition to their recognized flexibility in modeling data with complex characteristics.

Keywords: regression models; GAMLSS; zero-adjusted distributions; statistical modeling; biometric analysis; agricultural insurance.

INDICADORES DE IMPACTO

O estudo demonstra um grande potencial para impactar a sociedade, especialmente nas áreas de modelagem estatística e seguro agrícola, destacando a necessidade de métodos mais flexíveis e modernos para a modelagem de fenômenos complexos. Ao mostrar que modelos de regressão menos complexos não são adequados para explicar completamente a resposta devido às diferentes estruturas de regressão presentes na distribuição dos dados, o estudo enfatiza a importância de abordagens mais avançadas, como os Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS). Portanto, o estudo tem o potencial de contribuir significativamente para o avanço da ciência agrária, fornecendo métodos mais eficazes e precisos para modelar o valor da indenização. Isso pode beneficiar diversas áreas, desde que a estrutura do conjunto de dados apresente características semelhantes, incluindo agricultura, medicina, finanças, economia, gestão de desastres naturais e engenharia.

IMPACT INDICATORS

The study demonstrates significant potential to impact society, particularly in the fields of statistical modeling and agricultural insurance, highlighting the need for more flexible and modern methods in modeling complex phenomena. By showing that less complex regression models are inadequate to fully explain the response due to the different regression structures present in the data distribution, the study emphasizes the importance of more advanced approaches, such as Generalized Additive Models for Location, Scale, and Shape (GAMLSS). Therefore, the study has the potential to significantly contribute to the advancement of agricultural science by providing more effective and precise methods for modeling indemnity values. This could benefit various fields, as long as the data structure exhibits similar characteristics, including agriculture, medicine, finance, economics, disaster management, and engineering.

LISTA DE FIGURAS

Figura 2.1 – Diferentes tipos de falhas do modelo indicadas pelo <i>worm plot</i> : (a) e (b) indicam falha no ajuste correto da locação; (c) e (d) indicam falha no ajuste correto da escala; (e) e (f) indicam falha no ajuste correto da assimetria; e (g) e (h) indicam falha no ajuste correto da curtose.	35
Figura 3.1 – Screening procedure for selecting GAMLSS papers	46
Figura 3.2 – Total number of articles published over the 20-year period.	47
Figura 3.3 – Countries of the first authors	47
Figura 3.4 – Word cloud of author’s keywords in the papers	48
Figura 3.5 – Top ten first authors with the most publications on GAMLSS.	50
Figura 3.6 – Top ten authors with the highest <i>g</i> -index.	50
Figura 3.7 – Network visualization of co-authorship collaboration on GAMLSS.	51
Figura 3.8 – Top ten journals with the most publications on GAMLSS.	52
Figura 3.9 – Trend topics in the 20-year time span.	53
Figura 3.1 – Distribuição zBCCG para diferentes combinações de valores de parâmetros	122
Figura 3.2 – Distribuição zBCPE para diferentes combinações de valores de parâmetros: (a) forma leptocúrtica; e (b) forma platicúrtica	124
Figura 3.3 – Splines cúbicas ajustadas para explicar a relação entre a probabilidade de não ocorrência de um sinistro, em que: (a) latitude (°); (b) longitude (°); (c) área da lavoura (ha); (d) velocidade média mensal do vento (m/s); e (e) logaritmo da produtividade estimada (kg/ha)	128
Figura 3.4 – Splines cúbicas ajustadas para explicar a relação entre o valor mediano e variabilidade das indenizações (dado que um sinistro ocorreu) em que: (a) área da lavoura (ha); (b) média mensal de dias com precipitação; e (c) a capacidade de água disponível (CAD).	130
Figura 3.5 – Worm plot obtido a partir do modelo GAMLSS zero-ajustado com base na distribuição BCPE	131

LISTA DE QUADROS

Quadro 2.1 – Distribuições contínuas implementadas em <code>gamlss.dist</code> (com funções de ligação predefinidas).	27
Quadro 2.2 – Distribuições discretas implementadas em <code>gamlss.dist</code> (com funções de ligação predefinidas).	28
Quadro 2.3 – Distribuições mistas implementadas no pacote <code>gamlss.dist</code>	29
Quadro 2.4 – Diferentes formatos de <i>worm plot</i> e interpretações	34

LISTA DE TABELAS

Tabela 3.1 – Top ten most cited papers throughout the years	49
Tabela 3.1 – Estimativas médias, viéses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCCG, considerando diferentes cenários.	123
Tabela 3.2 – Estimativas médias, viéses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCPE, considerando diferentes cenários em que a forma da distribuição é leptocúrtica.	125
Tabela 3.3 – Estimativas médias, viéses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCPE, considerando diferentes cenários em que a forma da distribuição é platicúrtica.	126
Tabela 3.4 – Critério de Informação de Akaike (AIC) para os modelos GAMLSS ajustados com as distribuições zBCPE e zBCCG	129

LISTA DE SIGLAS

MG	Minas Gerais
UNIFAL	Universidade Federal de Alfenas
UEL	Universidade Estadual de Londrina
UFLA	Universidade Federal de Lavras
LM	Modelos lineares
GLM	Modelos lineares generalizados
GAM	Modelos aditivos generalizados
GAMLSS	Modelos aditivos generalizados para locação, escala e forma
ZAGA	Distribuição gama zero-ajustada
ZAIG	Distribuição normal inversa zero-ajustada
zBCCG	Distribuição Box-Cox Cole e Green Zero-Ajustada
zBCPE	Distribuição Box-Cox Exponencial Potência Zero-Ajustada
BCCG	Distribuição Box-Cox Cole e Green
BCPE	Distribuição Box-Cox Exponencial Potência
AIC	Critério de informação de Akaike
BIC	Critério de informação bayesiano
EQM	Erros quadráticos médios
IBGE	Instituto Brasileiro de Geografia e Estatística
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
INMET	Instituto Nacional de Meteorologia
ANA	Agência Nacional de Águas e Saneamento Básico
CAD	Capacidade de água disponível

SUMÁRIO

	PRIMEIRA PARTE - UM PANORAMA GERAL	15
1	INTRODUÇÃO	17
1.1	Objetivos	18
1.2	Estrutura da dissertação	19
1.3	Limitações	19
2	REFERENCIAL TEÓRICO	21
2.1	Modelos de regressão	21
2.1.1	Modelo de regressão linear	21
2.1.2	Modelo linear generalizado	22
2.1.3	Modelo aditivo generalizado	23
2.1.3.1	P-splines	23
2.1.4	Modelos aditivos generalizados para locação, escala e forma	24
2.1.4.1	Distribuições da família GAMLSS	26
2.1.4.1.1	Distribuições zero-ajustadas	30
2.1.4.2	Método de estimação dos parâmetros	31
2.1.4.3	Seleção das variáveis	32
2.2	Diagnóstico do modelo	33
3	METODOLOGIA	37
3.1	Estudo de caso: seguros agrícolas	38
	SEGUNDA PARTE - ARTIGOS	40
	ARTIGO 1 - Historical developments and research trends in GAMLSS . . .	41
1	Introduction	43
2	Materials and Methods	44
3	Results and discussion	45
4	Concluding remarks	54
	ARTIGO 2 - Novas distribuições zero-ajustadas baseada na família Box-Cox	113
1	Introdução	115
2	Estrutura dos GAMLSS	116
2.1	Família de distribuição Box-Cox: uma breve revisão	117
2.2	Distribuições BCCG e BCPE zero-ajustadas	118
2.3	Seleção de covariáveis e diagnóstico do modelo	119

3	Estudo de simulação	121
3.1	Distribuição zBCCG	121
3.2	Distribuição zBCPE	122
4	Aplicação: seguro rural	124
4.1	Modelo para $Y = 0$	127
4.2	Modelo para $Y > 0$	129
5	Considerações finais	131
	TERCEIRA PARTE - CONCLUSÃO	137
4	Conclusão	139
	REFERÊNCIAS	141

PRIMEIRA PARTE - UM PANORAMA GERAL

1 INTRODUÇÃO

A modelagem estatística preditiva por meio de modelos matemáticos estatísticos tem atraído o interesse de muitos pesquisadores. Gompertz (1825), em seu artigo histórico, apresentou o primeiro modelo matemático relacionado à mortalidade. Desde então, a proposta de modelos estatísticos para diferentes tipos de conjuntos de dados tem evoluído continuamente. Destacam-se os modelos lineares (LM), modelos lineares generalizados (GLM) (Nelder; Wedderburn, 1972), modelos aditivos generalizados (GAM) (Hastie; Tibshirani, 1990) e modelos aditivos generalizados para localização, escala e forma (GAMLSS) (Rigby; Stasinopoulos, 2005). Os GAMLSS, atualmente também denominados modelos de regressão distribucional (Heller; Robledo; Marschner, 2022), são uma alternativa para modelagem, permitindo a construção de estruturas de regressão não apenas para a média da variável resposta, mas também para outros parâmetros que escrevem sua distribuição, como por exemplo, variância, assimetria, excesso de curtose e, em casos específicos, inflação de zeros, dentre outros, possibilitando identificar características distintas que influenciam cada um desses parâmetros (Rigby; Stasinopoulos, 2005; Stasinopoulos et al., 2017; Rigby et al., 2019).

Com a flexibilidade oferecida pelos GAMLSS, diversas áreas do conhecimento têm se beneficiado de sua aplicação, como por exemplo, ciências atuariais (Nakamura et al., 2017), ciências imobiliárias (De Bastiani et al., 2018), ciências médicas (Ramires et al., 2018), agricultura de precisão (Righetto et al., 2019), ciências florestais (Oliveira et al., 2019), segurança pública (Ramires et al., 2019a), ciências naturais (He et al., 2021), psicologia (Timmerman; Voncken; Albers, 2021), ecologia marinha (Costa et al., 2022) e ciências animais (Nakamura et al., 2022b). Esse amplo espectro demonstra a versatilidade e a eficácia desses modelos para lidar com diferentes estruturas de dados.

Devido à sua ampla flexibilidade, os GAMLSS possibilitam o uso de uma vasta gama de distribuições. Essas distribuições são classificadas em três grandes grupos: contínuas, discretas e mistas (Rigby et al., 2019). As distribuições mistas combinam características de distribuições discretas e contínuas e incluem extensões mais complexas, como as distribuições inflacionadas, que modelam proporções no intervalo $[0, 1]$ e as zero-ajustadas, aplicáveis a variáveis no intervalo $(0, \infty)$ (Rigby et al., 2019), foco deste estudo.

Considerando que a variável resposta possui suporte em $[0, \infty)$, este estudo concentra-se em distribuições contínuas positivas zero-ajustadas. Tais distribuições são mistas, combinando uma distribuição discreta, que modela a probabilidade de não ocorrência de um evento ($Y = 0$),

e uma distribuição contínua, que descreve a ocorrência do evento ($Y > 0$) (Stasinopoulos et al., 2023; Rigby et al., 2019).

Neste sentido, duas das distribuições mais utilizadas nesta classe são a gama zero-ajustada (*zero-adjusted gamma*, ZAGA) e a normal inversa zero-ajustada (*zero-adjusted inverse Gaussian*, ZAIG). Essas distribuições são frequentemente aplicadas a dados com excesso de zeros e possíveis assimetrias. Rigby et al. (2019) destacaram que, para modelar adequadamente dados com essas características, como a presença de um número elevado de zeros combinada com uma distribuição assimétrica das respostas positivas, pode ser insuficiente acomodar uma única suposição. Assim, propuseram uma estrutura de modelagem para dados com excesso de zeros por meio de distribuições zero-ajustadas (Stasinopoulos et al., 2023).

Dado esse contexto, surge o seguinte questionamento: é possível construir novas distribuições zero-ajustadas no contexto dos modelos aditivos generalizados para locação, escala e forma?

Com base nesse problema, propomos a construção de dois modelos derivados das distribuições zero-ajustadas pertencentes à família Box-Cox: a distribuição Box-Cox Cole e Green Zero-Ajustada (*zero-adjusted Box-Cox Cole and Green*, zBCCG) e a Box-Cox Exponencial Potência Zero-Ajustada (*zero-adjusted Box-Cox power exponential*, zBCPE), com quatro e cinco parâmetros, respectivamente, uma vez oferecem alta flexibilidade na captura de padrões complexo em dados assimétricos e com excesso de curtose (Rigby; Stasinopoulos, 2004). Cada uma dessas distribuições recebe um parâmetro adicional devido à modelagem separada da probabilidade de zeros e do componente contínuo (Stasinopoulos et al., 2023).

1.1 Objetivos

O objetivo deste trabalho versa construir distribuições contínuas com probabilidade extra de ocorrência de zeros (zero-ajustadas) utilizando os modelos aditivos generalizados para locação, escala e forma (GAMLSS). Em complementaridade ao objetivo geral, o estudo orientou-se pelos seguintes objetivos específicos:

- a) investigar a evolução histórica dos GAMLSS, destacando suas principais contribuições e inovações ao longo do tempo (2004-2023) em diferentes áreas de conhecimento;
- b) elaborar matematicamente novas distribuições zero-ajustadas no contexto dos GAMLSS;
- c) implementar as distribuições desenvolvidas no *software* livre R;

- d) avaliar a eficácia dos modelos implementados no contexto dos seguros agrícolas.

1.2 Estrutura da dissertação

Esta dissertação está dividida em duas partes. A primeira apresenta um panorama geral que inclui introdução, objetivos, estrutura, limitações, referencial teórico e os aspectos metodológicos. A segunda parte inclui dois artigos científicos: o primeiro, intitulado “*Historical developments and research trends in GAMLSS*”, e o segundo, aborda “Novas distribuições zero-ajustadas baseada na família Box-Cox”, que propõe a construção de novas distribuições contínuas com probabilidade extra de ocorrência de zeros, utilizando os modelos aditivos generalizados para locação, escala e forma (GAMLSS). Além disso, são apresentadas, de forma sucinta, as considerações finais do estudo, bem como sugestões para trabalhos e pesquisas futuras. Por fim, apresentam-se as referências consultadas durante a elaboração desta dissertação e os apêndices.

1.3 Limitações

O desenvolvimento de qualquer trabalho científico enfrenta uma série de limitações, condicionadas por diversos fatores intrínsecos e extrínsecos, e este estudo não é uma exceção. As principais limitações deste trabalho foram as seguintes:

- a) dependência de ferramentas computacionais específicas e a complexidade do ajuste dos modelos zero-ajustados, que exigem maior esforço computacional e técnica;
- b) foi realizada uma análise bibliométrica ao invés de uma sistemática, devido à grande quantidade de trabalhos existentes, o que tornaria inviável a análise no escopo proposto;
- c) muitos artigos reportados e relacionados nesta busca não apresentavam claramente suas limitações e possíveis lacunas, dificultando a identificação de questões não abordadas;
- d) limitações no uso do *Biblioshiny* para manipulação e visualização da rede de colaboração de coautoria, com opções restritas de configuração e design.

2 REFERENCIAL TEÓRICO

Neste capítulo, serão apresentados os principais conceitos e modelos teóricos de regressão que fundamentam este estudo em questão, desde os primeiros modelos de regressão linear até os modelos aditivos generalizados para localização, escala e forma, abordando sua evolução até os dias atuais.

2.1 Modelos de regressão

Nesta seção, visa fornecer uma base sólida para a compreensão dos métodos utilizados e a interpretação dos resultados obtidos. Inicialmente, abordaremos os modelos de regressão linear, linear generalizado, aditivo generalizado e modelos aditivos generalizados para localização, escala e forma, que são ferramentas estatísticas essenciais para analisar a relação entre variáveis.

2.1.1 Modelo de regressão linear

O termo “regressão” define um conjunto amplo de técnicas estatísticas usadas para modelar relações entre variáveis e prever o valor de uma ou mais variáveis dependentes a partir de um conjunto de variáveis independentes (Rencher; Schaalje, 2008), tendo sido proposto pela primeira vez por Sir Francis Galton em 1885, em um estudo onde demonstrou que a altura dos filhos tende a não refletir exatamente a altura dos pais, mas sim a regredir para a média da população (Nelder; Wedderburn, 1972).

No caso dos modelos de regressão univariados, matematicamente, consideramos Y_i , $i = 1, \dots, n$, como uma variável resposta. Deste modo, o modelo de regressão linear (MRL) pode ser expresso como

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_r x_{ir} + \varepsilon_i, \quad (2.1)$$

em que x_1, \dots, x_r são variáveis independentes, cujos valores são supostamente conhecidos, $\beta_0, \beta_1, \dots, \beta_r$ são os parâmetros do modelo a serem estimados e $\varepsilon_i \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ são erros aleatórios supostos independentes e identicamente distribuídos (iid).

Ao estabelecer o modelo de regressão linear apresentado em (2.1), pressupomos que: (i) a relação entre as variáveis explicativas e a variável resposta é linear; (ii) os valores das variáveis explicativas são fixos; (iii) a média do erro é nula, isto é, $E(\varepsilon_i) = 0$; e (iv) a variância do erro σ^2 é constante.

O modelo (2.1) pode ser escrito de forma matricial, como,

$$\begin{aligned} \mathbf{Y} &\stackrel{\text{ind}}{\sim} N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \\ \boldsymbol{\mu} &= \mathbf{X}\boldsymbol{\beta}, \end{aligned} \quad (2.2)$$

em que $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ é o vetor da variável resposta, \mathbf{X} é a matriz de delineamento com dimensão $(n \times b)$, sendo $b = r + 1$, $\boldsymbol{\beta} = (\beta_0, \dots, \beta_r)^\top$ é o vetor de parâmetros, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ é o vetor de média, $\boldsymbol{\sigma}^2 = (\sigma^2, \dots, \sigma^2)^\top$ é um vetor de variância constante (Rencher; Schaalje, 2008; Stasinopoulos et al., 2017).

2.1.2 Modelo linear generalizado

Um dos pressupostos mais rigorosos de se trabalhar com os modelos lineares é que a variável resposta em estudo deve, necessariamente, seguir uma distribuição normal. Com o intuito de mitigar tal pressuposto, Nelder e Wedderburn (1972) propuseram os modelos lineares generalizados (GLM), nos quais é possível considerar outras distribuições de probabilidade para a variável resposta, sendo essas distribuições pertencentes à família exponencial (Paula, 2004). Uma distribuição pertencente à família exponencial ($\mathcal{F}\mathcal{E}$) é aquela cuja função (densidade) de probabilidade pode ser escrita na forma

$$f(y|\boldsymbol{\mu}, \phi) = \exp \left\{ \frac{y\boldsymbol{\theta} - b(\boldsymbol{\theta})}{\phi} + c(y, \phi) \right\}, \quad (2.3)$$

em que ϕ é o parâmetro de dispersão, $b(\cdot)$ e $c(\cdot)$ são funções conhecidas, $E(Y) = \boldsymbol{\mu} = b'(\boldsymbol{\theta})$ e $Var(Y) = \phi b''(\boldsymbol{\theta})$. Alguns exemplos de distribuições pertencentes à esta classe são: normal, Poisson, binomial, gama e normal inversa (Cordeiro; Demétrio; Moral, 2024).

De maneira geral, um GLM possui três componentes: (i) componente aleatório - uma variável resposta cuja distribuição pertence à $\mathcal{F}\mathcal{E}$; (ii) componente sistemático - as variáveis explicativas são consideradas no modelo na forma de uma soma linear de efeitos; e (iii) função de ligação - especifica o relacionamento entre o componente aleatório e o componente sistemático, isto é, vincula $E(Y)$ com as variáveis explicativas (Cordeiro; Demétrio; Moral, 2024).

Matematicamente, o GLM pode ser escrito como

$$Y_i \stackrel{\text{ind}}{\sim} \mathcal{F}\mathcal{E}(\mu_i, \phi), \quad (2.4)$$

em que

$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_r x_{ir} \quad (2.5)$$

e $g(\cdot)$ é chamada de função de ligação. Matricialmente, podemos reescrever

$$\begin{aligned} \mathbf{Y} &\stackrel{\text{ind}}{\sim} \mathcal{F}\mathcal{E}(\boldsymbol{\mu}, \boldsymbol{\phi}) \\ \boldsymbol{\eta} &= g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.6)$$

em que $\boldsymbol{\eta}$ é chamado de preditor linear e $\boldsymbol{\phi} = (\phi, \dots, \phi)^\top$ é um vetor de constantes ϕ .

2.1.3 Modelo aditivo generalizado

Uma das limitações dos GLM é considerar que a relação entre a média da resposta e as variáveis explicativas é linear. Neste sentido, Hastie e Tibshirani (1990) desenvolveram os modelos aditivos generalizados (GAM) que, basicamente, adicionam ao preditor linear, uma soma de funções de suavização não-paramétricas. Podemos escrever um GAM como

$$\begin{aligned} \mathbf{Y} &\stackrel{\text{ind}}{\sim} \mathcal{F}\mathcal{E}(\boldsymbol{\mu}, \boldsymbol{\phi}) \\ \boldsymbol{\eta} &= g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + \sum_{j=1}^J s_j(\mathbf{x}_j), \end{aligned} \quad (2.7)$$

em que $s_j(\cdot)$ é a função de suavização não-paramétrica aplicada à covariável \mathbf{x}_j . Existem várias funções de suavização, dentre elas, *splines* cúbicas, polinômios fracionais e potência, ajuste não linear, *P-splines*, *P-splines* cíclicos, *P-splines* monótonos, *P-splines* encolhidos em zero e *P-splines* de coeficientes variantes (Stasinopoulos et al., 2017). Ramires et al. (2019b) enfatizam que quando uma função de suavização é adicionada ao modelo, usualmente, não se realizam testes de hipóteses, mas sim, interpretamos seu relacionamento graficamente.

2.1.3.1 P-splines

Os *P-splines*, propostos por Eilers e Marx (1996), resultam da combinação da base dos *B-splines* (De Boor, 1978), que é uma base ortogonal de um polinômio por partes, com uma penalidade aplicada aos coeficientes do modelo. A ideia original foi baseada no trabalho de O'Sullivan (1986) e é derivada como a solução analítica do problema de mínimos quadrados

penalizados que minimiza em relação ao parâmetro $\boldsymbol{\gamma}$, dado por

$$\mathbf{S} = (\mathbf{y} - \mathbf{B}\boldsymbol{\gamma})^\top \mathbf{W}(\mathbf{y} - \mathbf{B}\boldsymbol{\gamma}) + \lambda \boldsymbol{\gamma}^\top \mathbf{P}\boldsymbol{\gamma}, \quad (2.8)$$

em que \mathbf{B} é uma matriz $n \times r$ de base B-spline, $\boldsymbol{\gamma}$ é um vetor paramétrico r -dimensional, \mathbf{W} é uma matriz diagonal de pesos de dimensão $n \times n$, $\mathbf{P} = \mathbf{D}^\top \mathbf{D}$ é uma matriz de penalidades de dimensão $r \times r$ e λ é um parâmetro de suavização. A solução para o problema dos mínimos quadrados penalizados é dada por:

$$\hat{\boldsymbol{\gamma}} = (\mathbf{B}^\top \mathbf{W} \mathbf{B} + \lambda \mathbf{D}^\top \mathbf{D})^{-1} \mathbf{B}^\top \mathbf{W} \mathbf{y}. \quad (2.9)$$

Do ponto de vista de interpretação, quanto maior o valor de λ , maior será a influência da penalidade e, portanto, mais suave será a função ajustada. Existem outros parâmetros implícitos no ajuste de um P-spline, como o número de nós e o grau do polinômio por partes na definição da base \mathbf{B} , e a ordem das diferenças na matriz de penalidade \mathbf{D} . Esses parâmetros afetam o ajuste do modelo, mas não tanto quanto o parâmetro de suavização λ (Stasinopoulos et al., 2017).

O uso do grau do polinômio por partes é assegurado pela continuidade das primeira e segunda derivadas da função ajustada, resultando em uma curva mais suave. A ordem padrão das diferenças na penalidade é dois, assumindo que os coeficientes $\boldsymbol{\gamma}$ se comportam como um passeio aleatório de segunda ordem. O número de nós B-splines excede o número de observações sem prejudicar o método. A simplicidade e flexibilidade dos P-splines são destacadas, além de uma interpretação de efeitos aleatórios (Hastie; Tibshirani, 1990; Stasinopoulos et al., 2017).

2.1.4 Modelos aditivos generalizados para locação, escala e forma

Apesar da flexibilização alcançada com os GAM, ainda há dois potenciais problemas em sua utilização: (i) apenas o parâmetro da média é explicitamente modelado com base nas covariáveis; e (ii) a distribuição da variável resposta ainda deve, necessariamente, pertencer à $\mathcal{F}\mathcal{E}$.

Para superar os entraves supracitados, os modelos aditivos generalizados para locação, escala e forma (GAMLSS) foram propostos por Rigby e Stasinopoulos (2005). Os GAMLSS são uma abordagem para aprendizado estatístico, sendo considerados como uma classe de mo-

delos de regressão univariados semi-paramétricos que generalizam os modelos anteriormente apresentados (LM, GLM e GAM). Nesta abordagem, qualquer distribuição \mathcal{D} pode ser assumida para representar a resposta e todo e qualquer parâmetro desta distribuição pode ser diretamente modelado por meio de funções lineares e de suavização não-paramétricas, isto é, são construídas diferentes estruturas de regressão para cada um dos parâmetros.

Matematicamente, um GAMLSS pode ser definido como

$$\begin{aligned} \mathbf{Y} &\overset{\text{ind}}{\sim} \mathcal{D}(\boldsymbol{\theta}_k) \\ g_k(\boldsymbol{\theta}_k) &= \mathbf{X}_k \boldsymbol{\beta}_k + \sum_{j=1}^{J_k} s_{kj}(\mathbf{x}_{kj}), \end{aligned} \quad (2.10)$$

$\boldsymbol{\theta}_k = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p)^\top$ é o vetor de parâmetros associados à distribuição \mathcal{D} da variável resposta, $g_k(\cdot)$, $k = 1, \dots, p$, denotam as funções de ligação relacionadas aos k -parâmetros da distribuição - usualmente escolhidas com base no suporte do parâmetro (De Bastiani et al., 2018), \mathbf{X}_k é uma matriz de delineamento, $\boldsymbol{\beta}_k = (\beta_{1k}, \dots, \beta_{J_k k})^\top$ é o vetor de parâmetros e $s_{kj}(\cdot)$ é uma função de suavização que explica o relacionamento da covariável \mathbf{x}_{kj} . Se $\sum_{j=1}^{J_k} s_{kj}(\mathbf{x}_{kj}) = 0$, então o GAMLSS é reduzido à sua versão linear paramétrica (Nakamura et al., 2019; Righetto et al., 2019).

No caso em que $k = 1, \dots, 4$, o modelo (1) se reduz ao caso

$$\begin{aligned} Y &\overset{\text{ind}}{\sim} \mathcal{D}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{v}, \boldsymbol{\tau}) \\ g_1(\boldsymbol{\mu}) &= \mathbf{X}_1 \boldsymbol{\beta}_1 + \sum_{j=1}^{J_1} s_{1j}(\mathbf{x}_{1j}) \\ g_2(\boldsymbol{\sigma}) &= \mathbf{X}_2 \boldsymbol{\beta}_2 + \sum_{j=1}^{J_2} s_{2j}(\mathbf{x}_{2j}) \\ g_3(\mathbf{v}) &= \mathbf{X}_3 \boldsymbol{\beta}_3 + \sum_{j=1}^{J_3} s_{3j}(\mathbf{x}_{3j}) \\ g_4(\boldsymbol{\tau}) &= \mathbf{X}_4 \boldsymbol{\beta}_4 + \sum_{j=1}^{J_4} s_{4j}(\mathbf{x}_{4j}), \end{aligned} \quad (2.11)$$

em que, usualmente, $\boldsymbol{\mu}$ e $\boldsymbol{\sigma}$ são parâmetros de localização e escala, respectivamente, e \mathbf{v} e $\boldsymbol{\tau}$ são parâmetros de forma.

Segundo Rigby e Stasinopoulos (2005), as funções de suavização dos GAMLSS podem, na maioria das situações, serem reescritas como $s(\mathbf{x}) = \mathbf{Z}\boldsymbol{\gamma}$, em que \mathbf{Z} é a matriz de base dependendo dos valores de \mathbf{x} e $\boldsymbol{\gamma}$ é um conjunto de parâmetros sujeitos à penalização $\boldsymbol{\lambda}\boldsymbol{\gamma}^\top \mathbf{G}\boldsymbol{\gamma}$, para uma matriz conhecida $\mathbf{G} = \mathbf{D}^\top \mathbf{D}$ em que \mathbf{D} é uma matriz de diferenças e $\boldsymbol{\lambda}$ é um vetor ou escalar de hiperparâmetros que regula o grau de suavização necessário no ajuste.

Dessa forma, podemos reescrever o modelo (2.12) como

$$\begin{aligned} Y &\stackrel{\text{ind}}{\sim} \mathcal{D}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{v}, \boldsymbol{\tau}) \\ g_1(\boldsymbol{\mu}) &= \mathbf{X}_1 \boldsymbol{\beta}_1 + \sum_{j=1}^{J_1} \mathbf{Z}_{1j}(\boldsymbol{\gamma}_{1j}) \\ g_2(\boldsymbol{\sigma}) &= \mathbf{X}_2 \boldsymbol{\beta}_2 + \sum_{j=1}^{J_2} \mathbf{Z}_{2j}(\boldsymbol{\gamma}_{2j}) \\ g_3(\mathbf{v}) &= \mathbf{X}_3 \boldsymbol{\beta}_3 + \sum_{j=1}^{J_3} \mathbf{Z}_{3j}(\boldsymbol{\gamma}_{3j}) \\ g_4(\boldsymbol{\tau}) &= \mathbf{X}_4 \boldsymbol{\beta}_4 + \sum_{j=1}^{J_4} \mathbf{Z}_{4j}(\boldsymbol{\gamma}_{4j}), \end{aligned}$$

em que $\boldsymbol{\gamma}_{kj} \stackrel{\text{iid}}{\sim} N(0, \mathbf{G}_{kj}^-(\boldsymbol{\lambda}_{kj}))$ são os parâmetros de efeito aleatório e $\mathbf{G}_{kj}^-(\boldsymbol{\lambda}_{kj})$ a inversa (generalizada) da matriz simétrica $\mathbf{G}_{kj}(\boldsymbol{\lambda}_{kj})$ de ordem $q_{kj} \times q_{kj}$.

2.1.4.1 Distribuições da família GAMLSS

De forma resumida, como mencionado na Seção 2.1.4, qualquer distribuição \mathcal{D} pode ser utilizada para representar a variável resposta, independentemente dela pertencer, ou não, à família exponencial. Podemos dividi-las em três grandes grupos: distribuições contínuas, discretas e mistas, ou seja, uma mistura de uma distribuição discreta (por exemplo, com probabilidade de valor zero, ou seja, $y = 0$) e uma distribuição contínua (por exemplo, $y > 0$). Destes, há distribuições com características que comportam dados fortemente assimétricos, bem como com caudas pesadas (leptocurtose) e leves (platicurtose). Há um pacote específico da família GAMLSS de distribuições denominado `gamlss.family` (Stasinopoulos; Rigby, 2008), no software R (R CORE TEAM, 2024), onde mais de 100 distribuições já estão implementadas. Algumas dessas distribuições contínuas, discretas e mistas são apresentadas nos Quadros 2.1, 2.2 e 2.3, respectivamente.

Quadro 2.1 – Distribuições contínuas implementadas em `gamlss.dist` (com funções de ligação predefinidas).

(Continua)

Ord.	Distribuição	gamlss	Intervalo	Função de ligação			
				μ	σ	ν	τ
1	beta	BE	(0, 1)	logit	logit	-	-
2	Box-Cox Cole-Green	BCCG	(0, ∞)	ident.	log	ident.	-
3	Box-Cox Cole-Green orig.	BCCGo	(0, ∞)	log	log	ident.	-
4	Box-Cox power exponencial	BCPE	(0, ∞)	ident.	log	ident.	log
5	Box-Cox Power Expon. orig.	BCPEo	(0, ∞)	log	log	ident.	log
6	Box-Cox t	BCT	(0, ∞)	ident.	log	ident.	log
7	Box-Cox t orig.	BCTo	(0, ∞)	log	log	ident.	log
8	exponencial	EXP	(0, ∞)	log	-	-	-
9	exponencial gaussiana	exGAUS	($-\infty$, ∞)	ident.	log	log	-
10	exponencial gen. beta 2	EGB2	($-\infty$, ∞)	ident.	log	log	log
11	gama	GA	(0, ∞)	log	log	-	-
12	beta generalizada tipo 1	GB1	(0, 1)	logit	logit	log	log
13	beta generalizada tipo 2	GB2	(0, ∞)	logit	logit	log	log
14	gama generalizada	GG	(0, ∞)	log	log	ident.	-
15	gaussiana inversa generalizada	GIG	(0, ∞)	log	log	ident.	-
16	t generalizado	GT	($-\infty$, ∞)	ident.	log	log	log
17	Gumbel	GU	($-\infty$, ∞)	ident.	log	-	-
18	gama inversa	IGAMMA	(0, ∞)	log	log	-	-
19	Gaussiana inversa	IGAMMA	(0, ∞)	log	log	-	-
20	reparação SU da Johnson's	JSU	($-\infty$, ∞)	ident.	log	ident.	log
21	SU original de Johnson	JSUo	($-\infty$, ∞)	ident.	log	ident.	log
22	logística	LO	($-\infty$, ∞)	ident.	log	-	-
23	logit normal	LOGITNO	(0, 1)	ident.	log	-	-
24	log normal	LOGNO	(0, ∞)	ident.	log	-	-
25	log normal 2	LOGNO2	(0, ∞)	log	log	-	-
26	log normal (Box-Cox)	LNO	(0, ∞)	ident.	log	fixed	-
27	NET	NET	($-\infty$, ∞)	ident.	log	fixed	fixed
28	normal	NO, NO2	($-\infty$, ∞)	ident.	log	-	-
29	família normal	NOF	($-\infty$, ∞)	ident.	log	-	-
30	Pareto 2	PARETO2	(0, ∞)	log	log	-	-
31	Pareto 2 original	PARETO2o	(0, ∞)	log	log	-	-
32	Pareto 2 repar	GP	(0, ∞)	log	log	-	-
33	potência exponencial	PE	($-\infty$, ∞)	ident.	log	log	-
34	extremo inverso gen.	RGE	$y > \mu - (\sigma/\nu)$	ident.	log	log	-
35	Gumbel invertido	RG	($-\infty$, ∞)	ident.	log	-	-
36	sinh-arcsinh	SHASH	($-\infty$, ∞)	ident.	log	log	log
37	sinh-arcsinh original	SHASHo	($-\infty$, ∞)	ident.	log	ident.	log
38	sinh-arcsinh original 2	SHASHo2	($-\infty$, ∞)	ident.	log	ident.	log
39	normal oblíqua tipo 1	SN1	($-\infty$, ∞)	ident.	log	ident.	-
40	normal oblíqua tipo 2	SN2	($-\infty$, ∞)	ident.	log	log	-

Quadro 2.1 – Distribuições contínuas implementadas em `gamlss.dist` (com funções de ligação predefinidas).

(Conclusão)

Ord.	Distribuição	gamlss	Intervalo	Função de ligação			
				μ	σ	ν	τ
41	potência oblíqua exp. tipo 1	SEP1	$(-\infty, \infty)$	ident.	log	ident.	log
42	potência oblíqua exp. tipo 2	SEP2	$(-\infty, \infty)$	ident.	log	ident.	log
43	potência oblíqua exp. tipo 3	SEP3	$(-\infty, \infty)$	ident.	log	log	log
44	potência oblíqua exp. tipo 4	SEP4	$(-\infty, \infty)$	ident.	log	log	log
45	skew t tipo 1	ST1	$(-\infty, \infty)$	ident.	log	ident.	log
46	skew t tipo 2	ST2	$(-\infty, \infty)$	ident.	log	log	log
47	skew t tipo 3	ST3	$(-\infty, \infty)$	ident.	log	log	log
48	skew t tipo 3 repar	SST	$(-\infty, \infty)$	ident.	log	log	log-2
49	skew t tipo 4	ST4	$(-\infty, \infty)$	ident.	log	ident.	log
50	skew t tipo 5	ST5	$(-\infty, \infty)$	ident.	log	log	log
51	família t	TF	$(-\infty, \infty)$	ident.	log	log	-
52	família t repar	TF2	$(-\infty, \infty)$	ident.	log	log-2	-
53	Weibull	WEI	$(0, \infty)$	log	log	-	-
54	Weibull (PH)	WEI2	$(0, \infty)$	log	log	-	-
55	Weibull à média)	WEI3	$(0, \infty)$	log	log	-	-

Fonte: Adaptado de Rigby et al. (2019)

Quadro 2.2 – Distribuições discretas implementadas em `gamlss.dist` (com funções de ligação predefinidas).

(Continua)

Ord.	Distribuição	gamlss	Intervalo	Função de ligação			
				μ	σ	ν	τ
1	beta binomial	BB	$\{0,1,\dots, n\}$	logit	log	-	-
2	beta neg. binomial	BNB	$\{0,1, 2, \dots\}$	log	log	log	-
3	binomial	BI	$\{0,1,\dots, n\}$	logit	-	-	-
4	geométrica	GEOM	$\{0,1, 2, \dots\}$	log	-	-	-
5	geométrica (original)	GEOMo	$\{0,1, 2, \dots\}$	logit	-	-	-
6	logarítmica	LG	$1, 2, \dots, \infty$	logit	-	-	-
7	Delaporte	DEL	$\{0,1, 2, \dots\}$	log	log	logit	-
8	duplo Poisson	DPO	$\{0,1, 2, \dots\}$	log	log	-	-
9	Multinomial	MULTIN	$1, 2, \dots, n$	R	R	R	-
10	binomial negativa tipo I	NBI	$\{0,1, 2, \dots\}$	log	log	log	-
11	binomial negativa tipo II	NBII	$\{0,1, 2, \dots\}$	log	log	log	-
12	família binomial negativa	NBF	$\{0,1, 2, \dots\}$	log	log	log	ident.
13	Poisson	PO	$\{0,1, 2, \dots\}$	log	-	-	-
14	Poisson inv. Gaussiano	PIG	$\{0,1, 2, \dots\}$	log	log	-	-
15	GIG deslocado de Poisson	PSGIG	$\{0,1, 2, \dots\}$	log	log	logit	logit

Quadro 2.2 – Distribuições discretas implementadas em `gamlss.dist` (com funções de ligação predefinidas).

(Conclusão)

Ord.	Distribuição	gamlss	Intervalo	Função de ligação			
				μ	σ	ν	τ
16	Sichel	SI	{0,1, 2, ...}	log	log	identity	-
17	Sichel (μ a média)	SICHEL	{0,1, 2, ...}	log	log	identity	-
18	Waring (μ a média)	WARING	{0,1, 2, ...}	log	log	-	-
19	Yule (μ a média)	YULE	{0,1, 2, ...}	log	-	-	-
20	zero alt. beta binomial	ZABB	{0,1,..., n}	logit	log	logit	-
21	zero alt. beta neg. binomial	ZABNB	{0,1, 2, ...}	log	log	ident.	logit
22	zero alt. binomial	ZABI	{0,1,..., n}	logit	logit	-	-
23	zero alt. logarítmica	ZALG	{0,1, 2, ...}	logit	logit	-	-
24	zero alt. neg. binomial	ZANBI	{0,1, 2, ...}	log	log	logit	-
25	zero alt. neg. binom. fam.	ZANBF	{0,1, 2, ...}	log	log	log	logit
26	zero alt. PIG	ZAPIG	{0,1, 2, ...}	log	log	logit	-
27	zero alt. Sichel	ZASICHEL	{0,1, 2, ...}	log	log	identity	logit
28	zero alt. poisson	ZAPIG	{0,1, 2, ...}	log	logit	-	-
29	zero alt. zipf	ZAZIPF	{0,1, 2, ...}	log	logit	-	-
30	zero inf. beta binomial	ZIBB	{0,1, 2, ...}	logit	log	logit	-
31	zero inf. beta neg. binom.	ZIBNB	{0,1, 2, ...}	log	log	log	logit
32	zero inf. Binomial	ZIBI	{0,1,..., n}	logit	logit	-	-
33	zero inf. neg. binomial	ZINBI	{0,1, 2, ...}	logit	log	logit	-
34	zero inf. neg. binom. fam.	ZINBF	{0,1, 2, ...}	log	log	log	logit
35	zero inf. Poisson	ZIP	{0,1, 2, ...}	logit	logit	-	-
36	zero inf. poisson (μ a média)	ZIP2	{0,1, 2, ...}	logit	logit	-	-
37	zero inf. PIG	ZIPIG	{0,1, 2, ...}	logit	log	logit	-
38	zero inf. Sichel	ZISICHEL	{0,1, 2, ...}	logit	log	log	identity
39	zipf	ZIPF	{0,1, 2, ...}	log	log	-	-

Fonte: Adaptado de Rigby et al. (2019)

Quadro 2.3 – Distribuições mistas implementadas no pacote `gamlss.dist`.

ord.	Distribuição	gamlss	Intervalo	Função de ligação			
				μ	σ	ν	τ
1	beta inflacionada (em 0)	BEOI	[0, 1)	logit	log	logit	-
2	beta inflacionada (em 0)	BEINFO	[0, 1)	logit	logit	log	-
3	beta inflacionada (em 1)	BEZI	(0, 1]	logit	log	logit	-
4	beta inflacionada (em 1)	BEINF1	(0, 1]	logit	logit	log	-
5	beta inflacionada (em 0 e 1)	BEINF	[0, 1]	logit	logit	log	log
6	GA zero-ajustada	ZAGA	[0, ∞)	log	log	logit	-
7	IG zero-ajustada	ZAIG	[0, ∞)	log	log	logit	-

Fonte: Adaptado de Rigby et al. (2019)

Outrossim, há ainda a possibilidade de se implementar novas distribuições caso necessário (Roquim et al., 2021), além de ser possível construir novas distribuições a partir de transformações daquelas já implementadas no pacote `gamlss.family` (Stasinopoulos et al., 2017). As funções disponíveis são apresentadas a seguir:

- a) `gen.Family()`: transformação logarítmica ou do tipo logit em qualquer distribuição que assuma que a variável resposta esteja no intervalo $-\infty < Y < \infty$;
- b) `gen.trun()`: função pertencente ao pacote `gamlss.tr` que realiza o truncamento de qualquer distribuição;
- c) `gen.cens()`: função pertencente ao pacote `gamlss.cens` que constrói versões censuradas de distribuições no intervalo $(0, \infty)$;
- d) `gamlssZadj()`: função pertencente ao pacote `gamlss.inf` que constrói versões ajustadas em zero de distribuições no intervalo $(0, \infty)$;
- e) `gen.Inf0to1()`: função pertencente ao pacote `gamlss.inf` que constrói versões inflacionadas em zero e/ou um de distribuições no intervalo $(0, 1)$.

Mais informações sobre as distribuições podem ser encontradas em Rigby et al. (2019).

2.1.4.1.1 Distribuições zero-ajustadas

Variáveis resposta que assumem valor zero ou uma quantidade positiva contínua, isto é, $0 \leq Y < \infty$, são comumente encontradas em diversas áreas de aplicação e podem ser representadas, por exemplo, pelas distribuições ZAGA e ZAIG (Tabela 2.3), que são distribuições zero-ajustadas (também denominadas como distribuições semicontínuas ou com peso zero).

As abordagens de modelagem para tais variáveis compreendem modelos de duas partes, nos quais a probabilidade zero e o componente contínuo são modelados separadamente, que é a abordagem que adotamos a seguir; e modelos Tobit (Tobin, 1958), que assumem uma distribuição qualquer subjacente para a resposta, que é censurada à esquerda em zero (Liu et al., 2019; Neelon; O'Malley; Smith, 2016). Esta segunda abordagem é mais restritiva, uma vez que a probabilidade de ocorrência de zero não pode ser modelada explicitamente.

Nesta estrutura, a distribuição $Y \sim \mathcal{D}(\boldsymbol{\theta}_k, \pi_0)$ de uma resposta zero-ajustada assume a forma

$$f_Y(y|\boldsymbol{\theta}_k, \pi_0) = \begin{cases} \pi_0 & \text{se } y = 0 \\ (1 - \pi_0)f_{Y_i}(y|\boldsymbol{\theta}_k) & \text{se } y > 0 \end{cases} \quad (2.12)$$

em que $\pi_0 = \mathbf{P}(Y = 0)$, $0 < \pi_0 < 1$, é a probabilidade do ponto em zero e $f_{Y_i}(y|\boldsymbol{\theta}_k)$ é qualquer função densidade de probabilidade com suporte na linha real positiva, com parâmetros $\boldsymbol{\theta}_k^\top = (\theta_1, \dots, \theta_p)$. O logaritmo da função de verossimilhança, neste caso, pode ser dividida em dois componentes: (i) um modelo binário, no qual a probabilidade de ocorrência do valor zero é modelado explicitamente a partir das variáveis explicativas disponíveis; e (ii) modelagem usual com base na estrutura dos GAMLSS para respostas positivas (Stasinopoulos et al., 2023). Desta maneira, Stasinopoulos et al. (2023) reescrevem um GAMLSS baseado em uma distribuição zero-ajustada como

$$\mathbf{y} \stackrel{\text{ind}}{\sim} \mathcal{D}(\boldsymbol{\pi}_0, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p) \quad (2.13)$$

$$g_\pi(\boldsymbol{\pi}_0) = \mathbf{X}_\pi \boldsymbol{\beta}_\pi + s_{\pi 1}(\mathbf{x}_{\pi 1}) + \dots + s_{\pi J_\pi}(\mathbf{x}_{\pi J_\pi}) \quad (2.14)$$

$$g_k(\boldsymbol{\theta}_k) = \mathbf{X}_k \boldsymbol{\beta}_k + s_{k 1}(\mathbf{x}_{k 1}) + \dots + s_{k J_k}(\mathbf{x}_{k J_k}) \quad (2.15)$$

em que, a equação (2.14) refere-se ao preditor do modelo para a probabilidade π_0 de ocorrência de zero.

2.1.4.2 Método de estimação dos parâmetros

Segundo Rigby et al. (2019), o processo de estimação dos parâmetros dos GAMLSS em sua versão linear paramétrica consiste em utilizar o método de máxima verossimilhança, isto é, maximizamos a função de máxima verossimilhança, dada por

$$\ell = \sum_{i=1}^n \log f(y_i | \boldsymbol{\mu}_i, \boldsymbol{\sigma}_i, \mathbf{v}_i, \boldsymbol{\tau}_i), \quad (2.16)$$

em que $f(\cdot)$ representa a função (densidade) de probabilidade da variável de resposta.

Para modelos que apresentam termos de suavização não paramétricos, como, por exemplo, os *P-splines*, é necessário recorrer ao método da máxima verossimilhança penalizada, que

consiste na maximização da função

$$\ell_p = \ell - \frac{1}{2} \sum_{k=1}^p \sum_{j=1}^{J_k} \lambda_{kj} \gamma'_{kj} \mathbf{P}_{kj} \gamma_{kj}. \quad (2.17)$$

Rigby e Stasinopoulos (2005) sugerem dois algoritmos para estimar um GAMLSS para valores fixos de λ , a fim de maximizar a função de verossimilhança penalizada, denominados algoritmos CG e RS. O CG, uma generalização do algoritmo proposto por Cole e Green (1992), requer informações sobre a primeira e segunda derivadas cruzadas do logaritmo da função de verossimilhança em relação aos parâmetros de distribuição $\boldsymbol{\theta}_k$; ao passo que o RS, uma generalização do algoritmo proposto por Rigby e Stasinopoulos (1996), não utiliza as derivadas cruzadas do logaritmo da função de verossimilhança. Uma terceira opção, composta por passos iniciais de RS e, posteriormente, por passos de CG, também é bastante utilizada (Stasinopoulos et al., 2017).

2.1.4.3 Seleção das variáveis

O processo de seleção de variáveis a serem introduzidas em cada uma das estruturas de regressão dos parâmetros $\boldsymbol{\theta}_k$ pode ser realizado de diferentes maneiras (Stasinopoulos et al., 2017). Segundo Ramires et al. (2021), a Estratégia A, implementada no pacote `gamlss` por meio da função `stepGAICAll.A()`, é a mais utilizada. Ela consiste em uma abordagem de seleção de variáveis com base no procedimento *stepwise*.

Considere uma distribuição com quatro parâmetros, isto é, $\boldsymbol{\theta}_k = (\mu, \sigma, \nu, \tau)^\top$, conforme Nakamura et al. (2017), a Estratégia A funciona da seguinte maneira:

- a) ajusta-se um modelo para μ por meio do procedimento *forward* baseado em algum critério de seleção, por exemplo, o critério de informação de Akaike (Akaike, 1974), considerando σ , ν e τ como constantes;
- b) usando o mesmo procedimento, o modelo para o parâmetro σ é ajustado, considerando o modelo ajustado para o parâmetro μ no passo (1) e os parâmetros restantes como constantes;
- c) o modelo para o parâmetro ν é ajustado utilizando um procedimento *forward*, considerando os modelos de μ e σ obtidos nos passos (1) e (2) e τ como fixo;

- d) o modelo para o parâmetro τ é ajustado por meio de um procedimento *forward* considerando os três modelos ajustados nos passos (1), (2) e (3);
- e) reajusta-se o parâmetro ν , pelo procedimento *backward*, dado todas as demais estruturas de regressão já ajustados;
- f) um procedimento *backward* é realizado para o modelo do parâmetro σ considerando todos os outros modelos ajustados;
- g) finalmente, um último procedimento *backward* é realizado para o parâmetro μ , considerando-se os modelos para σ , ν e τ .

Após os sete passos, o modelo final irá conter covariáveis para cada parâmetro da distribuição. Vale ressaltar que as estruturas construídas não serão necessariamente iguais.

2.2 Diagnóstico do modelo

Este processo é uma etapa fundamental em qualquer modelo de regressão. Em todos modelos de regressão, a análise de resíduos é uma prática indispensável. Nos GAMLSS, usualmente são utilizados os resíduos quantílicos (aleatorizados) normalizados (Dunn; Smyth, 1996), uma vez que, independentemente da distribuição assumida para a variável resposta, os resíduos sempre têm uma distribuição normal padrão se o modelo assumido é adequado aos dados. Estes resíduos são obtidos pela expressão

$$\hat{r}_i = \Phi^{-1}(\hat{u}_i), \quad (2.18)$$

em que $\Phi^{-1}(\cdot)$ é o inverso da função distribuição acumulada da normal padrão, e os \hat{u}_i são denominados resíduos quantílicos. Esses resíduos são obtidos de maneira distinta para variáveis contínuas e discretas. Para verificar as pressuposições do modelo baseados nos resíduos, usualmente utilizamos métodos gráficos, como, por exemplo, o *worm plot* (Buuren; Fredriks, 2001).

O *worm plot* é um gráfico eficaz para detectar áreas em que o modelo não está bem ajustado aos dados. Graficamente, os resíduos formam uma curva semelhante a uma minhoca, onde é possível identificar possíveis problemas nos resíduos e, conseqüentemente, no modelo ajustado, conforme mostra a Quadro 2.4.

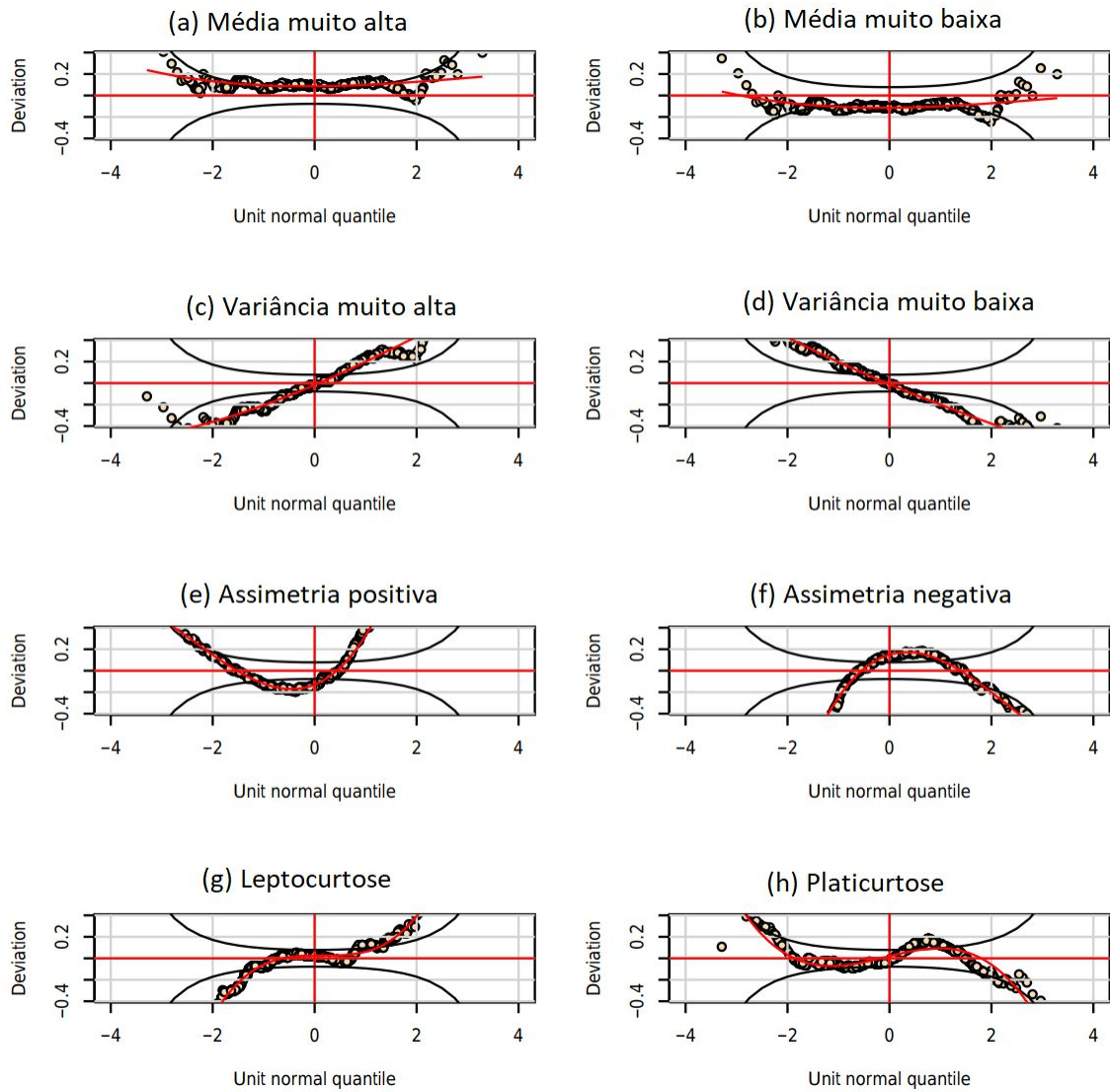
Quadro 2.4 – Diferentes formatos de *worm plot* e interpretações

Formato	Resíduos	Parâmetros ajustado
Nível: acima da origem	Média muito alta	Locação subestimada
Nível: abaixo da origem	Média muito baixa	Locação superestimada
Reta: inclinação positiva	Variância muito alta	Escala subestimada
Reta: inclinação negativa	Variância muito baixa	Escala superestimada
U	Assimetria positiva	Assimetria subestimada
U invertido	Assimetria negativa	Assimetria superestimada
S com curva esquerda pra baixo	Leptocurtose	Cauda muito leve
S com curva esquerda pra cima	Platicurtose	Cauda muito pesada

Fonte: Adaptado de Stasinopoulos et al. (2017)

Segundo Stasinopoulos et al. (2017), a Figura 2.1 apresenta os diferentes *worm plots*, que evidenciam problemas em parâmetros como variância, assimetria e curtose dos resíduos. Além disso, são relatados problemas relacionados à distribuição ajustada. O formato de cada *worm plot* correspondente da Quadro 2.4 pode ser caracterizado da seguinte maneira, conforme mostrado na Figura 2.1, o nível acima da origem corresponde ao worm plot do Painel (a); o nível abaixo da origem do Painel (b); a reta com inclinação positiva é representada pelo Painel (c); a reta com inclinação negativa é mostrada no Painel (d); a forma em U corresponde ao Painel (f); a U invertida é a do Painel (e); a forma em S com curva esquerda para baixo é vista no Painel (g); e, finalmente, a forma em S com curva esquerda para cima corresponde ao Painel (h).

Figura 2.1 – Diferentes tipos de falhas do modelo indicadas pelo *worm plot*: (a) e (b) indicam falha no ajuste correto da locação; (c) e (d) indicam falha no ajuste correto da escala; (e) e (f) indicam falha no ajuste correto da assimetria; e (g) e (h) indicam falha no ajuste correto da curtose.



Fonte: Stasinopoulos et al. (2017)

3 METODOLOGIA

Para o desenvolvimento dessa pesquisa foram estipuladas duas etapas. A primeira etapa adotou uma abordagem quali-quantitativa, baseada na busca e análise de documentos disponíveis na literatura sobre modelos aditivos generalizados para locação, escala e forma (GAMLSS), complementada por uma busca exploratória. As principais bases de dados utilizadas foram: Web of Science e Scopus (Elsevier), com acesso disponibilizado por meio do portal Periódicos CAPES (CAPES et al., 2024).

O *software* Zotero foi empregado para o gerenciamento de referências, organização de referências e armazenamento de bancos de dados de artigos relacionados ao tópico em estudo. Este *software* de código aberto destaca-se pela precisão na geração de bibliografias e por sua compatibilidade com formatos como RIS, BibTeX e Zotero RDF (Ivey; Crum, 2018; Ahmed; Dhubaib, 2011). Além disso, o Zotero foi utilizado para examinar manualmente a qualidade das publicações, verificando a relevância das palavras-chave selecionadas, com base no protocolo de pesquisa por Ferenhof e Fernandes (2016).

Entre diversas ferramentas de *software* voltadas para mapeamento científico e suporte à análise bibliométrica, como Bibliometrix (Aria; Cuccurullo, 2017), CitNetExplorer (Eck; Waltman, 2014), VOSviewer (Eck; Waltman, 2010), SciMAT (Cobo et al., 2012), BibExcel (Person; Danell; Schneider, 2009), Science of Science (Sci2) Tool (Team, 2009) e CiteSpace (Chen, 2006), foi utilizada a ferramenta Bibliometrix, disponibilizada como um pacote no *software* R (R CORE TEAM, 2024). O Bibliometrix oferece um conjunto robusto de ferramentas para pesquisa quantitativa em bibliometria e cienciometria (Aria; Cuccurullo, 2017; Cuccurullo; Aria; Sarto, 2016; Donthu et al., 2021). Desenvolvido na linguagem R, que é interativa e de código aberto, o pacote conta com suporte adicional da interface gráfica *Biblioshiny*, facilitando análises complexas. Maiores detalhes podem ser encontrados no Artigo 1 desta dissertação, intitulado: “*Historical developments and research trends in GAMLSS*”.

A segunda etapa da pesquisa envolveu a construção de distribuições contínuas de probabilidade adicional de não-ocorrência do evento de interesse, utilizando os GAMLSS. Foram exploradas distribuições pertencentes à família Box-Cox (Nakamura et al., 2022a), com foco nas distribuições Box-Cox Cole e Green (BCCG) (Cole; Green, 1992) e Box-Cox Exponencial Potência (BCPE) (Rigby; Stasinopoulos, 2004). Com o objetivo de modelar probabilidade de zeros e a componente contínua, foram propostas distribuições derivadas das BCCG e BCPE: BCCG zero-ajustada (zBCCG) e BCPE zero-ajustada (zBCPE). Essas distribuições foram de-

envolvidas a partir da base de transformação para distribuições zero-ajustadas, conforme proposto por Stasinopoulos et al. (2023), e a avaliação das novas distribuições foi realizada por meio de simulações em diferentes cenários estatísticos, além de uma aplicação a dados de seguros agrícolas.

3.1 Estudo de caso: seguros agrícolas

A agricultura é amplamente reconhecida como um motor crucial para o desenvolvimento econômico e o bem-estar de grande parte da população mundial, especialmente nas economias em desenvolvimento. Segundo o Banco Mundial, 43% da população mundial vive em zonas rurais. No Brasil, essa cifra é de 12%. Além disso, 70% das pessoas em situação de vulnerabilidade socioeconômica vivem em zonas rurais, onde a agricultura é a principal fonte de renda e emprego (Bank, 2024).

A atividade agrícola depende de condições climáticas incontrolláveis que afetam a vida dos agricultores, especialmente nos países em desenvolvimento, onde os métodos tradicionais, como o uso de ferramentas manuais e a falta de tecnologia avançada, ainda predominam. Essas condições climáticas adversas podem ser mitigadas por meio de métodos de gestão agrícola e transferência de riscos, como o seguro agrícola (Bilgin et al., 2022). Esse seguro é considerado um dos principais instrumentos de gestão de risco agrícola e um motor significativo do desenvolvimento agrícola mundial (Zou et al., 2022). No entanto, o seguro agrícola raramente está disponível fora das ofertas governamentais subsidiadas, que geralmente alcançam apenas uma pequena fração dos pequenos agricultores.

O seguro agrícola é um serviço que protege o produtor rural contra perdas causadas por desastres climáticos e outras adversidades (Gasparetto; Pozzo; Ozaki, 2023). Ele permite que os agricultores diversifiquem riscos que anteriormente não conseguiam gerir completamente, afetando diretamente seu rendimento e suas decisões de produção. Entretanto, a disponibilidade e eficácia do seguro agrícola variam globalmente, dependendo de fatores como o nível de desenvolvimento, a estrutura agrícola, as condições climáticas e políticas agrícolas (Kutlar; Akcaoz, 2021; Bilgin et al., 2022).

Formalmente, a incerteza para os agricultores pode ser classificada em quatro principais fontes: (i) Produção - incerteza na quantidade e qualidade da produção devido aos longos ciclos de produção agrícola; (ii) Preço - volatilidade dos preços agravada por características dos mercados agrícolas; (iii) Tecnologia - novas técnicas de produção que podem tornar obsoletos os

investimentos anteriores; (iv) Política - riscos associados a mudanças nas políticas governamentais, como subsídios e gestão de terras (Moschini; Hennessy, 2001; Türkten; Yıldırım; Ceyhan, 2023).

A gestão de risco agrícola em várias partes do mundo, como EUA, Gana, Etiópia, China, e Turquia, inclui estratégias de transferência de risco e seguros agrícolas baseados em índices que protegem contra perdas profundas. Além disso, essas formas de proteção são de fáceis compreensão e exigem pouco controle direto por parte do produtor (Asseldonk et al., 2019; Arias et al., 2018).

Deve-se perceber que, o seguro agrícola é especialmente relevante para pesquisadores e formuladores de políticas públicas e/ou privadas que buscam reduzir a vulnerabilidade e promover a produtividade entre as populações rurais pobres. A disponibilidade desse seguro pode atrair crédito, uma vez que famílias seguradas representam menor risco para credores. Além disso, combinar seguros com investimentos produtivos pode melhorar significativamente o bem-estar das famílias rurais (Lom; Caprian, 2023; Carter; Cheng; Sarris, 2016).

Apesar dos benefícios, existem desafios, como as altas taxas de inflação em países em desenvolvimento (Percy; Smith, 2015) e a redução de insumos agrícolas (Bulte; Lensink, 2023). Contudo, o impacto positivo do seguro agrícola inclui a melhoria da produtividade agrícola e a redução das disparidades de renda entre áreas urbanas e rurais (Wen et al., 2023).

No Brasil, o governo tem implementado programas para incentivar a adoção do seguro agrícola, como o Programa de Subvenção ao Seguro Rural (PSR), instituído pela Lei 10.823 de 19 de dezembro de 2003 e pelo Decreto 5.121 de 30 de junho de 2004, visando reduzir custos para produtores e aumentar a utilização desse instrumento de gestão de riscos (Carrer et al., 2020; Leite; Castro; Tateishi, 2023).

Existem quatro modalidades de seguro agrícola no Brasil: custeio agrícola, faturamento ou receita, produção e índices climáticos. Além disso, há dois tipos de contrato: multirrisco e risco nomeado (Gasparetto; Pozzo; Ozaki, 2023; Carrer et al., 2020). Dados do Ministério da Agricultura, Pecuária e Abastecimento - MAPA, 2024 mostram que a soja lidera o ranking de apólices históricas, com 732,61 mil (43,9%), seguida pelo milho da segunda safra, com 227,56 mil (13,5%), e pela uva, com 145,01 mil (8,6%). A quantidade de produtores e apólices emitidas é liderada pela Brasilseg, com 126,01 mil (23,7%) e 587,35 mil (34,9%), respectivamente (MAPA, 2024). Maiores detalhes podem ser encontrados no Artigo 2 desta dissertação, intitulado: Novas distribuições zero-ajustadas baseada na família Box-Cox.

SEGUNDA PARTE - ARTIGOS

ARTIGO 1 - Historical developments and research trends in GAMLSS

Historical developments and research trends in GAMLSS

Elias M. Sabe¹, Robert A. Rigby², D. Mikis Stasinopoulos², Andréa C. Konrath³, Luiz R. Nakamura⁴

¹Programa de Pós Graduação em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras, Lavras, Brazil.

²School of Computing & Mathematical Science, University of Greenwich, London, UK.

³Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

⁴Departamento de Estatística, Universidade Federal de Lavras, Lavras, Brazil.

ABSTRACT

This paper is a historical study starting on the first peer-reviewed publication on generalised additive models for location, scale and shape (GAMLSS), exploring its evolution and impact through a bibliometric analysis. GAMLSS emerged from the tradition of the statistical modelling community, integrating features from linear models (LM), generalised linear models (GLM), and generalised additive models (GAM). The PRISMA protocol was used to select 593 papers published between 2004 and 2023 from the Web of Science and Scopus databases. The results revealed an average annual publication growth rate of 26.87% and an average of 25.86 citations per article, with the United States, Brazil, and China leading in the number of publications. Furthermore, the most cited article is the seminal work by Rigby and Stasinopoulos (2005), with 2,024 citations and an average of 102.20 citations per year. The collaboration network highlights Stasinopoulos, Rigby, Ramires, Villarini, Kneib, Singh, and Wang as influential authors. Finally, in 2023, the most popular research terms included machine learning, drought propagation, newborns, COVID-19, growth charts, weight, and height.

Keywords: Bibliometric analysis; Distributional regression models; Statistical modelling.

1 Introduction

In 2024, we celebrate the 20th anniversary of the publication of the first peer-reviewed paper on GAMLSS in a scientific journal. GAMLSS created from the tradition of the statistical modelling community where stochastic models are used to explain data and where models are not the aim but the tool of the analysis. Within the statistical modelling community the idea of a “true” model used in the classical statistical methodology was challenged. This was captured perfectly by George Box’s phrase “all models are wrong but some are useful” Box (1979). GAMLSS followed the regression model tradition of the Generalized Linear Model (GLM), Nelder and Wedderburn (1972), and Generalized Additive models (GAM), Hastie and Tibshirani (1990). GLM’s introduced non-normal errors for the response variable while GAM introduced not linear “smooth” relationships in the explanatory variables. GAMLSS combines those approaches by trying to alleviate some of the restrictions related to the assumptions of the GLM and GAM models. It was soon realised that they also introduced a new way of thinking for regression analysis, focusing not only on the mean but also on other parts of the distribution of the response. GAMLSS had become a major player in “Beyond mean regression”, Kneib (2013), in the field which is now known as “Distributional Regression”. The original definition of a GAMLSS model is:

$$y_i \stackrel{ind}{\sim} \mathcal{D}(\theta_{1i}, \dots, \theta_{ki}) \quad (1)$$

$$g_k(\theta_{ki}) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} s_{jk}(x_{jk}),$$

where y_i for $i = 1, \dots, n$, is the response variable, $\mathcal{D}(\theta_{1i}, \dots, \theta_{ki})$ is a theoretical assumed distribution with K parameters $\theta_1, \dots, \theta_k$, which are all modelled, if needed, as functions of the explanatory variables. The functions g_1, \dots, g_k are *link* functions to ensure that the parameters of distributions are defined in the right range. For example if the range of the parameter θ_i is $(0, \infty)$ then the $\log(\theta_i)$ link will guaranty that the parameter is always positive. The functions $s(\cdot)$ are non-parametric smooth to account for possible non-linear relationships between the explanatory variables and the parameters of the distribution. If there are no smooth functions (and only parametric models) the GAMLSS model can be fitted using maximum likelihood estimation (or Bayesian methodology). If smothers exist then penalised likelihood can be used

for the fitting the model. Gradient boosting is available in GAMLSS and it is appropriate for the selection feature (explanatory variables). The penalised likelihood, the MCMC methodology for fitting a Bayesian GAMLSS and boosting are described in Stasinopoulos et al. (2024). The algorithm to fit the GAMLSS model using penalised likelihood, which is described in detail in Chapter 3 of Stasinopoulos et al. (2017), can be amended to enhance the model to;

$$g_1(\theta_{1i}) = ML_1(\mathbf{x}) \quad (3)$$

$$\dots = \dots$$

$$g_k(\theta_{ki}) = ML_K(\mathbf{x}) \quad (4)$$

where $ML_k(\mathbf{x}$ for $k = 1, \dots, K$ can be any sensible machine learning algorithm which can take prior weights. The enhance model of equation (4) could possibly cope better with interaction between the explanatory variables \mathbf{x} than the model in (3) where the interaction have to declared explicitly. GAMLSS has been used in a variety of applications where the focus is not the mean (or location) but other parts of the distribution as we will see in this article.

To celebrate the 20th anniversary of the first peer review paper of GAMLSS, this article, conducts a bibliometric analysis (Donthu et al., 2021) of GAMLSS over its 20-year history, following the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (Page, et al., 2021), supplemented by exploratory searches. This paper has the following objectives: i) analyse the growth of the literature on GAMLSS in scientific papers; ii) identify the main categories within the Web of Science and Scopus databases in which publications can be classified, along with the key terms prevalent within them; iii) identify the leading authors in the research field; iv) identify the main sources of publication; and v) determine networking and collaboration among authors and countries.

2 Materials and Methods

The subsequent search strategy included querying the following databases: Web of Science, and Scopus, both accessible through the *Periódico CAPES* platform, Brazil (CAPES, 2024). The search approach employed the terms (“GAMLSS” OR “Generalized Additive Models for Location, Scale* and Shape”), with the following inclusion and exclusion criteria: (i) document must be a peer-reviewed scientific article; (ii) full text must be in English-language;

(iii) abstract and full text must be available; and (iv) search period extends from 2004 to 2023. It is important to highlight that some articles in our list following this protocol appear with the year 2024; however, they were included in the analysis since the search strategy identified them as having been published online in 2023.

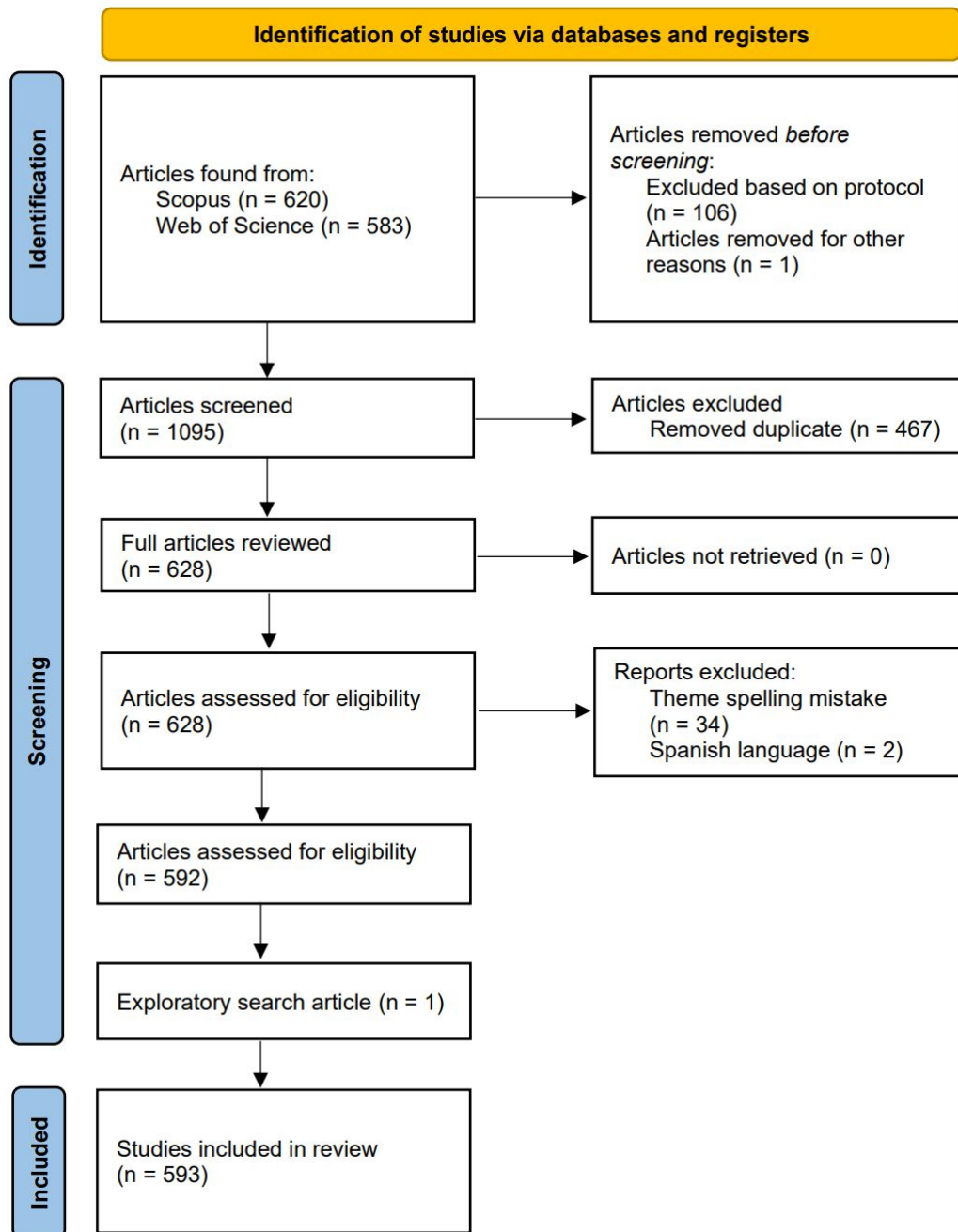
In total, 1203 articles were obtained from the Web of Science and Scopus databases during the initial search using the above mentioned protocol. The retrieved articles were exported in RIS (Research Information Systems) format and then imported into the Zotero[®] software (Ahmed; Al Dhubaib, 2011) for manual validation (by analysing titles and abstracts to determine the contextual relevance of chosen keywords), and deemed satisfactory. Following verification, the documents were imported into the R programming language (R CORE TEAM, 2024) for data conversion into BibTex and subsequent merging of databases. Continuing, articles from the databases were compared using the bibliometrix package (Aria; Cuccurullo, 2017) in R. Consequently, 467 duplicate articles between the databases were eliminated. Furthermore, 34 articles were removed due to spelling errors in their titles, while two articles were excluded since their full text were in Spanish. At the end of this process, 592 articles remained to compose the portfolio. Additionally, the publication by Rigby and Stasinopoulos (2004) was included through an exploratory search, as can be seen in Figure 3.1.

The 593 articles were organised alphabetically to facilitate the construction of a knowledge matrix in accordance with the established research protocol and the use of the systematic search flow (SSF) method (Ferenhof; Fernandes, 2016). This matrix serves as an aid for interpreting and writing the integrative review (Klopper et al., 2007). The full list of all articles, may be accessed in the Supplementary Material 1. It is vital to remember that different research protocols may result in varying numbers of articles and could influence the subsequent analyses.

3 Results and discussion

The bibliometric analysis began with the 593 articles selected for the final bibliographic portfolio composition. The chosen articles have an average of 25.86 citations per document. Furthermore, 340 different journals were identified, with an average of 6.04 authors per manuscript, 15 of whom were single-authored. Figure 3.2 displays the temporal evolution of the number of publications, which increased at an average yearly rate of 26.87%, peaking in 2022 with 99 articles.

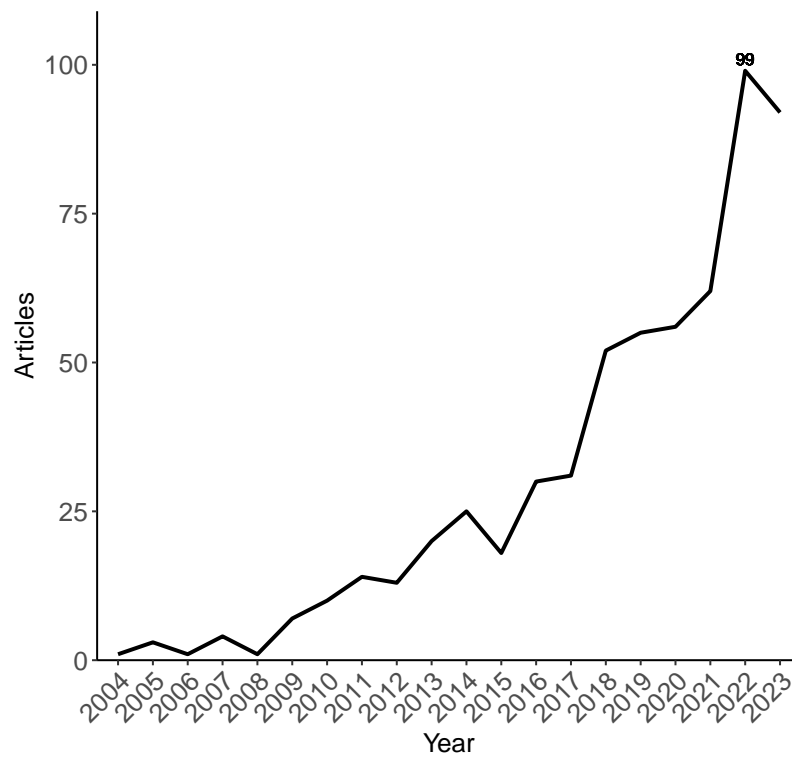
Figura 3.1 – Screening procedure for selecting GAMLSS papers



Source: Authors (2024)

Figure 3.3 shows the countries of origin for the institutions of the first authors of each article. The darker the shade of blue, the more frequently that country produces, whereas the grey colour indicates the absence of production, at least as the first author, from that country. The United States leads with 99 published papers, followed by Brazil (94), China (90), the United Kingdom (86), Germany (81), France (35), Australia (29), Italy (26), Spain (24), and Netherlands (20).

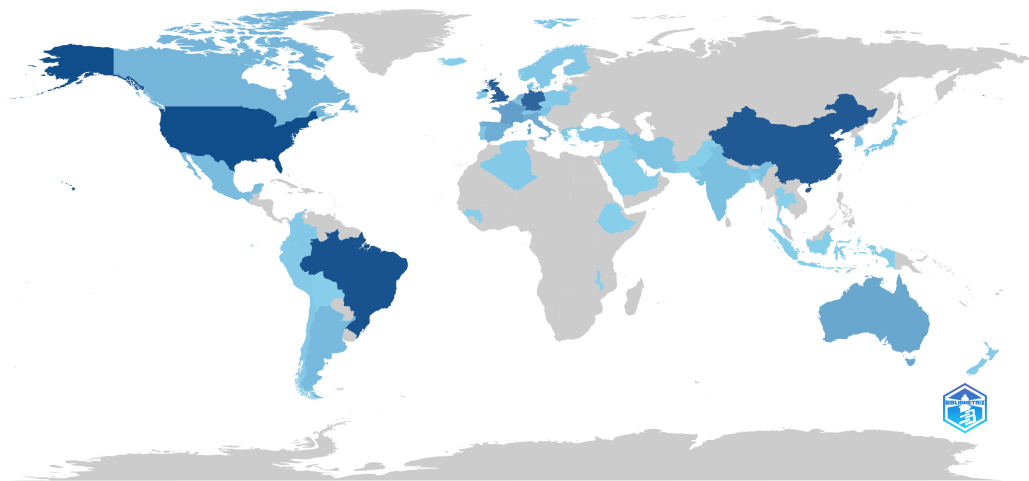
Figura 3.2 – Total number of articles published over the 20-year period.



Source: Authors (2024)

Figura 3.3 – Countries of the first authors

Country Scientific Production

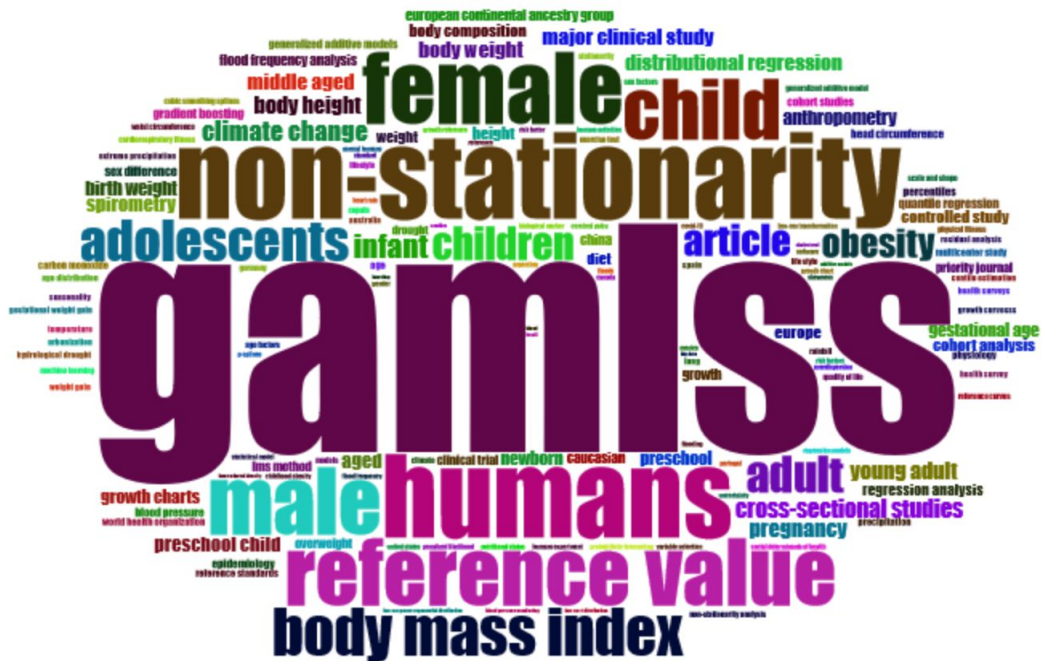


Source: Authors (2024)

Figure 3.4 shows the word cloud of author’s keywords collected from the selected papers. As expected, the term *gamlss* leads, appearing 235 times. *Non-stationarity* is the second

most frequent term with 89 occurrences, followed by *humans* (86), *female* (74), *reference value* (69), *male* (68), *child* (62), *body mass index* (49), *adolescents* (43), and *adult* (37).

Figura 3.4 – Word cloud of author’s keywords in the papers



Source: Authors (2024)

Table 3.1 lists the top ten most referenced papers in GAMLSS history. The most cited article is Rigby and Stasinopoulos’ original classic work (Rigby; Stasinopoulos, 2005), which has received 2,024 citations, with an average of 101.2 per year. Following that, Stasinopoulos and Rigby (2007), a paper describing the GAMLSS framework in R software, which received 1,010 citations. Wood et al. (2016) come in third with a smoothing parameter estimation and model selection method that can be applied to GAMLSS, with 694 citations.

Additionally, Villarini et al. (2009a) and Villarini et al. (2009b), cited 391 and 348 times respectively, present applications of GAMLSS in flood frequency analysis. Saari et al. (2011), with 344 citations, developed new Finnish reference curves (length/height-for-age, weight-for-length/height, and body mass index-for-age) for children and adolescents based on the GAMLSS framework. The article by Stanojevic et al. (2017) on the development of Global Lung Function Initiative reference values for single-breath transfer factor of the lung for carbon monoxide using GAMLSS reached 329 citations.

Furthermore, the first article considered temporally in this paper (Rigby; Stasinopoulos, 2004), which has 236 citations, introduces a new distribution called Box-Cox Power Exponential, along with the formulation of a GAMLSS-based regression framework for it. López and

Tabela 3.1 – Top ten most cited papers throughout the years

Reference	Total citations	Total citations
		per year
Rigby and Stasinopoulos (2005)	2,024	101.20
Stasinopoulos and Rigby (2007)	1,010	56.11
Wood et al. (2016)	694	77.11
Villarini et al. (2009a)	391	24.44
Villarini et al. (2009b)	348	21.75
Saari et al. (2011)	344	24.57
Stanojevic et al. (2017)	329	41.13
Rigby and Stasinopoulos (2004)	236	11.24
López and Francés (2013)	227	18.92
Brooks et al. (2011)	199	14.21

Source: Authors (2024)

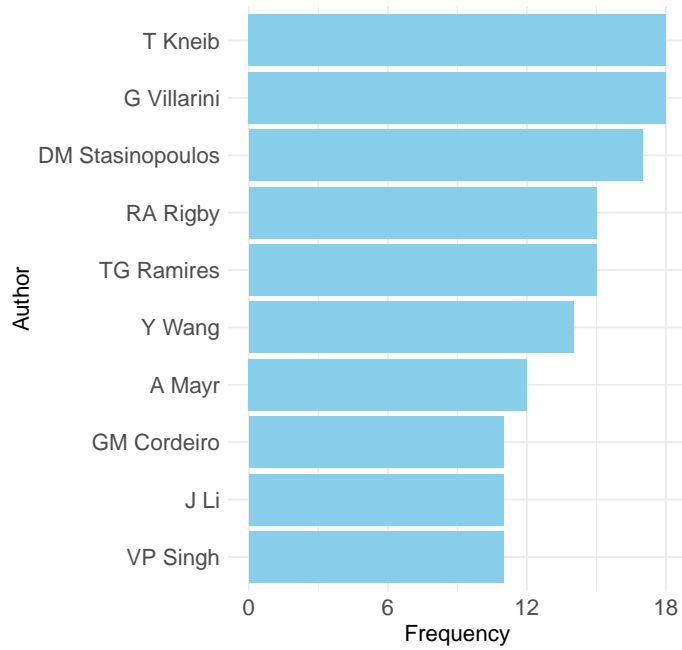
Francés (2013) also presents an application of GAMLSS in flood frequency analysis, now in Spanish rivers, and has 227 citations. The list ends with the article of Brooks et al. (2011), who used centile estimation within GAMLSS to develop reference curves of weight-for-age in cerebral palsy according to gender and Gross Motor Function Classifications System levels, with 199 citations.

Figure 3.5 presents the top ten first authors with the most publications related to GAMLSS. Authors Kneib and Villarini lead the list with 18 publications, followed by Stasinopoulos with 17, Rigby and Ramires with 15 each, Wang with 14, Mayr with 12, and Cordeiro, Li, and Singh with 11 each.

Figure 3.6 displays the top ten authors with the highest *g*-index (Egghe, 2006). Stasinopoulos leads the list with an index of 20 points, followed by Villarini and Kneib, both with 18 points, and Rigby with 15 points. The remaining authors in the list are: Wang (14), Singh (12), Mayr (12), Li (11), Cole (10), and Chen (10).

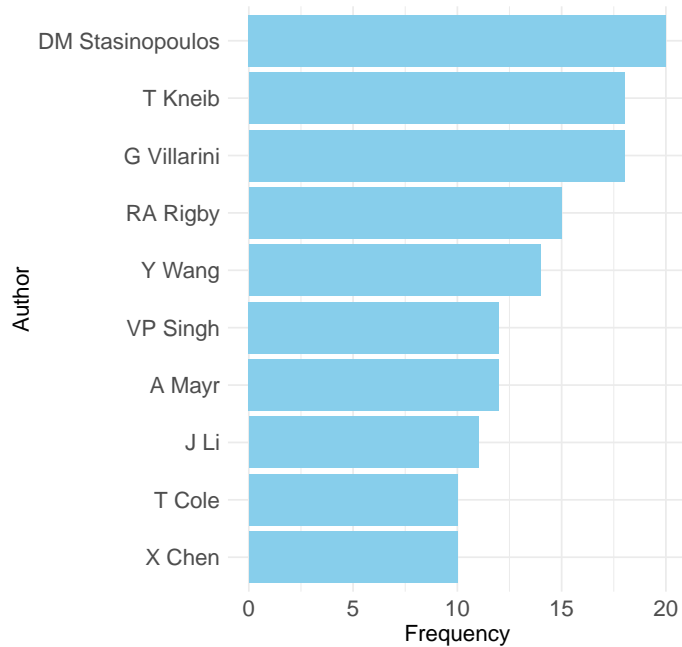
Figure 3.7 displays the collaboration network of the 30 authors with the highest number of publications in the GAMLSS field. Each node represents one of these authors, and larger nodes indicate a greater number of publications. The edges represent author collaboration, with thicker edges indicating stronger collaboration, i.e., more shared publications. The different colours represent five potential research groups, with three isolated nodes - Ahrens, Kittelson, and Cole - representing researchers with no shared publication within these groups. The size of the nodes can also help identify pivotal authors, like Stasinopoulos and Rigby (red group);

Figura 3.5 – Top ten first authors with the most publications on GAMLSS.



Source: Authors (2024)

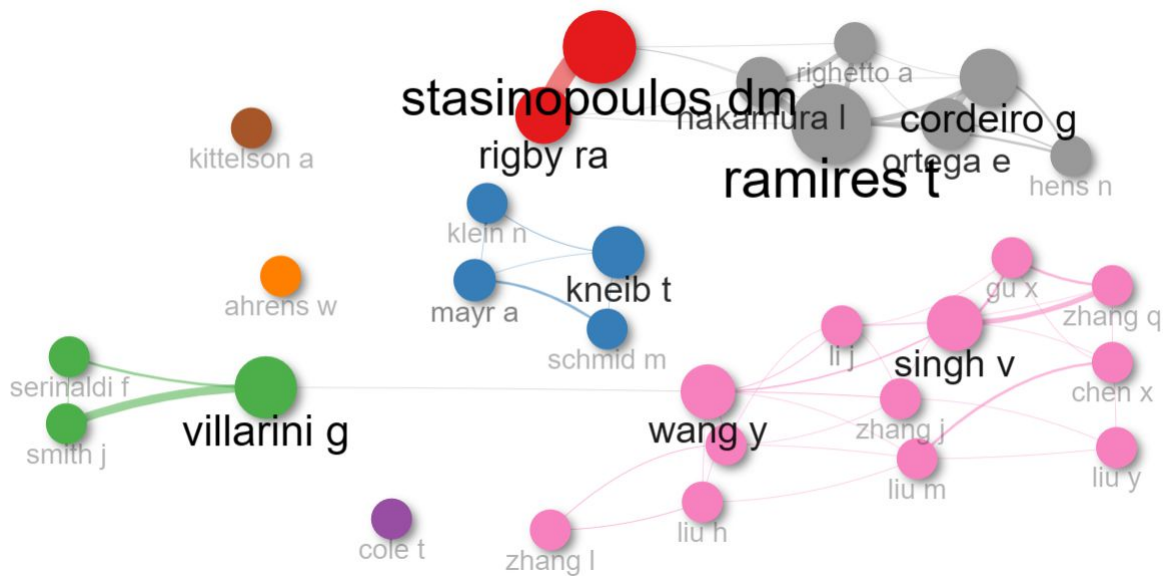
Figura 3.6 – Top ten authors with the highest g -index.



Source: Authors (2024)

Ramires and Cordeiro (grey group); Kneib (blue group); Villarini (green group); and Wang and Singh (pink group).

Figura 3.7 – Network visualization of co-authorship collaboration on GAMLSS.



Source: Authors (2024)

Obviously, publications are not limited to authors solely within each identified potential research group, meaning the potential research groups collaborate with one another. Researchers in the red group have collaborated on articles with those in the grey (Nakamura et al., 2019, 2022; Ramires et al., 2021b). Wang, from the pink group, collaborate with Villarini, from the green group, in Ren et al. (2022). Furthermore, despite there being no direct edge in Figure 3.7, one of the most recent publications regarding GAMLSS (Stasinopoulos et al., 2024), a book, includes researchers from both red and blue groups.

In Figure 3.8, we can see the top ten journals with the most articles related to GAMLSS. Leading the list is *Journal of Hydrology*, with a total of 18 publications, followed by *Statistical Modelling and Water (Switzerland)*, with 13 publications each. *Plos One* ranks third with 12 publications. Subsequently, we have *Water Resources Management* with 11 publications, *Communications in Statistics: Simulation and Computation* and *Journal of Applied Statistics* with 8 each, *Frontiers in Pediatrics* and *Statistics in Medicine* with 7 each, and *American Journal of Human Biology*, *International Journal of Climatology*, and *International Journal of Obesity* with 6 each.

Figure 3.9 shows the trend topics identified in GAMLSS applications over time. Each point on a line reflects the year in which a specific trend topic reached its peak frequency, while the intervals represent the periods during which a particular term or keyword was among the most popular.

Figura 3.8 – Top ten journals with the most publications on GAMLSS.



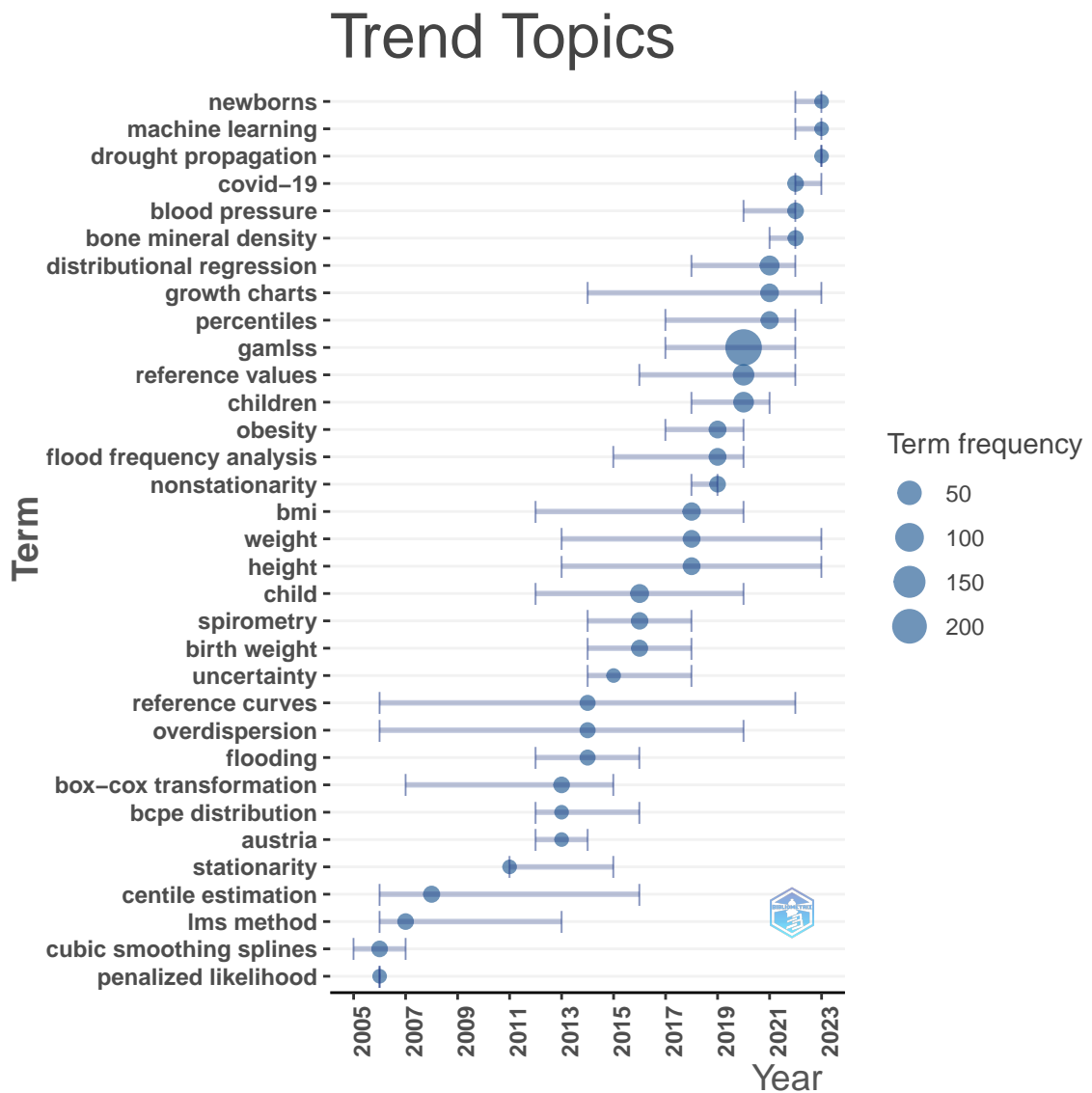
Source: Authors (2024)

It is worth noting, in particular, that the development and study of *reference curves* has been a popular topic throughout all the analysed time period, as seen in the papers by (Blagus et al., 2023; Bozek et al., 2023; Rand et al., 2023). The World Health Organization's use of GAMLSS as a foundational tool in producing growth curves (Stasinopoulos et al., 2018) may help to explain this phenomenon. A full list of the papers addressing this problem is available in Supplementary Material 2.

Another topic that remained throughout almost the entire analysed period was *overdispersion* (Ziel, 2022; Freitas et al., 2021; Marletta; Sciandra, 2020). A full list of these papers is also available in Supplementary Material 2. The explanation seems straightforward given the GAMLSS structure, which allows for the modelling of any distribution parameter, as well as the use of different distributions capable of capturing and modelling this type of problem, which is frequent in a variety of application areas Rigby et al. (2019).

Additionally, in the last year (2023), the most popular terms were *machine learning*, *drought propagation*, *newborns*, *COVID-19*, *growth charts*, *weight*, and *height* (please check

Figura 3.9 – Trend topics in the 20-year time span.



Source: Authors (2024)

Supplementary Material 2 for the full list of these papers). Machine learning has received increasing attention in the recent years, and it appears that major efforts have been made to integrate it with the GAMLSS framework, particularly in the past six years, see, for example, Bozek et al. (2023), Murray et al. (2023), and Ulrich et al. (2021). As for drought propagation (Wu et al., 2024; Jehanzaib et al., 2023; Sajeev; Kundapura, 2023; Zhang et al., 2023; Das et al., 2020). All other terms for the year 2023 may be related to the medical sciences - mainly also related to reference curves -, once again, that might be explained due to the WHO's use of GAMLSS (Aasim; Chand, 2023; Hans et al., 2023; Oliveira et al., 2023). On top of that, the predominance of COVID-19 among these trend topics is unsurprising given the recent end of a

pandemic period (Loedy et al., 2023; Reyes et al., 2022, 2023; Karim; Akter, 2022; Karim et al., 2022; Rojas et al., 2021).

Although most trending topics are related to health sciences, and a large portion of publications using GAMLSS are in this field, GAMLSS have been gaining popularity in a variety of other scientific domain, for instance, in the papers by Stasinopoulos et al. (2023), Ramires et al. (2021a), and Righetto et al. (2019).

Naturally, akin to science, there are always existing gaps that can be filled by future research. In this regard, as an example, 11 papers published - e.g. Almughlaq et al. (2023) and Reyes et al. (2022) - in the last five years specifically addressed gaps and suggested avenues for future research in their abstracts. Supplementary Material 3 provides a brief overview of each of these articles, including their gaps and future recommendations.

4 Concluding remarks

After a historical analysis of the development and application of the generalised additive models for location, scale, and shape (GAMLSS), since the first paper published back in 2004 (Rigby; Stasinopoulos, 2004), we can see that this class of models has garnered great interest and applicability in diverse fields of study. Despite its broad use, there are still gaps that can be investigated and addressed within this framework.

The importance of bibliometric analysis in the context of GAMLSS cannot be stressed, since it provides a thorough picture of research trends, major sources, and influential authors. Such analysis is a useful tool for detecting current research paths and projecting future trends in GAMLSS framework.

Furthermore, there is sufficient room for future research to improve the bibliometric mapping of GAMLSS. This can be achieved by broadening the scope to include more databases, using alternative bibliometric analysis software, and investigating different types of works, such as proceedings, review articles and books. By pursuing these paths of inquiry, researchers can contribute to a more thorough understanding of the landscape surrounding GAMLSS, allowing for their continuous improvement and implementation in theory and practice.

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. LR Nakamura was partially supported by FAPEMIG/Brazil RED-00133-21.

REFERENCES

AASIM, M.; CHAND, S. Height-for-age and weight-for-age growth charts for Pakistani infants under six months: derived from a novel case selection method using multiple indicator cluster survey data. **BMC Medical Research Methodology**, v. 23, p. 289, 2023.

AHMED, K. K. M.; AL DHUBAIB, B. E. Zotero: A bibliographic assistant to researcher. **Journal of Pharmacology and Pharmacotherapeutics**, v. 2, n. 4, p. 303-305, 2011.

ALMUGHLAQ, S.; AL-LAITH, A. A.; AL-THAWADI, S. Comparison between the World Health Organization and Bahraini children growth standards. **Revista Paulista de Pediatria**, v. 41, p. e2022050, 2023.

ARIA, M.; CUCCURULLO, C. bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959-975, 2017.

BLAGUS, R.; JURAK, G.; STARC, G.; LESKOŠEK, B. Centile reference curves of the SLOfit physical fitness tests for school-aged children and adolescents. **The Journal of Strength & Conditioning Research**, v. 37, n. 2, p. 328-336, 2023.

BOX, G. E. P. Robustness in the strategy of scientific model building. **Robustness in Statistics**, v. 1, p. 201-236, 1979.

BOZEK, J.; GRIFFANTI, L.; LAU, S.; JENKINSON, M. Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria. **Neuroimage**, v. 268, p. 119864, 2023.

BROOKS, J.; DAY, S.; SHAVELLE, R.; STRAUSS, D. Low weight, morbidity, and mortality in children with cerebral palsy: new clinical growth charts. **Pediatrics**, v. 128, n. 2, p. e299-307, 2011.

CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Periódico CAPES. 2024. URL: <<https://www.periodicos.capes.gov.br>>.

DAS, S.; DAS, J.; UMAMAHESH, N. V. Nonstationary modeling of meteorological droughts: Application to a region in India. **Journal of Hydrologic Engineering**, v. 26, n. 2, p. 0002039, 2020.

DONTHU, N.; KUMAR, S.; MUKHERJEE, D.; PANDEY, N.; LIM, W. M. How to conduct a bibliometric analysis: An overview and guidelines. **Journal of Business Research**, v. 133, p. 285-296, 2021.

EGGHE, L. Theory and practise of the g-index. **Scientometrics**, v. 69, n. 1, p. 131-152, 2006.

FERENHOF, H. A.; FERNANDES, R. F. Demystifying the literature review as basis for scientific writing: Ssf method. **Revista ACB**, v. 21, n. 3, p. 550-563, 2016.

FREITAS, L. A. C.; TACONELI, C. A.; SILVA, J. L. P.; TAMIOSO, P. R.; MOLENTO, C. F. M. A case study on animal behavior analysis using GAMLSS. **Brazilian Journal of Biometrics**, v. 39, n. 3, p. 412-433, 2021.

HANS, N.; KLEIN, N.; FASCHINGBAUER, F.; SCHNEIDER, M.; MAYR, A. Boosting distributional copula regression. **Biometrics**, v. 79, n. 3, p. 2298-2310, 2023.

HASTIE, T. J.; TIBSHIRANI, R. J. Generalized additive models. **Chapman & Hall**, London, 1990.

JEHANZAIB, M.; ALI, S.; KIM, M. J.; KIM, T. W. Modeling hydrological non-stationarity to analyze environmental impacts on drought propagation. **Atmospheric Research**, v. 286, p.

106699, 2023.

KARIM, R.; AKTER, N. Effects of climate variables on the COVID-19 mortality in Bangladesh. **Theoretical and Applied Climatology**, v. 150, p. 3-4, 2022.

KARIM, R.; AKTER, M. B.; HAQUE, S.; AKTER, N. Do temperature and humidity affect the transmission of SARS-CoV-2019 - a flexible regression analysis. **Annals of Data Science**, v. 9, p. 153-173, 2022.

KLOPPER, R.; LUBBE, S.; RUGBEER, H. The matrix method of literature review. **Alternation**, v. 14, n. 1, p. 262-276, 2007.

KNEIB, T. Beyond mean regression. **Statistical Modelling**, v. 13, n. 4, p. 275-303, 2013. ISSN 1471-082X.

LOEDY, N.; COLETTI, P.; WAMBUA, J.; HERMANS, L.; WILLEM, L.; JARVIS, C. I.; WONG, K. L. M.; EDMUNDS, W. J.; ROBERT, A.; LECLERC, Q. J.; GIMMA, A.; MOLENBERGHS, G.; BEUTELS, P.; FAES, C.; HENS, N. Longitudinal social contact data analysis: insights from 2 years of data collection in Belgium during the COVID-19 pandemic. **BMC Public Health**, v. 23, p. 1298, 2023.

LÓPEZ, J.; FRANCÉS, F. Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. **Hydrology and Earth System Sciences**, v. 17, n. 8, p. 3189-3203, 2013.

MARLETTA, A.; SCIANDRA, M. GAMLSS for high-variability data: an application to liver fibrosis case. **The International Journal of Biostatistics**, v. 16, n. 2, p. 20190113, 2020.

MURRAY, J.; AYERS, J.; BROOKFIELD, A. The impact of climate change on monthly baseflow trends across Canada. **Journal of Hydrology**, v. 618, p. 129254, 2023.

NAKAMURA, L. R.; CERQUEIRA, P. H. R.; RAMIRES, T. G.; PESCIM, R. R.; RIGBY, R. A.; STASINOPOULOS, D. M. A new continuous distribution on the unit interval applied to modelling the points ratio of football teams. **Journal of Applied Statistics**, v. 46, n. 3, p. 416-431, 2019.

NAKAMURA, L. R.; RAMIRES, T. G.; RIGHETTO, A. J.; PESCIM, R. R.; ROQUIM, F. V.; SAVIAN, T. V.; STASINOPOULOS, D. M. Cattle reference growth curves based on centile estimation: A GAMLSS approach. **Computers and Electronics in Agriculture**, v. 192, p. 106572, 2022.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society, Series A**, v. 135, p. 370-384, 1972.

OLIVEIRA, M. H.; ARAÚJO, J.; RAMOS, E.; CONDE, W. L. MULT: New height references and their efficiency in multi-ethnic populations. **American Journal of Human Biology**, v. 35, n. 5, p. e23859, 2023.

PAGE, M. J.; MOHER, D.; BOSSUYT, P. M.; BOUTRON, I.; HOFFMANN, T. C.; MULROW, C. D.; SHAMSEER, L.; TETZLAFF, J. M.; AKL, E. A.; BRENNAN, S. E.; CHOU, R.; GLANVILLE, J.; GRIMSHAW, J. M.; HRÓBJARTSSON, A.; LALU, M. M.; LI, T.; LODER, E. W.; MAYO-WILSON, E.; MCDONALD, S.; MCGUINNESS, L. A.; STEWART, L. A.; THOMAS, J.; TRICCO, A. C.; WELCH, V. A.; WHITING, P.; MCKENZIE, J. E. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. **BMJ**, v. 372, p. n160, 2021.

R CORE TEAM. R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, Vienna, Austria, 2024. URL: <<https://www.R-project.org>>.

RAMIRES, T. G.; NAKAMURA, L. R.; RIGHETTO, A. J.; CARVALHO, R. J.; VIEIRA, L. A.; PEREIRA, C. A. B. Comparison between highly complex location models and GAMLSS. **Entropy**, v. 23, n. 4, p. 469, 2021a.

RAMIRES, T. G.; NAKAMURA, L. R.; RIGHETTO, A. J.; PESCIM, R. R.; MAZUCHELI, J.; RIGBY, R. A.; STASINOPOULOS, D. M. Validation of stepwise-based procedure in GAMLSS. **Journal of Data Science**, v. 19, n. 1, p. 96-110, 2021b.

RAND, M. S.; DIEMAR, S. S.; MØLLEHAVE, L. T.; HEIDEMANN, M.; THUESEN, B. H.; PETERSEN, J. H.; JOHANNESSEN, J.; SCHOU, A. J.; WEDDERKOPP, N.; MØLGAARD, C.; JØRGENSEN, N. R. Z-scores of bone turnover markers calculated from new established sex- and age-specific reference curves are associated to future change in BMD in children and adolescents. **Bone**, v. 167, p. 116641, 2023.

REN, M.; VILLARINI, G.; PANG, B.; XU, Z.; DU, L.; WANG, Y. Hydrological response in a highly urbanized watershed in China. **Journal of Water and Climate Change**, v. 13, n. 5, p. 2171-2187, 2022.

REYES, H.; DIETHELM-VARELA, B.; MÉNDEZ, C.; REBOLLEDO-ZELADA, D.; LILLO-DAPREMONT, B.; MUÑOZ, S. R.; BUENO, S. M.; GONZÁLEZ, P. A.; KALERGIS, A. M. Contribution of two-dose vaccination toward the reduction of COVID-19 cases, ICU hospitalizations and deaths in Chile assessed through explanatory generalized additive models for location, scale, and shape. **Frontiers in Public Health**, v. 10, p. 815036, 2022.

REYES, H.; MÉNDEZ, C.; KALERGIS, A. M. Statistical explanation of the protective effect of four COVID-19 vaccine doses in the general population. **Frontiers in Public Health**, v. 11, p. 1253762, 2023.

RIGBY, R. A.; STASINOPOULOS, D. M. Smooth centile curves for skew and kurtotic data modelled using the box-cox power exponential distribution. **Statistics in Medicine**, v. 23, n. 19, p. 3053-3076, 2004.

RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 54, n. 3, p. 507-554, 2005.

RIGBY, R. A.; STASINOPOULOS, M. D.; HELLER, G. Z.; DE BASTIANI, F. **Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R**. CRC Press, 2019.

RIGHETTO, A. J.; RAMIRES, T. G.; NAKAMURA, L. R.; CASTANHO, P. L. D. B.; FAES, C.; SAVIAN, T. V. Predicting weed invasion in a sugarcane cultivar using multispectral image. **Journal of Applied Statistics**, v. 46, n. 1, p. 1-12, 2019.

ROJAS, F.; LEIVA, V.; HUERTA, M.; MARTIN-BARREIRO, C. Lot-size models with uncertain demand considering its skewness/kurtosis and stochastic programming applied to hospital pharmacy with sensor-related COVID-19 data. **Sensors**, v. 21, n. 15, p. 5198, 2021.

SAARI, A.; SANKILAMPI, U.; HANNILA, M. L.; KINIVIEMI, V.; KESSELI, K.; DUNKEL, L. New Finnish growth references for children and adolescents aged 0 to 20 years: Length/height-for-age, weight-for-length/height, and body mass index-for-age. **Annals of Medicine**, v. 43, n. 3, p. 235-248, 2011.

SAJEEV, A.; KUNDAPURA, S. Temporal assessment of meteorological drought events using stationary and nonstationary drought indices for two climate regions in India. **Journal of Hydrologic Engineering**, v. 28, n. 11, p. 6011, 2023.

STANOJEVIC, S.; GRAHAM, B. L.; COOPER, B. G.; THOMPSON, B. R.; CARTER, K. W.; FRANCIS, R. W.; HALL, G. L. Official ERS technical standards: Global lung function initiative reference values for the carbon monoxide transfer factor for Caucasians. **European Respiratory Journal**, v. 50, p. 1700010, 2017.

STASINOPOULOS, D. M.; RIGBY, R. A. Generalized additive models for location scale and shape (GAMLSS) in R. **Journal of Statistical Software**, v. 23, n. 7, p. 1-46, 2007.

STASINOPOULOS, D. M.; RIGBY, R. A.; HELLER, G. Z.; VOUDOURIS, V.; DE BASTIANI, F. **Flexible Regression and Smoothing: Using GAMLSS in R**. Chapman & Hall/CRC, Boca Raton, 2017.

STASINOPOULOS, D. M.; RIGBY, R. A.; HELLER, G. Z.; DE BASTIANI, F. P-splines and GAMLSS: a powerful combination, with an application to zero-adjusted distributions.

Statistical Modelling, v. 23, n. 5-6, p. 510-524, 2023.

STASINOPOULOS, M. D.; RIGBY, R. A.; DE BASTIANI, F. GAMLSS: A distributional regression approach. **Statistical Modelling**, v. 18, n. 3-4, p. 248-273, 2018.

STASINOPOULOS, M. D.; KNEIB, T.; KLEIN, N.; MAYR, A.; HELLER, G. Z. **Generalized Additive Models for Location, Scale and Shape: A Distributional Regression Approach, with Applications**. Cambridge University Press, 2024.

ULRICH, M.; JAHNKE, H.; LANGROCK, R.; PESCH, R.; SENGE, R. Distributional regression for demand forecasting in e-grocery. **European Journal of Operational Research**, v. 294, n. 3, p. 831-842, 2021.

VILLARINI, G.; SERINALDI, F.; SMITH, J. A.; KRAJEWSKI, W. F. On the stationarity of annual flood peaks in the continental united states during the 20th century. **Water Resources Research**, v. 45, n. 8, p. w08417, 2009a.

VILLARINI, G.; SMITH, J. A.; SERINALDI, F.; BALES, J.; BATES, P. D.; KRAJEWSKI, W. F. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. **Advances in Water Resources**, v. 32, n. 8, p. 1255-1266, 2009b.

WOOD, S. N.; PYA, N.; SÄFKEN, B. Smoothing parameter and model selection for general smooth models. **Journal of the American Statistical Association**, v. 111, n. 516, p. 1548-1563, 2016.

WU, C.; XU, Y.; JIN, J.; ZHOU, Y.; NIE, B.; LI, R.; CUI, Y.; TONG, F.; ZHANG, L. Meteorological to agricultural drought propagation time analysis and driving factors recognition considering time-variant characteristics. **Water Resources Management**, v. 38, p. 991-1010, 2024.

ZHANG, T.; SU, X.; WU, L.; CHU, J. Identification of dynamic drought propagation from a nonstationary perspective and its application to drought warnings. **Journal of Hydrology**, v. 626, p. 130372, 2023.

ZIEL, F. M5 competition uncertainty: Overdispersion, distributional forecasting, GAMLSS and beyond. **International Journal of Forecasting**, v. 38, n. 4, p. 1546-1554, 2022.

Supplementary Material 1: Knowledge matrix

ID	AUTHORS	JOURNAL	TITLE	YEAR
1	Hu Y;Duan X;Zhang Z;Lu C;Zhang Y	Children	Effects of adverse events and 12-week group step aerobics on sleep quality in Chinese adolescents	2023
2	Morkuniene R;Cole T;Jakimaviciene E;Bankauskiene A;Isakova J;Drazdiene N;Basys V;Tutkuvieni J	Frontiers in Pediatrics	Regional references vs international standards for assessing weight and length by gestational age in Lithuanian neonates	2023
3	Speller J;Staerk C;Gude F;Mayr A	Advances in Data Analysis And Classification	Robust gradient boosting for generalized additive models for location scale and shape	2023
4	Rand M;Diemar S;Møllehave L;Heidemann M;Thuesen B;Petersen J;Johannesen J;Schou A;Wedderkopp N;Mølgaard C;Jørgensen N	Bone	Z-scores of bone turnover markers calculated from new established sex-and age-specific reference curves are associated to future change in Bmd in children and adolescents	2023
5	Zhang T;Su X;Wu L;Chu J	Journal of Hydrology	Identification of dynamic drought propagation from a nonstationary perspective and its application to drought warnings	2023
6	Regis R;Ospina R;Bernardino W;Cribari-Neto F	Empirical Economics	Asset pricing in the Brazilian financial market: five-factor gamlss modeling	2023
7	Plummer J;Willmering M;Cleveland Z;Towe C;Woods J;Walkup L	Magnetic Resonance in Medicine	Childhood to adulthood: Accounting for age dependence in healthy-reference distributions in ¹²⁹ Xe gas-exchange MRI	2023
8	Lago-Deibe F;Valladares-Cabaleiro M;Fernández-Domínguez M;Fernández-Fernández I;Clavería A;Rodríguez-Pastoriza S;Roca-Pardinas J;Martín-Miguel M	Frontiers In Medicine	Effectiveness and safety of tetanus vaccine administration by intramuscular vs. subcutaneous route in anticoagulated patients: Randomized clinical trial in primary care	2022
9	Marmolejo-Ramos F;Barrera-Causil C;Kuang S;Fazlali Z;Wegener D;Kneib T;De Bastiani;Martinez-Flórez G	Cognitive Neurodynamics	Generalised exponential-Gaussian distribution: a method for neural reaction time analysis	2023
10	Cui H;Jiang S;Gao B;Ren L;Xiao W;Wang M;Ren M;Xu C	Journal Of Hydrology	On method of regional non-stationary flood frequency analysis under the influence of large reservoir group and climate change	2023
11	Heller G;Robledo K;Marschner I	BMC Medical Research Methodology	Distributional regression in clinical trials: treatment effects on parameters other than the mean	2022
12	Bröde P;Claus M;Gajewski P;Getzmann S;Wascher E;Watzl C	International Journal Of Molecular Sciences	From immunosenescence to aging types-establishing reference intervals for immune age biomarkers by centile estimation	2023
13	Restrepo-Mesa S;Rincón M;Restrepo A;Carrilho T;Kac G;Pulgarín J;Cano-Pulgarín K;Severi C;Sinisterra O;Zimmer S M;Ocampos M;Bannout M;Chico-Barba G;Arteaga N;Grandi C;Samur E;Escobar C	Plos One	Gestational weight gain charts for Latin American adolescents	2023

14	Dittrich A;Dumke M;Kapl F;Schneider P;Wege S;Gräber S;Stahl M;Herth F;Naehrlich L;Mall M;Sommburg O	Respiration International Journal of Thoracic Medicine	Survival-Adjusted FEV1 and BMI Percentiles for Patients with Cystic Fibrosis before the Era of Triple CFTR Modulator Therapy in Germany	2023
15	Kocevar G;Rioland M;Laxalde J;Mugnier A;Adib-Lesaux A;Gaillard V;Bodin J	Veterinary Research Communications	Growth charts for small sample sizes using unsupervised clustering: Application to canine early growth	2023
16	Figueroa M;Martínez J;Román E;Maury-Sintjago E;Lomaglio D;Dipierrri J	American Journal Of Human Biology	Chilean reference of length at birth	2022
17	Baur S;Kauffert J;Hewison A;Reinermann S; König A;Menzel A;Peters W	Ecology And Evolution	Spatial scaling in bed-site selection by roe deer fawns: Implications for mitigating neonatal mortality during mowing	2023
18	Lu Y;Kiechl S;Wang J;Xu Q;Kiechl S;Pechlaner R;Aguilar D;Al-Hashmi K;Alvim R;Al-Zakwani I;Antza C;Cicero A;Avramovska M;Avramovski P;Baek H;Bäck M;Bailey K;Baldo M;Batista R;Benetos A;Benjamin E;Bia D;Borghi C;Roux S;Breet Y;Burgner D;Cardoso V;Cececlja M;Ceponiene I;Chen C;Cheung M;Cheng H;Cho J;Chowienczyk P;Coelho E;Cseprekal O;Silva A;Dallaire F;Cunha R;Diaz A;Ferreira A;Ferrières J;Furuta Y;Gómez-Marcos M;Gómez-Sánchez L;Halcox J;Hanis C;Herzig K;Jaeggi E;Kavousi M;Kiechl-Kohlendorfer U;Kim H;Kim M;Kim Y;Kis E;Knoflach M;Kotsis V;Koyama T;Kozakova M;Kruger R;Kullo I;Kweon S;Lambrinoudaki I;Liu C;Loeffler M;Logan J;Maddock J;Magalhães P;Maldonado J;Mattace-Raso F;Messner A;Meyer M;Mi J;Mill J;Mitchell G;Mu J;Muhammad I;Nairz J;Nakagomi A;Nakamura M;Nilson P;Ninomiya	eBiomedicine	Global distributions of age- and sex-related arterial stiffness: systematic review and meta-analysis of 167 studies with 509,743 participants	2023
19	Chan K;Zhu H;Yu M;Yuen H;Dai S;Chin H;Choy J;Chan J;Tsoi D;Siu B;Au C;Li A	Pediatric Pulmonology	Applicability of the Global Lung Function Initiative prediction equations in Hong Kong Chinese children	2023

20	Marmolejo-Ramos F;Tejo M;Brabec M;Kuzilek J;Joksimovic S;Kovanovic V;González J;Kneib T;Bühlmann P;Kook L;Briseño-Sánchez G;Ospina R	Wiley Interdisciplinary Reviews: Data Mining And Knowledge Discovery	Distributional regression modeling via generalized additive models for location, scale, and shape: An overview through a data set from learning analytics	2023
21	Bai X;Jin J;Wu C;Zhou Y;Zhang L;Cui Y;Tong F	Water	Construction of a Time-Variant Integrated Drought Index Based on the GAMLSS Approach and Copula Function	2023
22	Jin H;Willems P;Chen X;Liu M	Journal of Hydrology	Nonstationary flood and its influencing factors analysis in the Hanjiang River Basin, China	2023
23	Rodrigues V;Martins R;De Sousa B	BMC Women's Health	Spatio-temporal trends of the age-at-menarche percentiles among Portuguese women since 1920	2023
24	Sajeev A;Kundapura S	Journal of Hydrologic Engineering	Temporal Assessment of Meteorological Drought Events Using Stationary and Nonstationary Drought Indices for Two Climate Regions in India	2023
25	Bafei S;Yang S;Chen C;Gu X;Mu J;Liu F;Sun J;Zhuang Q;Wei P;Zhao X;Chen Y;Yin Y;Xie H;Shen C	Mechanisms of Ageing and Development	Sex and age differences in the association between high sensitivity C-reactive protein and all-cause mortality: A 12-year prospective cohort study	2023
26	Tsehay B;Seyoum G	Canadian Journal of Kidney Health and Disease	Sonographic Evaluation of the Morphology of Kidneys Among Apparently Healthy Children in Northwest Ethiopia, 2021: The Normal Limits and Percentile Curves	2023
27	Lee J;Shim Y;Cho T;Kim S;Ko J;Phi J	Orphanet Journal of Rare Diseases	Growth patterns of young achondroplasia patients in Korea and predictability of neurosurgical procedures	2023
28	Rodrigues G;Ortega E;Vila R;Cordeiro G	Communications in Statistics: Simulation and Computation	A new extended normal quantile regression model: properties and applications	2023
29	Berazategui M;Raftópulos P;Farias A;Bernardi R	Forest Ecology And Management	Subregional to local topographic features shape resources land use and disturbances determining forest distribution in oldgrowth grasslands of subtropical South America	2023
30	Reyes H;Méndez C;Kalergis A	Frontiers in Public Health	Statistical explanation of the protective effect of four COVID-19 vaccine doses in the general population	2023
31	Tassi R;Seidel E;Motta-Marques D;Villanueva A;Kalin L	Water	Wildlife Roadkill Driven by Hydrological Regime in a Subtropical Wetland	2023
32	Aasim M;Chand S	BMC Medical Research Methodology	Height-for-age and weight-for-age growth charts for Pakistani infants under six months: derived from a novel case selection method using multiple indicator cluster survey data	2023
33	Eriksen B;Fasiolo M;Mathisen U;Jenssen T;Stefansson V;Melsom T	Scientific Reports	Ambulatory blood pressure as risk factor for long-term kidney function decline in the general population: a distributional regression approach	2023
34	Liu W;Yu L;Deng Q;Li Y;Lu P;Yang J;Chen F;Li F;Zhou X;Bergeron M;Ashford J;Xu Q	Alzheimer's And Dementia	Toward digitally screening and profiling AD: A GAMLSS approach of MemTrax in China	2023
35	Lolli L;Gregson W;Bonanno D;Kuitunen S;Di S V	International Journal of Sports Physiology and Performance	Age-Related Reference Intervals for Physical Performance Test Outcomes Relevant to Male Youth Middle Eastern Football Players	2023

36	Hanaoka M;Wada Y;Goto N;Kitaguchi Y;Koarai A;Kubota M;Oyamada Y;Koto H	Respiratory Investigation	Referential equations for pulmonary diffusing capacity generated from the Japanese population using the Lambda, Mu, or Sigma method and their comparisons with prior referential equations	2023
37	Wang Y;Peng T;He Y;Singh V;Lin Q;Dong X;Fan T;Liu J;Guo J;Wang G	Journal of Hydrology	Attribution analysis of non-stationary hydrological drought using the GAMLSS framework and an improved SWAT model	2023
38	Wu C;Xu Y;Jin J;Zhou Y;Nie B;Li R;Cui Y;Tong F;Zhang L	Water Resources Management	Meteorological to Agricultural Drought Propagation Time Analysis and Driving Factors Recognition Considering Time-Variant Characteristics	2023
39	De Oliveira M;Araújo J;Ramos E;Conde W	American Journal of Human Biology	MULT: New height references and their efficiency in multi-ethnic populations	2023
40	Jehanzaib M;Ali S;Kim M;Kim T	Atmospheric Research	Modeling hydrological non-stationarity to analyze environmental impacts on drought propagation	2023
41	Lauffer P;Pals G;Zwinderman A;Postema F;Baars M;Dulfer E;Hilhorst-Hofstee Y;Houweling A;Kempers M;Krapels I;Van D L I;Loeys B;Spaans A;Warnink-Kavelaars J;De W V;Wit J;Menke L	American Journal of Medical Genetics, Part A	Growth charts for Marfan syndrome in the Netherlands and analysis of genotype–phenotype relationships	2023
42	Hecht J;Barth N;Ryberg K;Gregory A	Journal of Hydrology X	Simulation experiments comparing nonstationary design-flood adjustments based on observed annual peak flows in the conterminous United States	2022
43	Murray J;Ayers J;Brookfield A	Journal of Hydrology X	The impact of climate change on monthly baseflow trends across Canada	2023
44	Genowska A;Strukcinskiene B; J;Motkowski R; Jamiolkowski J;Abramowicz P;Konstantynowicz J	Journal of Clinical Medicine	Reference Values for Birth Weight in Relation to Gestational Age in Poland and Comparison with the Global Percentile Standards	2023
45	Sonnenschein B;Ziel F	Water Resources Research	Probabilistic Intraday Wastewater Treatment Plant Inflow Forecast Utilizing Rain Forecast Data and Sewer Network Sensor Data	2023
46	Mohammadi P;Ebrahimi K;Bazrafshan J	Journal of Hydrologic Engineering	Flood Frequency Analysis under Nonstationary Modeling in the Gorganrood River Basin, Iran	2023
47	Blagus R;Jurak G;Starc G; Leskošek Ek B	Journal of Strength and Conditioning Research	Centile Reference Curves of the SLOfit Physical Fitness Tests for School-Aged Children and Adolescents	2023
48	Agudo-Domínguez A;Pérez-Blanco C;Gil-García L;Ortega J;Dasgupta S	Agricultural Water Management	Climate-sensitive hydrological drought insurance for irrigated agriculture under deep uncertainty. Insightful results from the Cega River Basin in Spain	2022
49	Desiree L;C W J;Armando P;Mario C;Patricia C	European Journal of Clinical Nutrition	Reference values for body composition in healthy urban Mexican children and adolescents	2023
50	Miranda-Soberanis V;Yee T	Computational Statistics	Two-parameter link functions, with applications to negative binomial, Weibull and quantile regression	2023

51	Eratne D;Kang M;Malpas C;Simpson-Yap S;Lewis C;Dang C;Grewal J;Coe A;Dobson H;Keem M;Chiu W;Kalincik T;Ooi S;Darby D;Brodthmann A;Hansson O;Janelidze S;Blennow K;Zetterberg H;Walker A;Dean O;Berk M;Wannan C;Pantelis C;Loi S;Walterfang M;Berkovic S;Santillo A;Velakoulis D	Australian and New Zealand Journal of Psychiatry	Plasma neurofilament light in behavioural variant frontotemporal dementia compared to mood and psychotic disorders	2023
52	Talwar P;Deantoni M;Van E M;Muto V;Chylinski D;Koshmanova E;Jaspar M;Meyer C;Degueldre C;Berthomier C;Luxen A;Salmon E;Collette F;Dijk D;Schmidt C;Phillips C;Maquet P;Sherif S;Vandewalle G	Scientific Reports	In vivo marker of brainstem myelin is associated to quantitative sleep parameters in healthy young men	2023
53	Klawitter S;Kacprowski T	Journal of Laboratory Medicine	A visualization tool for continuous reference intervals based on GAMLSS	2023
54	Morkuniene R;Tutkuvieni J;Cole T;Jakimaviciene E;Isakova J;Bankauskiene A;Drazdiene N;Basys V	Scientific Reports	Neonatal head circumference by gestation reflects adaptation to maternal body size: comparison of different standards	2022
55	Agudo-Domínguez A;Pérez-Blanco C;Gil-García L;Ortega J;Dasgupta S	Agricultural Water Management	Climate-sensitive hydrological drought insurance for irrigated agriculture under deep uncertainty. Insightful results from the Cega River Basin in Spain	2023
56	Jin H;Zhong R;Liu M;Ye C;Chen X	Stochastic Environmental Research and Risk Assessment	Nonstationary analysis of water and sediment in the Jinsha River Basin based on GAMLSS model	2023
57	Prado J;Daudt N;Perez M;Castilho P;Monteiro D	ICES Journal of Marine Science	Intensive and wide-ranging beach surveys uncover temporal and spatial stranding patterns of marine megafauna	2023
58	Ramgopal S;Sepanski R;Martin-Gill C	Annals of Emergency Medicine	Empirically Derived Age-Based Vital Signs for Children in the Out-of-Hospital Setting	2023
59	Giacomet C;Ramos A;Moura H;Berra T;Alves Y;Delpino F;Farley J;Reynolds N;Alonso J;Teibo T;Arcêncio R	Archives of Public Health	A distributional regression approach to modeling the impact of structural and intermediary social determinants on communities burdened by tuberculosis in Eastern Amazonia – Brazil	2023
60	Tosunoglu F;Slater L;Kowal K;Gu X;Yin J	Theoretical and Applied Climatology	Non-stationary modeling of seasonal precipitation series in Turkey: estimating the plausible range of seasonal extremes	2023
61	Akbar N;Aslam M;Petersen J;Mustafa G	Acta Paediatrica, International Journal of Paediatrics	The 2022 Pakistani references from birth to 60 months for length/height, weight and body mass index	2023

62	Cheung M;Cole T;Arundel P;Bridges N;Burren C;Cole T;Davies J;Hagenäs L;Högler W;Hulse A;Mason A;Mcdonnell C;Merker A;Mohnike K;Sabir A;Skae M;Rothenbuhler A;Warner J;Irving M	American Journal of Medical Genetics, Part A	Growth reference charts for children with hypochondroplasia	2023
63	Martins R;De S B;Rodrigues V	Scientific Reports	The geography of the age at menopause in central Portugal since the early twentieth century	2022
64	De Oliveira M;Araújo J;Severo M;Rodrigues K;Conde W	American Journal of Human Biology	MULT: A new BMI reference to assess nutritional status of multi-ethnic children and adolescents	2023
65	García-Peña J;Ramírez-Gil J;Góngora-Canul C;Calderón L;Cuellar J;Cruz C	Plant Pathology	Multiscale temporal and spatiotemporal analysis of wheat blast infection-like dynamics using vertical plant stratification, regression and Markov chain approaches	2023
66	Bozek J;Griffanti L;Lau S;Jenkinson M	NeuroImage	Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria	2023
67	Çakmak E;Kılıç S;Özdemir P;Karaağaoğlu E	Balkan Medical Journal	Body Mass Index Growth Curves for Birth to 24 Months Children in Ankara with RefCurv Software	2023
68	Zhan Z;Bastide-Van G S;Wiersum M;Heineman K;Hadders-Algra M;Heuvel E	Communications in Statistics: Simulation and Computation	A comparison of statistical methods for age-specific reference values of discrete scales	2023
69	Turcotte R;Boucher J	North American Actuarial Journal	GAMLSS for Longitudinal Multivariate Claim Count Models	2023
70	Zida D;Bognounou F;Sanou L;Savadogo P;Jonas K;Savadogo L	Trees, Forests and People	Transplanted seedling age and watering effects on the field performance of <i>Senegalia macrostachya</i> (Rchb. ex DC.) Kyal. & Boatwr., a high-valued indigenous fruit tree species in Burkina Faso	2023
71	Berger M;Kowark A;Rossaint R;Coburn M;Schmid M	Journal of the American Statistical Association	Modeling Postoperative Mortality in Older Patients by Boosting Discrete-Time Competing Risks Models	2023
72	Ramgopal S;Sepanski R;Crowe R;Martin-Gill C	Jacep Open	Age-based centiles for diastolic blood pressure among children in the out-of-hospital emergency setting	2023

73	Ortega F; Leskošek Ek B; Blagus R; Gil-Cosano J; Mäestu J; Tomkinson G; Ruiz J; Mäestu E; Starc G; Milanovic I; Tammelin T; Sorić M; Scheuer C; Carraro A; Kaj M; Csányi T; Sardinha L; Lenoir M; Emeljanovas A; Mieziene B; Sidossis L; Pihu M; Lovecchio N; Konstabel K; Tambalis K; Štefan L; Drenowatz C; Rubiñ L; Gontarev S; Castro-Piñero J; Vanhelst J; O'keeffe B; Veiga O; Gisladóttir T; Sandercock G; Misigoj-Durakovic M; Niessner C; Riso E; Popovic S; Kuu S; Chinapaw M; Clavel I; Labayen I; Dobosz J; Colella D; Kriemler S; Salaj S; Noriega M; Bös K; Sánchez-López M; Lakka T; Tabacchi G; Novak D; Ahrens W; Wedderkopp N; Jurak G	British Journal of Sports Medicine	European fitness landscape for children and adolescents: updated reference values, fitness maps and country rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network	2023
74	Constable P; Loh L; Prem-Senthil M; Marmolejo-Ramos F	Attention, Perception, and Psychophysics	Visual search and childhood vision impairment: A GAMLSS-oriented multiverse analysis approach	2023
75	Tu C; Tu Y; Jiang C; Zhang S	Iranian Journal of Public Health	The Establishment of the Cut-off Value of Weight-for-Height for Preschool Children in China and an Empirical Study on the Influence of Overweight and Obesity on Physical Performance	2023
76	Kittelson A; Loyd B	Archives of Physical Medicine and Rehabilitation	Personalized Reference Values for the Two-Minute Walk Test: An Analysis of Cross-Sectional Data From the National Institutes of Health Toolbox Study	2023
77	Kneib T; Silbersdorff A; Säfken B	Econometrics and Statistics	Rage Against the Mean – A Review of Distributional Regression Approaches	2023
78	Nogueira D; SáFadi T; De L R; Da M A; De C G M; De P B J; Barãante T; Dourado S	Cadernos de Saude Publica	Spatio-temporal GAMLSS modeling of the incidence of schistosomiasis in the central region of the State of Minas Gerais, Brazil	2023
79	Almiray-Soto A; Denova-Gutiérrez E; Lopez-Gonzalez D; Medeiros M; Clark P	Calcified Tissue International	Muscle Strength Reference Values and Correlation with Appendicular Muscle Mass in Mexican Children and Adolescents	2022
80	Wang Y; Liu M; Xing Z; Liu H; Song J; Hou Q; Xu Y	Water	Study of Nonstationary Flood Frequency Analysis in Songhua River Basin	2023
81	Hans N; Klein N; Faschingbauer F; Schneider M; Mayr A	Biometrics	Boosting Distributional Copula Regression	2023
82	Yousuf E; Rochow N; Li J; Simioni J; Gunn E; Hutton E; Morrison K	Pediatric Research	Growth and body composition trajectories in infants meeting the WHO growth standards study requirements	2022

83	Wu W;Chen J;Mo M;Si S;Huang K;Chen R;Maimaiti M;Chen S;Gong C;Zhu M;Wang C;Su Z;Liang Y;Yao H;Wei H;Zheng R;Du H;Yang Y;Luo F;Li P;Cui L;Dong G;Yu Y;Fu J	BMC Public Health	Construction of a new complete growth reference for urban Chinese children	2022
84	Gao Y;Xia J;Chen X;Zou L;Huang J;Yu J	Journal of Hydrology: Regional Studies	Analysis of the nonstationarity characteristics and future trends of flood extremes in the Dongting Lake Basin	2022
85	Kretschmer L;Salali G;Andersen L;Hallal P;Northstone K;Sardinha L;Dyble M;Bann D	International Journal of Behavioral Nutrition and Physical Activity	Gender differences in the distribution of children's physical activity: evidence from nine countries	2023
86	Stasinopoulos D M;Rigby R A;Heller G;De Bastiani	Statistical Modelling	P-splines and GAMLSS: a powerful combination, with an application to zero-adjusted distributions	2023
87	Strömer A;Klein N;Staerk C;Klinkhammer H;Mayr A	Statistics in Medicine	Boosting multivariate structured additive distributional regression models	2023
88	Wattanawongwan S;Mues C;Okhrati R;Choudhry T;So M	International Journal of Forecasting	A mixture model for credit card exposure at default using the GAMLSS framework	2023
89	Loedy N;Coletti P;Wambua J;Hermans L;Willem L;Jarvis C;Wong K;Edmunds W;Robert A;Leclerc Q;Gimma A;Molenberghs G;Beutels P;Faes C;Hens N	BMC Public Health	Longitudinal social contact data analysis: insights from 2 years of data collection in Belgium during the COVID-19 pandemic	2023
90	Siegfried S;Kook L;Hothorn T	American Statistician	Distribution-Free Location-Scale Regression	2023
91	Ildefonso Pg;Nieto L A;Librero Ab; Alonso Mm;Cerceño Mlh;Serrano Eg;Prieto-Matos P	Anales De Pediatria	Normal range for acid-labile subunit in paediatric patients in Spain and its association with age, sex, pubertal stage and other growth factors	2023
92	Roquim F;Lima R;Nakamura L;Ramires T;Blanco Y	Brazilian Journal of Biometrics	Dealing with highly kurtotic count data with excess zeroes: comparing different treatments in the control of dairy cattle gastrointestinal parasites	2023
93	Michels M;Luo H;Weller V A P;Muãÿhoff O	Journal of Behavioral and Experimental Economics	Compliance with pre-harvest interval rules in apple production—A comparative analysis of green nudges among fruit growers and agricultural students in Germany	2023
94	Fairclough S;Rowlands A;Del P C B;Crotti M;Foweather L;Graves L;Hurter L;Jones O;Macdonald M;Mccann D;Miller C;Noonan R;Owen M;Rudd J;Taylor S;Tyler R;Boddy Y;Lu F;Zhou Y;Ruan B;Dai Y;Wang K	International Journal of Behavioral Nutrition and Physical Activity	Reference values for wrist-worn accelerometer physical activity metrics in England children and adolescents	2023
95	Nakamura L;Ramires T;Righetto A;Silva V;Konrath A	Brazilian Journal of Biometrics	Using the Box-Cox family of distributions to model censored data:a distributional regression approach	2022
96	Reyes H;Diethelm-Varela B;Méndez C;Rebolledo-Zelada D;Lillo-Dapremont B;Muã'oz S;Bueno S;González P;Kalergis A	Frontiers in Public Health	Contribution of Two-Dose Vaccination Toward the Reduction of COVID-19 Cases, ICU Hospitalizations and Deaths in Chile Assessed Through Explanatory Generalized Additive Models for Location, Scale, and Shape	2022

97	Dixit S;Jayakumar K	Water Resources Management	A Non-stationary and Probabilistic Approach for Drought Characterization Using Trivariate and Pairwise Copula Construction (PCC) Model	2022
98	Malavasi M;Peters G;Shevchenko P;Trück S;Jang J;Sofronov G	Insurance: Mathematics and Economics	Cyber risk frequency, severity and insurance viability	2022
99	Areekal S;Goel P;Khadilkar A;Khadilkar V;Cole T	Annals of Human Biology	Assessment of height growth in Indian children using growth centiles and growth curves	2022
100	Ma N;Dang J;Liu Y;Zhong P;Yan X;Zhang J;Dong Y;Song Y;Ma J;Lau P	Frontiers in Nutrition	Percentile Curves for Multiple Physical Fitness Components Among Chinese Han Children and Adolescents Aged 7–18 Years From a National Survey Based on the Total and the Normal Weight Population	2022
101	Ducharme F;Smyrnova A;Lawson C;Miles L	Pediatric Pulmonology	Reference values for respiratory sinusoidal oscillometry in children aged 3 to 17 years	2022
102	Kousali M;Salarijazi M;Ghorbani K	Natural Resources Research	Estimation of Non-Stationary Behavior in Annual and Seasonal Surface Freshwater Volume Discharged into the Gorgan Bay, Iran	2022
103	León C;Úsuga O;Peñuela G	Environmental Monitoring and Assessment	Statistical model applied in the qualitative and quantitative analysis of the relationships between environmental variables and cyanotoxin concentrations in tropical reservoirs	2022
104	Jancikova V;Pridalova M;Kaplanova T	Children	Reference Curves of Selected Circumferential Parameters for Czech Children Aged 6 to 11 Years	2021
105	Ziel F	International Journal of Forecasting	M5 competition uncertainty: Overdispersion, distributional forecasting, GAMLSS, and beyond	2022
106	Wu D;Yoon H;Lee J;Kim J	Applied Sciences	Regional Analysis of Hotspot and Coldspot Areas Undergoing Nonstationary Drought Characteristics in a Changing Climate	2022
107	De Oliveira J;Corrêa C;Rocha R;Santos E;Beltrame R;Borges A	Revista Brasileira de Enfermagem	Implications of health conditions on rural soybean workers' quality of life	2022
108	Graber J;Juarez-Colunga E;Thigpen C;Waugh D;Bade M;Stevens-Lapsley J;Kittelson A	Disability and Rehabilitation	Development of reference charts for monitoring quadriceps strength with handheld dynamometry after total knee arthroplasty	2022
109	Scala P;Cipolla G;Treppiedi D;Noto L	Water	The Use of GAMLSS Framework for a Non-Stationary Frequency Analysis of Annual Runoff Data over a Mediterranean Area	2022
110	Stadlmann S;Kneib T	Sage Journals (Statistical Modelling)	Interactively visualizing distributional regression models with distreg.vis	2022
111	Wang M;Jiang S;Ren L;Xu C;Shi P;Yuan S;Liu Y;Fang X	Journal of Hydrology	Nonstationary flood and low flow frequency analysis in the upper reaches of Huaihe River Basin, China, using climatic variables and reservoir index as covariates	2022
112	Mattiello R;Mundstock E;Ziegelmann P	Frontiers in Nutrition	Brazilian Reference Percentiles for Bioimpedance Phase Angle of Healthy Individuals	2022
113	Wu Z;Liu S;Wang H	Journal of Water and Climate Change	Calculation method of short-duration rainstorm intensity formula considering nonstationarity of rainfall series: impacts on the simulation of urban drainage system	2021
114	Chhavi N;Ojha S;Awasthi A;Shalimar S;Goel A	Journal of Clinical and Experimental Hepatology	Serum Level of Alanine- and Aspartate-Aminotransferase Levels in Newborns in India	2022
115	Lu D;Tong X;Liu Y;Zhang H	Frontiers in Pediatrics	Reference Values for Point-of-Care Echocardiographic Measurements of Preterm Infants in China	2022

116	Karim R;Aker N	Theoretical and Applied Climatology	Effects of climate variables on the COVID-19 mortality in Bangladesh	2022
117	Voncken L;Albers C;Timmerman M	Assessment	Bias-Variance Trade-Off in Continuous Test Norming	2021
118	Marmolejo-Ramos F;Ospina R;Hernández-Barajas F	Australian and New Zealand Journal Of Statistics	The place of probability distributions in statistical learning. A commented book review of 'Distributions for modeling location, scale, and shape using GAMLSS in R' by Rigby et al. (2021)	2022
119	Hao W;Shao Q;Wei P;Zhu C;Chen X;Chen R	Journal Of Water And Climate Change	Evaluation and projection of the annual maximum streamflow in response to anthropogenic and climatic effects under nonstationary conditions in the Hanjiang River Basin, China	2022
120	Al-Qerem W	International Journal Of Clinical Practice	Spirometry reference equations for children from a Middle Eastern population	2021
121	Ma C;Li L;Wang X;Hou L;Xia L;Yin Y;Cheng X;Qiu L	Frontiers In Cardiovascular Medicine	Establishment of Reference Interval and Aging Model of Homocysteine Using Real-World Data	2022
122	Markus C;Tan R;Lim C;Rankin W;Matthews S;Loh T;Hague W	Clinical Chemistry And Laboratory Medicine	Performance of four regression frameworks with varying precision profiles in simulated reference material commutability assessment	2022
123	Gilmour M;Felis J;Hester M;Young L;Adams J	Marine Ecology Progress Series	Laysan albatross exhibit complex behavioral plasticity in the subtropical and subarctic North Pacific Ocean	2022
124	Hoang T;Coletti P;Kifle Y;Kerckhove K;Vercruyse S;Willem L;Beutels P;Hens N	BMC Infectious Diseases	Close contact infection dynamics over time: insights from a second large-scale social contact survey in Flanders, Belgium, in 2010-2011	2021
125	Garcia-Rojas M;Keatley M;Roslan N	Plos One	Citizen science and expert opinion working together to understand the impacts of climate change	2022
126	Thiruvengadam R;Desiraju B;Natchu U;Wadhwa N;Sachdeva K;Misra S;Parmar N;Juyal M;Mittal P;Bharti R;Tripathi R;Ramji S;Sachdev H;Bhatnagar S	European Journal Of Clinical Nutrition	Gestational weight gain trajectories in GARBH–Ini pregnancy cohort in North India and a comparative analysis with global references	2022
127	Liew B;Rugamer D;Duffy K;Taylor M;Jackson J	Plos One	The mechanical energetics of walking across the adult lifespan	2021
128	Wu Q;Zhang H;Zhang L;Xu Y;Sun J;Gao N;Qiao X;Li Y	Frontiers In Pediatrics	A New Birthweight Reference by Gestational Age: A Population Study Based on the Generalized Additive Model for Location, Scale, and Shape Method	2022
129	Nakamura L;Ramires T;Righetto A;Pescim R;Roquim F;Savian T;Stasinopoulos D M	Computers And Electronics In Agriculture	Cattle reference growth curves based on centile estimation: A GAMLSS approach	2022
130	Jiménez-U M;Peña L;López J	Heliyon	Non-stationary analysis for road drainage design under land-use and climate change scenarios	2022
131	Ulrich M;Jahnke H;Langrock R;Pesch R;Senge R	European Journal Of Operational Research	Distributional regression for demand forecasting in e-grocery	2021
132	Wiedermann W;Zhang B;Reinke W;Herman K;Von E A	Psychological Methods	Distributional causal effects: Beyond an “averagarian” view of intervention effects.	2022
133	Flatley C;Sole-Navais P;Vaudel M;Helgeland ;Modzelewska D;Johansson S;Jacobsson B;Njølstad P	Placenta	Placental weight centiles adjusted for age, parity and fetal sex	2022
134	Wada Y;Goto N;Kitaguchi Y;Yasuo M;Hanaoka M	Plos One	Referential equations for pulmonary diffusing capacity using GAMLSS models derived from Japanese individuals with near-normal lung function	2022

135	Costa C;De Oliveira A;Caraciolo J;De Lucena L;Leite M	Acta Scientiarum - Agronomy	A GAMLSS approach to predicting growth of <i>Nopalea cochenillifera</i> Giant Sweet clone submitted to water and saline stress	2022
136	Singh J;Ghosh S;Simonovic S;Karmakar S	Hydrological Processes	Identification of flood seasonality and drivers across Canada	2021
137	Neuhauser H;Büschges J;Schaffrath R A;Schienkiewitz A;Sarganas G;Königstein K;Schweizer D;Schmidt-Trucksäss A	Hypertension	Carotid Intima-Media Thickness Percentiles in Adolescence and Young Adulthood and Their Association With Obesity and Hypertensive Blood Pressure in a Population Cohort	2022
138	Lei X;Gao L;Ma M;Wei J;Xu L;Wang L;Lin H	Journal Of Hydrology: Regional Studies	Does non-stationarity of extreme precipitation exist in the Poyang Lake Basin of China?	2021
139	Ren M;Villarini G;Pang B;Xu Z;Du L;Wang Y	Journal Of Water And Climate Change	Hydrological response in a highly urbanized watershed in China	2022
140	Prataviera F;Cordeiro G;Ortega E;Hashimoto E;Cancho V	Journal Of Applied Statistics	A new regression model for rates and proportions data with applications	2022
141	Mateo-Abad M;Vrotsou K;Ruiz M;Montiel-Luque A;Del C S F M;Machón M;Ruiz F;Vergara I	Plos One	Use of health care services according to functional performance in community-dwelling older adults in Spain. An approach using GAMLSS models	2022
142	Zavorsky G;Cao J	BMJ Open Respiratory Research	Reference equations for pulmonary diffusing capacity using segmented regression show similar predictive accuracy as GAMLSS models	2022
143	Mota A;Santos-Neto M;Neto M;Leão J;Tomazella V;Louzada F	Communications In Statistics - Simulation And Computation	Weighted Lindley regression model with varying precision: estimation, modeling and its diagnostics	2022
144	Babochkina T;Gerlinskaya L;Anisimova M;Kontsevaya G;Feofanova N;Stanova A;Moshkin M;Moshkin Y	International Journal Of Molecular Sciences	Mother–Fetus Immune Cross-Talk Coordinates “Extrinsic”/“Intrinsic” Embryo Gene Expression Noise and Growth Stability	2022
145	Schaak H;Musshoff O	European Review Of Agricultural Economics	The distribution of the rent–price relationship of agricultural land in Germany	2022
146	Hohberg M;Donat F;Marra G;Kneib T	Journal Of The Royal Statistical Society. Series C: Applied Statistics	Beyond Unidimensional Poverty Analysis Using Distributional Copula Models for Mixed Ordered-Continuous Outcomes	2021
147	Dejsomritrutai W;Chuaychoo B;Chierakul N	Journal Of The Medical Association Of Thailand	A New Derivation of the Reference Equations of Spirometric Values in the Thai Population Using the GAMLSS Method	2022
148	Yan L;Lu D;Hu P;Yan P;Xu Y;Qi J;Liu F;Li J	Hydrological Sciences Journal	Estimation of design precipitation in Beijing–Tianjin–Hebei region under a changing climate	2022
149	Xu P;Wang D;Wang Y;Singh V	Journal Of Hydrologic Engineering	A Stepwise and Dynamic C-Vine Copula–Based Approach for Nonstationary Monthly Streamflow Forecasts	2022
150	Prataviera F;Ortega E;Cordeiro G;Cancho V	Communications In Statistics - Simulation And Computation	The exponentiated power exponential semiparametric regression model	2022
151	Zhou Y;Guo S;Xu C;Xiong L;Chen H;Ngongondo C;Li L	Hydrology Research	Probabilistic interval estimation of design floods under non-stationary conditions by an integrated approach	2022

152	Do Réngo N;De Andrade H;Duarte-Neto P	Boletim Do Instituto De Pesca	Length structure analysis of the Atlantic thread herring (<i>Opisthonema oglinum</i>) captured in the Santa Cruz Channel (Pernambuco, Brazil), using generalized additive models for location, scale and shape	2022
153	Van B I;Van D S M;De F K;Spaans S;Hokken-Koelega A	Clinical Nutrition	Body composition and bone mineral density by Dual Energy X-ray Absorptiometry: Reference values for young children	2022
154	Bazrafshan J;Cheraghalizadeh M;Shahgholian K	Water Resources Management	Development of a Non-stationary Standardized Precipitation Evapotranspiration Index (NSPEI) for Drought Monitoring in a Changing Climate	2022
155	Mikis D;Robert A;Nikolaos G;Fernanda D	Sage Journals (Statistical Modelling)	Principal component regression in GAMLSS applied to Greek-German government bond yield spreads	2022
156	Datson N;Weston M;Drust B;Atkinson G;Lolli L;Gregson W	Science And Medicine In Football	Reference values for performance test outcomes relevant to English female soccer players	2022
157	Marker R;Kittelson A;Jankowski C;Scorsone J;Peters J;Purcell W	Rehabilitation Oncology	Development of a Reference Chart for Monitoring Cancer-Related Fatigue Throughout a Supervised Exercise Program	2022
158	Meng F;Yuan Y;Jung S;Stimm B;Estrella N;Menzel A	Agricultural And Forest Meteorology	Long-term flowering intensity of European tree species under the influence of climatic and resource dynamic variables	2022
159	Dangat K;Gupte S;Wagh G;Lalwani S;Randhir K;Madiwale S;Pisal H;Kadam V;Gundu S;Chandhiok N;Kulkarni B;Joshi S;Fall C;Sachdev H	Frontiers In Medicine	Gestational weight gain in the REVAMP pregnancy cohort in Western India: Comparison with international and national references	2022
160	Tosunoglu F;Slater L	River Research And Applications	Decreasing flood hazard evaluated in Turkey using nonstationary models	2022
161	Almasi F;Nguyen H;Heydarian D;Sohi R;Nikbin S;Jenvey C;Halliwell E;Ponnampalam E;Desai A;Jois M;Stear M	Animal Production Science	Quantification of behavioural variation among sheep grazing on pasture using accelerometer sensors	2022
162	Shao S;Zhang H;Singh V;Ding H;Zhang J;Wu Y	Journal Of Hydrology	Nonstationary analysis of hydrological drought index in a coupled human-water system: Application of the GAMLSS with meteorological and anthropogenic covariates in the Wuding River basin, China	2022
163	Fitzgerald K;Sotirchos E;Smith M;Lord H;Duval A;Mowry E;Calabresi P	Annals Of Neurology	Contributors to Serum NFL Levels in People without Neurologic Disease	2022
164	De A H;Arroyo L;Yamamura M;Ramos A;De A C J;Berra T;Neto M;Pinto I;Palha P;Monroe A;Alexander K;Arcêcio R	Journal Of Infection In Developing Countries	Social inequalities associated with the onset of tuberculosis in disease-prone territories in a city from northeastern Brazil	2021
165	Godoy A;Marques A;Teixeira G;Freire F;Costa R;Fransozo A	Regional Studies In Marine Science	Modulating factors influencing abundances and distribution of brachyuran communities (Crustacea, Decapoda, Brachyura) close to the islands in Ubatuba, Brazil	2022
166	Calabrese R;Zanin L	Expert Systems With Applications	Modelling spatial dependence for Loss Given Default in peer-to-peer lending	2022
167	Ma X;Mao Y;Wang J;Zewangzhandui Z;Wang X	Frontiers In Nutrition	Anthropometric indices, body function, and physical fitness reference values for Tibetan ethnic children aged 6-17 residing at 3,650 meters above sea level	2022

168	Han S;Slater L;Wilby R;Faulkner D	Journal Of Hydrology	Contribution of urbanisation to nonstationary river flow in the uk	2022
169	Jia H;Wan B;Bu T;Luo Y;Ma W;Huang S;Gang L;Deng W;Liu Z	Frontiers In Physiology	Chinese physical fitness standard for campus football players: A pilot study of 765 children aged 9 to 11	2022
170	Liu H;Zou L;Xia J;Chen T;Wang F	Sustainable Cities And Society	Impact assessment of climate change and urbanization on the nonstationarity of extreme precipitation: A case study in an urban agglomeration in the middle reaches of the Yangtze river	2022
171	Zhang L;Gao N;Liu H;Wu Q;Liu J;Zhang T;Sun J;Qi J;Qiao X;Zhao Y;Li Y	Frontiers In Pediatrics	Differences in Postnatal Growth of Preterm Infants in Northern China Compared to the INTERGROWTH-21st Preterm Postnatal Growth Standards: A Retrospective Cohort Study	2022
172	Bann D;Wright L;Cole T	eLife	Risk factors relate to the variability of health outcomes as well as the mean: A GAMLSS tutorial	2022
173	Cheng Y;Liu M;Liu Y;Xu H;Chen X;Zheng H;Wu X;Shen Z;Shen C	BMC Nephrology	Chronic kidney disease: prevalence and association with handgrip strength in a cross-sectional study	2021
174	Wang Y;Peng T;Lin Q;Singh V;Dong X;Chen C;Liu J;Chang W;Wang G	Water Resources Management	A New Non-stationary Hydrological Drought Index Encompassing Climate Indices and Modified Reservoir Index as Covariates	2022
175	Zhang B;Hepp T;Grevén S;Bergherr E	Computational Statistics	Adaptive step-length selection in gradient boosting for Gaussian location and scale models	2022
176	Arnold N;Ascherl R;Thome U	Frontiers In Pediatrics	Charts and LMS Tables of Transfontanelar and Transvertical Ear-to-Ear Distances for Gestational Age	2022
177	Clemente G;Savelli N;Spedicato G;Zappa D	North American Actuarial Journal	Modeling General Practitioners' Total Drug Costs through GAMLSS and Collective Risk Models	2022
178	Mariën B;Papadimitriou D;Kotilainen T;Zuccarini P;Dox I;Verlinden M;Heinecke T;Mariën J;Willems P;Decoster M;Gascó A;Lange H;Peñuelas J;Campioli M	Agricultural And Forest Meteorology	Timing leaf senescence: A generalized additive models for location, scale and shape approach	2022
179	Narajewski M	Energies	Probabilistic Forecasting of German Electricity Imbalance Prices	2022
180	Anjos A;Alvarado S;Solé M;Benchimol M	Journal For Nature Conservation	Patch and landscape features drive fire regime in a Brazilian flammable ecosystem	2022
181	Peterson M;Cherukuri V;Paulson J;Ssentongo P;Kulkarni A;Warf B;Monga V;Schiff S	Journal Of Neurosurgery: Pediatrics	Normal childhood brain growth and a universal sex and anthropomorphic relationship to cerebrospinal fluid	2021
182	Matsumoto T;Bunn D;Yamada Y	Quantitative Finance	Pricing electricity day-ahead cap futures with multifactor skew-t densities	2022
183	Adam T;Mayr A;Kneib T	Econometrics And Statistics	Gradient boosting in markovswitching generalized additive models for location scale and shape	2022
184	De Bastiani;Stasinopoulos D M;Rigby R A;Heller G;Silva L	Brazilian Journal Of Probability And Statistics	Bucket plot a visual tool for skewness and kurtosis comparisons	2022
185	Karim M;Akter M;Haque S;Akter N	Annals Of Data Science	Do temperature and humidity affect the transmission of sarscov2a flexible regression analysis	2022
186	Sinnige A;Kittelson A;Van D W P;Tejjink J;Hoogeboom G;Carrilho T;Rasmussen K;Reichenheim M;Farias D;Hutcheon J	European Journal of Vascular and Endovascular Surgery	Personalised outcomes forecasts of supervised exercise therapy in intermittent claudication: an application of neighbours based prediction methods with routinely collected clinical data	2022

187	De A A;Marques A;Dantas A;De M J M;De M G;Tinoco M	Aquatic Sciences	Assisted phoresy of invertebrates by anurans in tank bromeliads interspecific relationship	2020
188	Zhang S;Jiang C;Tu C	Iranian Journal Of Public Health	Chinese preschool children growth reference standard and related curve based on gamlss approach	2021
189	Geirsson P;Hrafnkelsson B;Simpson D;Sigurdarson H	Statistical Science	Lgm split sampler an efficient mcmc sampling scheme for latent gaussian models	2020
190	Valencia-Sosa E;Chávez-Palencia C;Vallarta-Robledo J;Romero-Velarde E;Larrosa-Haro A;Vásquez-Garibay E;Ramos-García C	Children	Percentile reference values for the neck circumference of Mexican children	2021
191	Lopez-Gonzalez D;Wells J;Cortina-Borja M;Fewtrell M;Partida-Gaytán A;Clark P	Bone	Reference values for bone mineral density in healthy Mexican children and adolescents	2021
192	Bergerat M;Heude B;Taine M;Nguyen T T S;Werner A;Frاندji B;Blauwblomme T;Sumanaru D;Charles M;Chalumeau M;Scherdel P	The Lancet Regional Health - Europe	Head circumference from birth to five years in France new national reference charts and comparison to who standards	2021
193	Zahan R;Khan S;Rennie D;Karunanayake C;Fenton M;Seeseequasis J;Arnault D;Gardipy J;Dosman J;Pahwa P	Model Assisted Statistics And Applications	Lung function reference equations and lower limit of normal for cree first nations children and adolescents living in rural saskatchewan Canada	2020
194	Ribeiro M;Andrade L;Spyrides M;Lima K;Da S P;Batista D;De L I	Weather, Climate, And Society	Environmental disasters in northeast brazil hydrometeorological social and sanitary factors	2021
195	Riepsamen A;Donoghoe M;Baerwald A;Pankhurst M;Lien S;Chong Y;Robertson D;Ledger W;Gilchrist R	Fertility And Sterility	Exploratory analysis of serum concentrations of oocyte biomarkers growth differentiation factor 9 and bone morphogenetic protein 15 in ovulatory women across the menstrual cycle	2021
196	Zbořilová V;Přidalová M;Kaplanová T	Children	Body fat mass percent body fat fatfree mass and skeletal muscle mass reference curves for Czech children aged 6-11 years	2021
197	Cajias M	Journal Of European Real Estate Research	Is there room for another hedonic model? The advantages of the gamlss approach in real estate research	2018
198	Wang J;Chen F;Doan Q;Xu Y	Urban Climate	Exploring the effect of urbanization on hourly extreme rainfall over yangtze river delta of china	2021

199	Xi B;Zong X;Kelishadi R;Litwin M;Hong Y;Poh B;Steffen L;Galcheva S;Herter-Aeberli I;Nawarycz T;Krzywińska-Wiewiorowska M;Khadiikar A;Schmidt M;Neuhauser H;Schienkiewitz A;Kulaga Z;Kim H;Stawińska Witoszyńska B;Motlagh M;Ruzita A;Iotova V;Grajda A;Ismail M;Krzyżaniak A;Heshmat R;Stratev V;Rózdzyńska-Świątkowska A;Ardalan G;Qorbani M;Świąder-Leśniak A;Ostrowska-Nawarycz L;Yotov Y;Ekbote V;Khadiikar V;Venn A;Dwyer T;Zhao M;Magnussen C;Bovet P	Journal Of Clinical Endocrinology And Metabolism	International waist circumference percentile cutoffs for central obesity in children and adolescents aged 6 to 18 years	2020
200	Guo Z;Boeing W;Xu Y;Yan C;Faghihinia M;Liu D	Water Research	Revisiting seasonal dynamics of total nitrogen in reservoirs with a systematic framework for mining data from existing publications	2021
201	Abramova E;Bunn D	Energies	Forecasting the intraday spread densities of electricity prices	2020
202	Gonçalves J;Barreto-Souza W	Metron	Flexible regression models for counts with highinflation of zeros	2020
203	Hibatullah F;Kuswanto H;Sutikno S	Journal Of Advanced Research In Dynamical And Control Systems	Nonstationary generalized additive modelbased modelling of extreme precipitation indices in east nusa tenggara Indonesia	2020
204	Kittelson A;Elings J;Colborn K;Hoogeboom T;Christensen J;Van M N;Van B S;Stevens-Lapsley J	Bmc Musculoskeletal Disorders	Reference chart for knee flexion following total knee arthroplasty a novel tool for monitoring postoperative recovery	2020
205	Jehanzaib M;Shah S;Yoo J;Kim T	Journal Of Hydrology	Investigating the impacts of climate change and human activities on hydrological drought using nonstationary approaches	2020
206	Zouiouich S;Lofffield E;Huybrechts I;Viallon V;Louca P;Vogtmann E;Wells P;Steves C;Herzig K;Menni C;Jarvelin M;Sinha R;Gunter M	Diabetologia	Markers of metabolic health and gut microbiome diversity findings from two populationbased cohort studies	2021
207	Das S;Das J;Umamahesh N	Journal Of Hydrologic Engineering	Nonstationary modeling of meteorological droughts application to a region in India	2021
208	Yang L;Li J;Kang A;Li S;Feng P	Water Resources Management	The effect of nonstationarity in rainfall on urban flooding based on coupling swmm and mike21	2020
209	Molina V;Belmar L;Levipan H;RamiRez-Flandes S;Anguita C;Galán A;Montes I;Ulloa O	Frontiers In Marine Science	Spatiotemporal distribution of key pelagic microbes in a seasonal oxygendeficient coastal upwelling system of the eastern south Pacific Ocean	2020
210	Pitt D;Trück S;Van D H R;Wong W	Journal Of Environmental Management	Modeling risks from natural hazards with generalized additive models for location scale and shape	2020

211	Baer S;Tuzin N;Kang P;Mohammed S;Kubota M;Van I Y;Busa T;Rossi M;Morel G;Michot C;Baujat G;Durand M;Obringer C;Le M N;Calmels N;Laugel V	European Journal Of Medical Genetics	Growth charts in cockayne syndrome type 1 and type 2	2021
212	Newman L;Nolan H;Carey D;Reilly R;Kenny R	Archives Of Gerontology And Geriatrics	Age and sex differences in frontal lobe cerebral oxygenation in older adultsnormative values using novel scalable technology findings from the irish longitudinal study on ageing tilda	2020
213	De F L;Taconeli C;Da S J;Tamioso P;Molento C	Revista Brasileira De Biometria	A case study on animal behavior analysis using gamlss um estudo de caso em analise de comportamento animal usando gamlss	2021
214	Tadphale S;Zurakowski D;Bird L;Yohannan T;Agrawal V;Lloyd H;Allen K;Waller B;Hall A;Sathanandam S	Pediatric Cardiology	Construction of femoral vessel nomograms for planning cardiac interventional procedures in children 04 years old	2020
215	Shiau J	Water Resources Management	Effects of gammadistribution variations on spibased stationary and nonstationary drought analyses	2020
216	Voigt M;Rochow N;Landau-Crangle E;Meyer-Kahrweg L;Olbertz D;Kunze M;Nikischin W;Wittwer-Backofen U;Rochow M;Däbritz J;Hentschel R	Journal Of Perinatal Medicine	Individualized sexspecific birth weight percentiles for gestational age based on maternal height and weight	2021
217	Rydén J	Acta Geophysica	A note on analysis of extreme minimum temperatures with the gamlss framework	2019
218	Aeberhard W;Cantoni E;Marra G;Radice R	Statistics And Computing	Robust fitting for generalized additive models for location scale and shape	2021
219	Yagli G;Yang D;Srinivasan D	Solar Energy	Ensemble solar forecasting using datadriven models with probabilistic postprocessing through gamlss	2020
220	Rodríguez-González A;May-Tec A;Herrera-Silveira J;Puch-Hau C;Quintanilla-Mena M;Villafuerte J;Velázquez-Abunader I;Aguirre-Macedo M;Vidal-Martínez V	Ecological Indicators	Fluctuating asymmetry of sclerotized structures of haliotrematoides spp monogenea dactylogyridae as bioindicators of aquatic contamination	2020
221	Gonçalves M;Pedro J;Silva C;Magalhães P;Brito M	Journal Of Electrocardiology	Normal limits of the electrocardiogram in angolans	2020
222	Li K;Hu L;Peng Y;Yan R;Li Q;Peng X;Song W;Ni X	Bmc Medical Research Methodology	Comparison of four algorithms on establishing continuous reference intervals for pediatric analytes with agedependent trend	2020
223	Cole T	Statistical Methods In Medical Research	Sample size and sample composition for constructing growth reference centiles	2021
224	Rivas D;Torres N;Pino M;Vásquez M;Casana Y;Bazan E;Carrera C;Shimabuku Y R;Santos R;Calderón L;Chamiqut C	Pakistan Journal Of Biological Sciences	Using an algorithm to fit a gamlss model on dry matter data from brachiaria brizantha	2021

225	Funk S;Palomo G B;Bueno Z A;Ickowitz A;Poni N;Abdou M;Sibama Y;Penda R;Ros B G;Abossolo M;Martin E;Okale R;Ze B;Moreno C A;Sebastián C G ;Ruiz D L G C;López-Romero S F;Amazia H;Álvarez R I;Sánchez E R;Fa J	Human Ecology	Understanding growth and malnutrition in baka pygmy children	2020
226	Navin P;Moynagh M;Atkinson E;Tirumanisetty P;Lebrasseur N;Kumar A;Khosla S;Takahashi N	Clinical Nutrition	Establishment of normative biometric data for body composition based on computed tomography in a North American cohort	2021
227	Loyd B;Kittelson A;Forster J;Stackhouse S;Stevens-Lapsley J	Disability And Rehabilitation	Development of a reference chart to monitor postoperative swelling following total knee arthroplasty	2020
228	Timmerman M;Voncken L;Albers C	Psychological Methods	A tutorial on regressionbased norming of psychological tests with gamlss	2020
229	Van E A;Eerens S;Trouet D;Lauwers E;Wouters K;De W B;Van D L J;Van H K;Ledeganck K	European Journal Of Pediatrics	Body composition monitoring in children and adolescents reproducibility and reference values	2021
230	Berger M;Schmid M	Statistical Methods In Medical Research	Flexible modeling of ratio outcomes in clinical and epidemiological research	2020
231	Rodrigues J;Bazán J;Suzuki A	Communications In Statistics - Theory And Methods	A flexible procedure for formulating probability distributions on the unit interval with applications	2020
232	Yamada G;Jones-Smith J;Castillo-Salgado C;Moulton L	European Journal Of Clinical Nutrition	Differences in magnitude and rates of change in bmi distributions by socioeconomic and geographic factors in Mexico Colombia and Peru 20052010	2020
233	Leidhin C;Mcmorrow J;Carey D;Newman L;Williamson W;Fagan A;Chappell M;Kenny R;Meaney J;Knight S	Neuroimage	Agerelated normative changes in cerebral perfusion data from the irish longitudinal study on ageing tilda	2021
234	Klein N;Entwistle A;Rosenberger A;Kneib T;Bickeböller H	Journal Of Applied Statistics	Candidategene association analysis for a continuous phenotype with a spike at zero using parentoffspring trios	2020
235	Kessels R;Hoorweg A;Thanh B T;Erreygers G	International Journal For Equity In Health	A distributional regression approach to incomerelated inequality of health in Australia	2020
236	Truckenbrod C;Meigen C;Brandt M;Vogel M;Wahl S;Jurkutat A;Kiess W	Plos One	Reference curves for refraction in a German cohort of healthy children and adolescents	2020
237	Shinde P;Hesamzadeh M;Date P;Bunn D	Ieee Transactions On Power Systems	Optimal dispatch in a balancing market with intermittent renewable generation	2021
238	Vaccari F;Fiori F;Bravo G;Parpinel M;Messina G;Malavolta R;Lazzer S	European Journal Of Pediatrics	Physical fitness reference standards in Italian children	2021
239	Shiau J;Liu Y	Water (Switzerland)	Nonstationary analyses of the maximum and minimum streamflow in tamsui river basin Taiwan	2021
240	Wang Y;Xu Y;Song S;Wang J;Zhang J	Environmental Science And Pollution Research	Assessing the impacts of climatic and anthropogenic factors on water level variation in the taihu plain based on nonstationary statistical models	2020
241	Boucher J;Turcotte R	Risks	A longitudinal analysis of the impact of distance driven on the probability of car accidents	2020

242	Marletta A;Sciandra M	International Journal Of Biostatistics	Gamlss for highvariability data an application to liver fibrosis case	2020
243	Quintino N;Sabino E;Da S J;Ribeiro A;Ferreira A;Davi G;Oliveira C;Cardoso C	Plos Neglected Tropical Diseases	Factors associated with quality of life in patients with chagas disease samitrop project	2020
244	Lee H;Park B;Park E;Cho S;Kim H;Choi E;Kim N;Park H	Journal Of Korean Medical Science	Blood pressure curve for children less than 10 years of age findings from the ewha birth and growth cohort study	2020
245	Rojas F;Leiva V;Huerta M;Martin-Barreiro C	Sensors	Lotsize models with uncertain demand considering its skewnesskurtosis and stochastic programming applied to hospital pharmacy with sensorrelated covid19 data	2021
246	Ruiz B M;Cuestas E;Cieri M;Cuestas E	American Journal Of Human Biology	Reference ranges for knee height in argentine children and adolescents aged 2 to 18 years	2020
247	Xu P;Wang D;Singh V;Lu H;Wang Y;Wu J;Wang L;Liu J;Zhang J	Journal Of Hydrology	Copulabased seasonal rainfall simulation considering nonstationarity	2020
248	Sun X;Li Z;Tian Q	Hydrology Research	Assessment of hydrological drought based on nonstationary runoff data	2020
249	Ivanovski I;Ivanovski I;Djuric O;Broccoli S;Caraffi S;Accorsi P;Adam M;Avela K;Badura-Stronka M;Bayat A;Clayton-Smith J;Cocco I;Cordelli D;Cuturilo G;Di P V;Dupont G J;Gastaldi R;Giordano L;Guala A;Hoei-Hansen C;Inaba M;Iodice A;Nielsen J;Kuburovic V;Lazalde-Medina B;Malbora B;Mizuno S;Moldovan O;Møller R;Muschke P;Otelli V;Pantaleoni C;Piscopo C;Poch-Olive M;Prpic I;Marin R P;Raviglione F;Ricci E;Scarano E;Simonte G;Smigiel R;Tanteles G;Tarani L;Trimouille A;Valera E;Schrier V S;Writzl K;Callewaert B;Savasta S;Street M;lughetti L;Bernasconi S;Giorgi R P;Garavelli L	Orphanet Journal Of Rare Diseases	Mowatwilson syndrome growth charts	2020
250	Seidel E;Pazini J;Tomazella V;Vieira A;Silva F;Martins J;Barrigossi J	Environmental Entomology	Predicting rice stem stink bug population dynamics based on gamlss models	2020
251	Tu K;Yan Z	International Journal Of Climatology	Nonstationary climate changes in summer hightemperature extremes in shanghai since the late 19th century	2021
252	Qu C;Li J;Yan L;Yan P;Cheng F;Lu D	Water (Switzerland)	Nonstationary flood frequency analysis using cubic bsplinebased gamlss model	2020
253	Winkler C;Linden K;Mayr A;Schultz T;Welchowski T;Breuer J;Herberg U	Software Impacts	Refcurv a software for the construction of pediatric reference curves	2020

254	Heude B;Scherdel P;Werner A;Le G M;Gelbert N;Walther D;Arnould M;Bellaïche M;Chevallier B;Cheymol J;Jobez E;N'guyen S;Pietrement C;Reynaud R;Salaün J;Khoshnood B;Zeitlin J;Maccario J;Breart G;Thalabard J;Charles M;Botton J;Frاندji B;Chalumeau M	The Lancet Digital Health	A bigdata approach to producing descriptive anthropometric references a feasibility and validation study of paediatric growth charts	2019
255	Shiau J;Wu P	Asia-Pacific Journal Of Atmospheric Sciences	Nonstationary distributional changes of annual rainfall indices in Taiwan	2021
256	Casals P;Fernández J;Batet A;Taüll M;Ruiz-Olmo J	European Journal Of Wildlife Research	Disentangling landscape features and vegetation composition effects on wild ungulate use of hay meadows in high pyrenees	2021
257	He C;Chen F;Long A;Luo C;Qiao C	Water (Switzerland)	Frequency analysis of snowmelt flood based on gamlss model in manas river basin china	2021
258	Hossein F;Momayyezi M;Masoud M	Journal Of Nutrition And Food Security	Reference values for serum lipid profiles in Iranian adults a splinebased quantile regression method	2021
259	Groll A;Hambuckers J;Kneib T;Umlauf N	Computational Statistics And Data Analysis	Lassotype penalization in the framework of generalized additive models for location scale and shape	2019
260	Hohberg M;Pütz P;Kneib T	Plos One	Treatment effects beyond the mean using distributional regression methods and guidance	2020
261	Udy J;Wing S;Jowett T;O'connell-Milne S;Durante L;Mcmullin R;Kolodzey S	Frontiers In Psychology	Percentiles and principal component analysis of physical fitness from a big sample of children and adolescents aged 6-18 years the dafis project	2021
262	Wen Q;Sun P;Zhang Q;Li H	Water (Switzerland)	Nonstationary ecological instream flow and relevant causes in the huai river basin china	2021
263	Brockhaus S;Rügamer D;Greven S	Journal Of Statistical Software	Boosting functional regression models with fdboost	2020
264	Nevill A;Myers J;Kaminsky L;Arena R;Myers T	Erj Open Research	Comparing individual and population differences in minute ventilationcarbon dioxide production slopes using centile growth curves and loglinear allometry	2021
265	Shypailo R;Wong W	American Journal Of Clinical Nutrition	Fat and fatfree mass index references in children and young adults assessments along racial and ethnic lines	2020
266	Sodja C	Journal Of Computational And Graphical Statistics	Detecting anomalous time series by gamlss Akaike weights scoring	2021
267	Hasan E;Tarhule A;Kirstetter P	Remote Sensing	Twentieth and twentyfirst century water storage changes in the Nile river basin from gracegracefo and modeling	2021
268	Ramires T;Nakamura L;Righetto A;Carvalho R;Vieira L;Pereira C	Entropy	Comparison between highly complex location models and gamlss	2021
269	Herberg U;Smit F;Winkler C;Dalla-Pozza R;Breuer J;Laser K	Quantitative Imaging In Medicine And Surgery	Realtime 3dechocardiography of the right ventriclepaediatric reference values for right ventricular volumes using knowledgebased reconstruction a multicentre study	2021

270	Ramos A;Alonso J;Berra T;Alves L;Martoreli J J;Santos F;Alves Y;Andrade H;Costa F;Crispim J;Yamamura M;Alves J;Santos N M;Fuentealba-Torres M;Pinto I;Arcêncio R	Acta Tropica	Social inequalities and their association with the leprosy burden in a Brazilian city of low endemicity an ecological study	2021
271	Agarwal D;Parker R;Pinnock H;Roy S;Ghorpade D;Salvi S;Khatavkar P;Juvekar S	European Respiratory Journal	Normal spirometry predictive values for the western Indian adult population	2020
272	Chen M;Papadikis K;Jun C	Journal Of Hydrology	An investigation on the nonstationarity of flood frequency across the UK	2021
273	Cajias M;Freudenreich P;Freudenreich A;Schäfers W	Journal Of Business Economics	Liquidity and prices a cluster analysis of the German residential real estate market	2020
274	Rauhut A	Zeitschrift Fur Anglistik Und Amerikanistik	Exploring the effect of conversion on the distribution of inflectional suffixes a multivariate corpus study	2021
275	Voigt M;Meyer-Kahrweg L;Landau-Crangle E;So H;Däbritz J;Rochow M;Kunze M;Rochow N	Journal Of Perinatal Medicine	Individualized birth length and head circumference percentile charts based on maternal body weight and height	2020
276	Zhang L;Lin J;Liang S;Sun J;Gao N;Wu Q;Zhang H;Liu H;Cheng X;Cao Y;Li Y	Frontiers In Pediatrics	Comparison of postnatal growth charts of singleton preterm and term infants using world health organization standards at 40160 weeks postmenstrual age a Chinese singlecenter retrospective cohort study	2021
277	Briseño S G;Hohberg M;Groll A;Kneib T	Journal Of The Royal Statistical Society. Series A: Statistics In Society	Flexible instrumental variable distributional regression	2020
278	Stöcker A;Brockhaus S;Schaffer S;Bronk B;Opitz M;Greven S	Statistical Modelling	Boosting functional response models for location scale and shape with an application to bacterial competition	2021
279	Bigirinama R;Ntaongo J;Batumbo D;Sam-Agudu N;Katoto P;Byamungu L;Karume K;Nachega J;Bompangue D	Tropical Medicine And International Health	Environmental and anthropogenic factors associated with increased malaria incidence in southkivu province democratic republic of the Congo	2020
280	Ribeiro T;Seidel E;Guerra R;Peña-Ramirez F;Silva M	Communications In Statistics Case Studies Data Analysis And Applications Journal Of European Real Estate Research	Soybean production value in the Rio Grande do Sul under the gamlss framework is there room for another hedonic model the advantages of the gamlss approach in real estate research	2021
281	Ramires T;Ortega E;Cordeiro G;Paula G;Hens N	Communications In Statistics: Simulation And Computation	New regression model with four regression structures and computational aspects	2018
282	Peluso A;Vinciotti V;Yu K	Journal Of The Royal Statistical Society. Series C: Applied Statistics	Discrete weibull generalized additive model an application to count fertility data	2019
283	Ricci R A;Raggi A P;Gasalla M	Fisheries Research	Cost structure and financial performance of marine commercial fisheries in the south brazil bight	2019
284	Wang Y;Zeng Q;Xiong F;Zeng Y	Asian Journal Of Andrology	Male external genitalia growth curves and charts for children and adolescents aged 0 to 17 years in chongqing china	2018
285	Panayi E;Peters G;Danielsson J;Zigrand J	Econometrics And Statistics	Designating market maker behaviour in limit order book markets	2018

286	Rashidi A;Kiani O;Heidarzadeh M;Imani B;Nematy M;Taghipour A;Sadr M;Norouzy A	Iranian Journal Of Pediatrics	Reference curves of birth weight length and head circumference for gestational age in Iranian singleton births	2018
287	Müller A;Correa M;Führ C;Padoin T;De Q D;Schmitt J	International Journal Of Biometeorology	Neotropical ferns community phenology climatic triggers in subtropical climate in araucaria forest	2019
288	Udy J;Wing S;Jowett T;O'connell-Milne S;Durante L;Mcmullin R;Kolodzey S	Ecosphere	Regional differences in kelp forest interaction chains are influenced by both diffuse and localized stressors	2019
289	Li J;Ma Q;Tian Y;Lei Y;Zhang T;Feng P	Natural Hazards	Flood scaling under nonstationarity in daqinghe river basin china	2019
290	Gu X;Zhang Q;Li J;Singh V;Sun P	Journal Of Hydrology	Impact of urbanization on nonstationarity of annual and seasonal precipitation extremes in china	2019
291	Sun P;Wen Q;Zhang Q;Singh V;Sun Y;Li J	Journal Of Hydrology	Nonstationaritybased evaluation of flood frequency and flood risk in the huai river basin china	2018
292	Stasinopoulos D M;Rigby R A;De Bastiani	Statistical Modelling	Gamlss a distributional regression approach	2018
293	Thomas J;Mayr A;Bischl B;Schmid M;Smith A;Hofner B	Statistics And Computing	Gradient boosting for distributional regression faster tuning and improved variable selection via noncyclical updates	2018
294	Alizadeh M;Ramires T;Mirmostafae S;Samizadeh M;Ortega E	Communications In Statistics: Simulation And Computation	A new useful fourparameter extension of the gumbel distribution properties regression model and applications using the gamlss framework	2019
295	Samadi S;Tufford D;Carbone G	Stochastic Environmental Research And Risk Assessment	Estimating hydrologic model uncertainty in the presence of complex residual error structures	2018
296	Land P;Bailey T;Taberner M;Pardo S;Sathyendranath S;Zenouz K;Brammall V;Shutler J;Quartly G	Remote Sensing	A statistical modeling framework for characterising uncertainty in large datasets application to ocean colour	2018
297	Schlosser L;Hothorn T;Stauffer R;Zeileis A	Annals Of Applied Statistics	Distributional regression forests for probabilistic precipitation forecasting in complex terrain	2019
298	Wojtyś M;Marra G;Radice R	Computational Statistics And Data Analysis	Copula based generalized additive models for location scale and shape with nonrandom sample selection	2018
299	Muyumba E;Nkulu D;Mukeng C;Musung J;Kakoma P;Kakisingi C;Luboya O;Malonga F;Kizonde J;Mukuku O;Yan W	Bmc Cardiovascular Disorders	Oscillometric blood pressure by age and height for non-overweight children and adolescents in Lubumbashi democratic republic of Congo	2018
300	Yamada G;Castillo-Salgado C;Jones-Smith J;Moulton L	Public Health Nutrition	Differences in magnitude and rate of change in adult obesity distribution by age and sex in Mexico Colombia and Peru 20052010	2019
301	Su C;Chen X	Advances In Water Resources	Assessing the effects of reservoirs on extreme flows using nonstationary flood frequency models with the modified reservoir index as a covariate	2019
302	Li M;Zhang T;Feng P	Hydrological Processes	A nonstationary runoff frequency analysis for future climate change and its uncertainties	2019
303	Correia R;Ruete A;Stropp J;Malhado A;Dos S J;Lessa T;Alves J;Ladle R	Ecological Indicators	Using ignorance scores to explore biodiversity recording effort for multiple taxa in the caatinga	2019
304	De J T;Moshkin Y;Guryev V	Physiological Genomics	Gene expression variability the other dimension in transcriptome analysis	2019
305	Hambuckers J;Kneib T;Langrock R;Silbersdorff A	Quantitative Finance	A markovswitching generalized additive model for compound poisson processes with applications to operational loss models	2018

306	Gao L.;Huang J;Chen X;Chen Y;Liu M	Atmospheric Research	Contributions of natural climate changes and human activities to the trend of extreme precipitation	2018
307	Sun P;Zhang Q;Gu X;Shi P;Singh V;Song C;Zhang X	International Journal Of Disaster Risk Science	Nonstationarities and atsite probabilistic forecasts of seasonal precipitation in the east river basin China	2018
308	Doerstling S;Hedberg P;Öhrvik J;Leppert J;Henriksen E	Upsala Journal Of Medical Sciences	Growth differentiation factor 15 in a community based sample age dependent reference limits and prognostic impact	2018
309	Rashid M;Beecham S	Science Of The Total Environment	Development of a nonstationary standardized precipitation index and its application to a south Australian climate	2019
310	Yu J;Kim T;Park D	Water (Switzerland)	Future hydrological drought risk assessment based on nonstationary joint drought management index	2019
311	Righetto A;Ramires T;Nakamura L;Castanho P;Faes C;Savian T	Journal Of Applied Statistics	Predicting weed invasion in a sugarcane cultivar using multispectral image	2019
312	Zhao X;Xia Y;Zhang H;Baker P;Norris T	Bmc Pediatrics	Birth weight charts for a chinese population an observational study of routine newborn weight data from chongqing	2019
313	Umlauf N;Klein N;Zeileis A	Journal Of Computational And Graphical Statistics	Bamlss bayesian additive models for location scale and shape and beyond	2018
314	Voncken L;Albers C;Timmerman M	Behavior Research Methods	Improving confidence intervals for normed test scores include uncertainty due to sampling variability	2019
315	Rashid M;Beecham S	Hydrological Sciences Journal	Simulation of streamflow with statistically downscaled daily rainfall using a hybrid of wavelet and gamlss models	2019
316	Jacinto T;Amaral R;Malinovsky A;Janson C;Fonseca J;Alving K	Journal Of Applied Physiology	Exhaled no reference limits in a large populationbased sample using the lambdamusigma method	2018
317	Wanke P;Chen Z;Liu W;Antunes J;Azad M	Habitat International	Investigating the drivers of railway performance evidence from selected Asian countries	2018
318	Chamitava L;Garcia-Larsen V;Cazzoletti L;Degan P;Pasini A;Bellisario V;Corsico A;Nicolis M;Olivieri M;Pirina P;Ferrari M;Stasinopoulos D M;Zanolin M	Plos One	Determination of adjusted reference intervals of urinary biomarkers of oxidative stress in healthy adults using gamlss models	2018
319	Spencer P;Sanders K;Judge D	American Journal Of Physical Anthropology	Growth curvess and the international standard how childrens growth reflects challenging conditions in rural Timor-Leste	2018
320	Zhu T;Yip B;Hung V;Choy C;Cheng K;Kwok T;Cheng J;Qin L	Journal Of Bone And Mineral Research	Normative standards for hrpqt parameters in Chinese men and women	2018
321	Alderson L;Joksaitte S;Kemp J;Main E;Watson T;Platt F;Cortina-Borja M	Archives Of Disease In Childhood	Agerelated gait standards for healthy children and young people the gosich paediatric gait centiles	2019
322	Poon L;Sahota D;Chaemsaitong P;Nakamura T;Machida M;Naruse K;Wah Y;Leung T;Pooh R	Ultrasound In Obstetrics And Gynecology	Transvaginal threedimensional ultrasound assessment of sylvian fissures at 1830 weeks gestation	2019
323	Zou L;Xia J;Ning L;She D;Zhan C	Water (Switzerland)	Identification of hydrological drought in eastern china using a timedependent drought index	2018
324	Barbini S;Sabadin D;Lucifora L	Reviews In Fish Biology And Fisheries	Comparative analysis of feeding habits and dietary niche breadth in skates the importance of body size snout length and depth	2018

325	Anne-Isola N K;Weldon A;Imron M;Maynard K;Nijman V;Poindexter S;Morcatty T	Toxins	Venom in furs facial masks as aposematic signals in a venomous mammal	2019
326	Smith E;Wilkinson K;Wyatt S;Vaish P	Journal Of Physical Education And Sport	Reference curves for a fitness battery developed for children ages 5-12 years in England	2019
327	Frank J;Castle M;Westfall J;Weiskittel A;Macfarlane D;Baral S;Radtko P;Pelletier G	Forestry	Variation in occurrence and extent of internal stem decay in standing trees across the eastern us and Canada evaluation of alternative modelling approaches and influential factors	2018
328	Hepp T;Schmid M;Mayr A	International Journal Of Biostatistics	Significance tests for boosted location and scale models with linear baselearners	2019
329	Marincioni V;Marra G;Altamirano-Medina H	Building And Environment	Development of predictive models for the probabilistic moisture risk assessment of internal wall insulation	2018
330	Voncken L;Albers C;Timmerman M	Assessment	Model selection in continuous test norming with gamlss	2019
331	Al-Qerem W;Hammad A;Gassar E;Al-Qirim R;Ling J	Expert Review Of Respiratory Medicine	Spirometry reference equations for an adult middle eastern population	2019
332	Vasiljevic M;Fuller G;Pilling M;Hollands G;Pechey R;Jebb S;Marteau T	Appetite	What is the impact of increasing the prominence of calorie labelling a stepped wedge randomised controlled pilot trial in worksite cafeterias	2019
333	Hao W;Shao Q;Hao Z;Ju Q;Baima W;Zhang D	International Journal Of Climatology	Nonstationary modelling of extreme precipitation by climate indices during rainy season in hanjiang river basin china	2019
334	Denuit M;Trufin J	Journal Of Computational And Applied Mathematics	Collective loss reserving with two types of claims in motor third party liability insurance	2018
335	Ho N;Li F;Wang S;Kuhn L	Bmc Bioinformatics	Metamicrobiome an r package for analysis of microbiome relative abundance data using zeroinflated beta gamlss and metaanalysis across studies using random effects models	2019
336	Nakamura L;Cerqueira P;Ramires T;Pescim R;Rigby R A;Stasinopoulos D M	Journal Of Applied Statistics	A new continuous distribution on the unit interval applied to modelling the points ratio of football teams	2019
337	De M E;De L R;De O R;Dantas L;Dos S C	Water (Switzerland)	Spacetime kriging of precipitation modeling the largescale variation with model gamlss	2019
338	Dunn L;Sherrell H;Bligh L;Alsolai A;Flatley C;Kumar S	Placenta	Reference centiles for maternal placental growth factor levels at term from a lowrisk population	2019
339	Alías-Hernández I;Galera-Martínez R;García-García E;Muñoz-Vico F;Vázquez L M;Olvera-Porcel M;Bonillo P A	Pediatric Diabetes	Insulinaemia and insulin resistance in caucasian general paediatric population aged 2 to 10 years associated risk factors	2018
340	De C L;Marsico E;Arnaud V;Coupé C	Plos One	Assessing sound symbolism investigating phonetic forms visual shapes and letter fonts in an implicit boubakiki experimental paradigm	2018
341	Gunapalasingham G;Frithioff-Bøjsøe C;Lund M;Hedley P;Fonvig C;Dahl M;Pedersen O;Christiansen M;Hansen T;Lausten-Thomsen U;Holm J	Scandinavian Journal Of Clinical And Laboratory Investigation	Reference values for fasting serum concentrations of thyroidstimulating hormone and thyroid hormones in healthy danishnorthern european white children and adolescents	2019
342	Sepanski R;Godambe S;Zaritsky A	Frontiers In Pediatrics	Pediatric vital sign distribution derived from a multicentered emergency department database	2018
343	Miao H;Yao F;Wu Y;Zhang X;He R;Li B;Zhao Q	Scientific Reports	Birth weight percentiles by sex and gestational age for twins born in southern china	2019

344	Coupé C	Frontiers In Psychology	Modeling linguistic variables with regression models addressing nongaussian distributions nonindependent observations and nonlinear predictors with random effects and generalized additive models for location scale and shape	2018
345	Bazrafshan J;Hejabi S	Water Resources Management	A nonstationary reconnaissance drought index nrldi for drought monitoring in a changing climate	2018
346	Silva-Alonso T;Iglesias-Pérez M;García-Soidán J	Journal Of Human Sport And Exercise	Percentile curves and reference values for 2000m rowing ergometer performance time in international rowers aged 1470 years	2018
347	Merker A;Neumeyer L;Hertel N;Grigelioniene G;Mohnike K;Hagenäs L	American Journal Of Medical Genetics, Part A	Development of body proportions in achondroplasia sitting height leg length arm span and foot length	2018
348	Wang J;Xu Y;Wang Y;Yuan J;Wang Q;Xiang J	Stochastic Environmental Research And Risk Assessment	Nonstationarity analysis of extreme water level in response to climate change and urbanization in the taihu basin china	2019
349	Zou L;Xia J;She D	Water Resources Management	Analysis of impacts of climate change and human activities on hydrological drought a case study in the wei river basin china	2018
350	Li J;Lei Y;Tan S;Bell C;Engel B;Wang Y	Water Resources Management	Nonstationary flood frequency analysis for annual flood peak and volume series in both univariate and bivariate domain	2018
351	Galjaard S;Ameje L;Lees C;Pexsters A;Bourne T;Timmerman D;Devlieger R	Biology Of Sex Differences	Sex differences in fetal growth and immediate birth outcomes in a lowrisk caucasian population	2019
352	Dong Q;Zhang X;Lall U;Sang Y;Xie P	Hydrological Sciences Journal	An improved nonstationary model for flood frequency analysis and its implication for the three gorges dam china	2019
353	Zhang T;Wang Y;Wang B;Tan S;Feng P	Water (Switzerland)	Nonstationary flood frequency analysis using univariate and bivariate timevarying models based on gamlss	2018
354	Hajmohammadi H;Marra G;Heydecker B	Transportation Research Part D: Transport And Environment	Datadriven models for microscopic vehicle emissions	2019
355	Chang S;Tsai H;Tzeng J;Yeh K;Chen L;Lai S;Liao S;Hua M;Tsai M;Huang J;Yao T	World Allergy Organization Journal	Reference equations for spirometry in healthy asian children aged 5 to 18 years in taiwan	2019
356	Stander J;Dalla V L;Taglioni C;Liseo B;Wade A;Cortina-Borja M	Statistics In Medicine	Analysis of paediatric visual acuity using bayesian copula models with sinharcsinh marginal densities	2019
357	Balder J;Lansberg P;Hof M;Wiegman A;Hutten B;Kuivenhoven J	Journal Of Clinical Lipidology	Pediatric lipid reference values in the general population the dutch lifelines cohort study	2018
358	SáíA;Turkman M;Pereira J	Modeling Earth Systems And Environment	Exploring fire incidence in portugal using generalized additive models for location scale and shape gamlss	2018
359	Ramires T;Ortega E;Hens N;Cordeiro G;Paula G	Journal Of Applied Statistics	A flexible semiparametric regression model for bimodal asymmetric and censored data	2018
360	Barceló M;Van B P;Soudzilovskaia N	Journal Of Ecology	Climate drives the spatial distribution of mycorrhizal host plants in terrestrial ecosystems	2019
361	De Bastiani;Rigby Ra;Stasinopoulous D;Cysneiros A;Uribe-Opazo M	Journal Of Applied Statistics	Gaussian markov random field spatial models in gamlss	2018
362	Ötting M;Langrock R;Deutscher C	Statistical Modelling	Integrating multiple data sources in matchfixing warning systems	2018

363	Brockhaus S;Fuest A;Mayr A;Greven S	Journal Of The Royal Statistical Society. Series C: Applied Statistics	Signal regression models for location scale and shape with an application to stock returns	2018
364	Norris T;Seaton S;Manktelow B;Baker P;Kurinczuk J;Field D;Draper E;Smith L	Archives Of Disease In Childhood: Fetal And Neonatal Edition	Updated birth weight centiles for england and wales	2018
365	Bae Y;Zeidler R;Baber R;Vogel M;Wirkner K;Loeffler M;Ceglarek U;Kiess W;Körner A;Thiery J;Kratzsch J	Journal Of Steroid Biochemistry And Molecular Biology	Reference intervals of nine steroid hormones over the lifespan analyzed by lcmsms effect of age gender puberty and oral contraceptives	2019
366	Ramires T;Nakamura L;Righetto A;Pescim R;Mazucheli J;Cordeiro G	Journal Of Applied Statistics	A new semiparametric weibull cure rate model fitting different behaviors within gamlss	2019
367	Raut D;Raje D;Dandge V;Singh D	Indian Journal Of Radiology And Imaging	Percentile reference curves for normal pancreatic dimensions in indian children	2018
368	Smith A;Hofner B;Lamb J;Osenkowski J;Allison T;Sadoti G;Mcwilliams S;Paton P	Ecology And Evolution	Modeling spatiotemporal abundance of mobile wildlife in highly variable environments using boosted gamlss hurdle models	2019
369	Merker A;Neumeyer L;Hertel N;Grigelioniene G;Mäkitie O;Mohnike K;Hagenäs L	American Journal Of Medical Genetics, Part A	Growth in achondroplasia development of height weight head circumference and body mass index in a european cohort	2018
370	Soysal P;Perera G;Isik A;Onder G;Petrovic M;Cherubini A;Maggi S;Shetty H;Molokhia M;Smith L;Stubbs B;Stewart R;Veronese N;Mueller C	Experimental Gerontology	The relationship between polypharmacy and trajectories of cognitive decline in people with dementia a large representative cohort study	2019
371	Murillo-Posada J;Salas S;Velázquez-Abunader I	Peerj	Factors affecting relative abundance of lowmobility fishing resources spiny lobster in the galapagos marine reserve	2019
372	Ramires T;Hens N;Cordeiro G;Ortega O;Guedes T;Gonçalves-Zuliani A;Nunes W	Computational Statistics Acta Scientiarum - Biological Sciences	Estimating nonlinear effects in the presence of cure fraction using a semi-parametric regression model	2018
373	Rashid M;Beecham S;Chowdhury R	Theoretical And Applied Climatology	Statistical downscaling of rainfall a nonstationary and multiresolution approach	2016
374	Buchholz M;Holst G;Musshoff O	Water Resources Research	Irrigation water policy analysis using a business simulation game	2016
375	Argyropoulos C;Etheridge A;Sakhanenko N;Galas D	Nucleic Acids Research	Modeling bias and variation in the stochastic processes of small rna sequencing	2017
376	Sankilampi U;Hannila M;Saari A;Gissler M;Dunkel L	Annals Of Medicine	New populationbased references for birth weight length and head circumference in singletons and twins from 23 to 43 gestation weeks	2013
377	Machado M;Botero B;López J;Francés F;Diez-Herrero A;Benito G	Hydrology and Earth System Sciences (An interactive open-access journal of the European Geosciences Union)	Flood frequency analysis of historical flood data under stationary and non-stationary modelling	2015
378	Hothorn T;Kneib T;Bühlmann P	Journal Of The Royal Statistical Society. Series B: Statistical Methodology	Conditional transformation models	2014
379	Debele S;Bogdanowicz E;Strupczewski W	Acta Geophysica	Around and about an application of the gamlss package to nonstationary flood frequency analysis	2017

380	Brognez H;Fallourd R;Mallet C;Siviraa R;Dufour C	Journal Of Atmospheric And Oceanic Technology	Estimating confidence intervals around relative humidity profiles from satellite observations application to the saphir sounder	2016
381	Urlacher S;Blackwell A;Liebert M;Madimenos F;Cepon-Robins T;Gildner T;Snodgrass J;Sugiyama L	American Journal Of Human Biology	Physical growth of the shuar height weight and bmi references for an indigenous amazonian population	2016
382	Peplies J;Jiménez-Pavón D;Savva S;Buck C;Günther K;Fraterman A;Russo P;Iacoviello L;Veidebaum T;Tornaritis M;De H S;Márlid S;Molnár D;Moreno L;Ahrens W	International Journal Of Obesity	Percentiles of fasting serum insulin glucose hba1c and homair in prepubertal normal weight european children from the idefics cohort	2014
383	Landi M;Zoccola A;Bacaro G;Angiolini C	Plant Species Biology	Phenology of dryopteris affinis ssp affinis and polystichum aculeatum modeling relationships to the climatic variables in a mediterranean area	2014
384	Levin D;Marryat L;Cole T;Mccoll J;Harjunmaa U;Ashorn P;Wright C	Archives Of Disease In Childhood	Fit to who weight standard of european infants over time	2016
385	De M P;Gracia-Marco L;Ortega F;Intemann T;Foraita R;Lissner L;Oja L;Barba G;Michels N;Tornaritis M;Molnár D;Pitsiladis Y;Ahrens W;Moreno L	International Journal Of Obesity	Physical fitness reference standards in european children the idefics study	2014
386	Wang Y;Li J;Feng P;Hu R	Water Resources Management	A timedependent drought index for nonstationary precipitation series	2015
387	Debele S;Strupczewski W;Bogdanowicz E	Acta Geophysica	A comparison of three approaches to nonstationary flood frequency analysis	2017
388	Kohler L;Silverstein J;Rajagopalan B	Environmental Engineering Science	Predicting life cycle failures of onsite wastewater treatment systems using generalized additive models	2016
389	Gjessing H;Grøttum P;Økland I;Eik-Nes S	Ultrasound In Obstetrics And Gynecology	Fetal size monitoring and birthweight prediction a new population-based approach	2017
390	Barba G;Buck C;Bammann K;Hadjigeorgiou C;Hebestreit A;Márlid S;Molnár D;Russo P;Veidebaum T;Vyncke K;Ahrens W;Moreno L	International Journal Of Obesity	Blood pressure reference values for european no overweight school children the idefics study	2014
391	Greven S;Scheipl F	Statistical Modelling	A general framework for functional regression modelling	2017
392	Gheorghe M;Brouwer W;Van B P	Medical Decision Making	Quality of life and time to death have the health gains of preventive interventions been underestimated	2015
393	De J R;Van B S;Spiess M	Communications In Statistics: Simulation And Computation	Multiple imputation of predictor variables using generalized additive models	2016
394	Vanhelst J;Labreuche J;Béghin L;Drumez E;Fardy P;Chapelot D;Mikulovic J;Ulmer Z	Journal Of Strength And Conditioning Research	Physical fitness reference standards in french youth the bouge program	2017
395	Bongers B;De V S;Obeid J;Van B S;Helders P;Takken T	Physical Therapy	The steep ramp test in dutch white children and adolescents age and sexrelated normative values	2013

396	Garcia G S;Gimenez P;Giraldo-Osorio J	Water (Switzerland)	Assessing nonstationary spatial patterns of extreme droughts from longterm high-resolution observational dataset on a semiarid basin Spain	2015
397	Herrmann D;Intemann T;Lauria F;Märild S;Molnár D;Moreno L;Sioen I;Tornaritis M;Veidebaum T;Pigeot I;Ahrens W	International Journal Of Obesity	Reference values of bone stiffness index and cterminal telopeptide in healthy european children	2014
398	Burt L;Liang Z;Sajobi T;Hanley D;Boyd S	Journal Of Bone And Mineral Research	Sex and sitespecific normative data curves for hrpqt	2016
399	Stauffer R;Mayr G;Messner J;Umlauf N;Zeileis A	International Journal Of Climatology	Spatiotemporal precipitation climatology over complex terrain using a censored additive regression model	2017
400	López J;Francés F	Hydrology And Earth System Sciences	Nonstationary flood frequency analysis in continental spanish rivers using climate and reservoir indices as external covariates	2013
401	Cribari-Neto F;Lucena S	Journal Of Statistical Computation And Simulation	Nonnested hypothesis testing inference for gamlss models	2017
402	Zhang J;Hu X;Shan G	Respirology	Spirometry reference values for population aged 780 years in china	2017
403	Marra G;Radice R	Computational Statistics And Data Analysis	Bivariate copula additive models for location scale and shape	2017
404	Kainu A;Timonen K;Toikka J;Qaiser B;Pitkaniemi J;Kotaniemi J;Lindqvist A;Vanninen E;Länsimies E;Sovijärvi A	Clinical Physiology And Functional Imaging	Reference values of spirometry for finnish adults	2016
405	Wood S;Pya N;Säfken B	Journal Of The American Statistical Association	Smoothing parameter and model selection for general smooth models	2016
406	Mayer M;Gleiss A;Häusler G;Borkenstein M;Kapelari K;Köstl G;Lassi M;Schemper M;Schmitt K;Blümel P	Annals Of Human Biology	Weight and body mass index bmi current data for austrian boys and girls aged 4 to under 19 years	2015
407	Xiong L;Jiang C;Du T	Water Science And Technology	Statistical attribution analysis of the nonstationarity of the annual runoff series of the weihe river	2014
408	Rigby R A;Stasinopoulos D M	Statistical Methods In Medical Research	Automatic smoothing parameter selection in gamlss with an application to centile estimation	2014
409	Sezer A;Kan K B;Yazici B	Applied Ecology And Environmental Research	Modelling extreme rainfalls using generalized additive models for location scale and shape parameters	2016
410	Ramires T;Ortega E;Cordeiro G;Hens N	Journal Of Statistical Computation And Simulation	A bimodal flexible distribution for lifetime data	2016
411	Bertapelli F;Agiowlasitis S;Machado M;Rosso R;Guerra-Junior G	Journal Of Epidemiology	Growth charts for Brazilian children with down syndrome birth to 20 years of age	2017
412	Cardoso F;Almodhy M;Pepera G;Stasinopoulos D M;Sandercock G	Journal Of Sports Sciences	Reference values for the incremental shuttle walk test in patients with cardiovascular disease entering exercisebased cardiac rehabilitation	2017
413	Bonander C;Andersson R;Nilson F	Accident Analysis And Prevention	The effect of stricter licensing on road traffic injury events involving 15 to 17yearold moped drivers in Sweden a time series intervention study	2015
414	Kuhle S;Ashley-Martin J;Maguire B;Hamilton D	Peerj	Percentile curves for skinfold thickness for Canadian children and youth	2016

415	Wagner-Riddle C;Werner S;Caramori P;Ricce W;Nitsche P;Von B P;De S E	International Journal Of Biometeorology	Determining the influence of itaipu lake on thermal conditions for soybean development in adjacent lands	2015
416	Zhang R;Chen X;Wang H;Cheng Q;Zhang Z;Shi P	Hydrological Processes	Temporal change of spatial heterogeneity and its effect on regional trend of annual precipitation heterogeneity indices	2017
417	Chen H;Kelly C;Xavier C F;He Y;Zuo X;Reiss P	Neuroimage	Quantile rank maps a new tool for understanding individual brain development	2015
418	De H S;Michels N;Vyncke K;Hebestreit A;Russo P;Intemann T;Peplies J;Frateman A;Eiben G;De L M;Tornaritis M;Molnar D;Veidebaum T;Ahrens W;Moreno L	International Journal Of Obesity	Blood lipids among young children in europe results from the european idefics study	2014
419	Ferrara G;Vidoli F	European Journal Of Operational Research	Semiparametric stochastic frontier models a generalized additive model approach	2017
420	Zhang D;Yan D;Wang Y;Lu F;Liu S	Natural Hazards	Gamlssbased nonstationary modeling of extreme precipitation in beijingtianjinhebei region of china	2015
421	Green P	Statistical Modelling	Discussion of beyond mean regression	2013
422	Cuny H;Kiessé T	Communications In Statistics: Simulation And Computation	On modeling wood formation using parametric and semiparametric regressions for count data	2016
423	Klein N;Denuit M;Lang S;Kneib T	Insurance: Mathematics And Economics	Nonlife ratemaking and risk management with bayesian generalized additive models for location scale and shape	2014
424	Bisiacchi P;Cona G;Tarantino V;Schiff S;Montagnese S;Amodio P;Capizzi G	Metabolic Brain Disease	Assessing inter and intraindividual cognitive variability in patients at risk for cognitive impairment the case of minimal hepatic encephalopathy	2014
425	Kiose D;Voudouris V	Expert Systems With Applications	The acewem framework an integrated agentbased and statistical modelling laboratory for repeated power auctions	2015
426	Ertl D;Gleiss A;Sagmeister S;Haeusler G	Wiener Medizinische Wochenschrift	Determining the normal range for igfi igfbp3 and als new reference data based on current internal standards serumbestimmung von igfi igfbp3 und als neue referenzwerte laut aktuellen internen standards	2014
427	Gao L;Huang J;Chen X;Chen Y;Liu M	Journal Of Hydrometeorology	Risk of extreme precipitation under nonstationarity conditions during the second flood season in the southeastern coastal region of china	2017
428	Nagy P;Kovacs E;Moreno L;Veidebaum T;Tornaritis M;Kourides Y;Siani A;Lauria F;Sioen I;Claessens M;Mårild S;Lissner L;Bammann K;Intemann T;Buck C;Pigeot I;Ahrens W;Molnár D	International Journal Of Obesity	Percentile reference values for anthropometric body composition indices in european children from the idefics study	2014
429	Premkumar P;Antonisamy B;Mathews J;Benjamin S;Regi A;Jose R;Kuruville A;Mathai M	Bmc Pregnancy And Childbirth	Birth weight centiles by gestational age for twins born in South India	2016
430	López-De L C J;Francés F	Tecnologia Y Ciencias Del Agua	Lowfrequency climate variability in the nonstationary modeling of flood regimes in the sinaloa and presidio san pedro hydrologic regions	2014
431	Caroni C;Panagoulia D	Revstat Statistical Journal	Nonstationary modelling of extreme temperatures in a mountainous area of Greece	2016

432	Zararsiz G; Çiçek B;Kondolot M; Mazicioğlu M;Öztürk A;Kurtoglu S	Jcrpe Journal Of Clinical Research In Pediatric Endocrinology	Comparison of updated weight and height percentiles with previous references in 617yearold children in kayseri turkey	2017
433	Luo J;Chen L;Liu H	Physica A: Statistical Mechanics And Its Applications	Distribution characteristics of stock market liquidity	2013
434	Hofner B;Mayr A;Schmid M	Journal Of Statistical Software	Gamboostlss an r package for model building and variable selection in the gamlss framework	2016
435	Gu X;Zhang Q;Singh V;Chen X;Liu L	Global And Planetary Change	Nonstationarity in the occurrence rate of floods in the tarim river basin china and related impacts of climate indices	2016
436	Durán P;Merker A;Briceño G;Colón E;Line D;Abad V;Del T K;Chahin S;Matallana A;Lema A;Llano M;Céspedes J;Hagenäs L	Acta Paediatrica, International Journal Of Paediatrics	Colombian reference growth curves for height weight body mass index and head circumference	2016
437	Georgiadis L;Keski-Nisula L;Harju M;Räisänen S;Georgiadis S;Hannila M;Heinonen S	Placenta	Umbilical cord length in singleton gestations a finnish populationbased retrospective register study	2014
438	Rawal D;Kasel S;Keatley M;Nitschke C	Agricultural And Forest Meteorology	Climatic and photoperiodic effects on flowering phenology of select eucalypts from southeastern australia	2015
439	Li J;Tan S	Water Resources Management	Nonstationary flood frequency analysis for annual flood peak series adopting climate indices and check dam index as covariates	2015
440	Gu X;Zhang Q;Singh V;Shi P	International Journal Of Climatology	Nonstationarities in the occurrence rate of heavy precipitation across china and its relationship to climate teleconnection patterns	2017
441	He J;Xia H;Liu Y;Xia X;Mo W;Wang P;Cheng K;Leung G;Feng Q;Schooling C;Qiu X	Archives Of Disease In Childhood: Education And Practice Edition	A new birthweight reference in guangzhou southern china and its comparison with the global reference	2014
442	Tamminen S;Juutilainen I;Röning J	Expert Systems With Applications	Exceedance probability estimation for a quality test consisting of multiple measurements	2013
443	Cromwell E;Dube Q;Cole S;Chirambo C;Dow A;Heyderman R;Van R A	European Journal Of Paediatric Neurology	Validity of us norms for the bayley scales of infant developmentiii in malawian children	2014
444	Zhang Q;Gu X;Singh V;Xiao M;Xu C	Journal Of Hydrology	Stationarity of annual flood peaks during 19512010 in the pearl river basin china	2014
445	Mayr A;Schmid M;Pfahlberg A;Uter W;Gefeller O	Statistical Methods In Medical Research	A permutation test to analyse systematic bias and random measurement errors of medical devices via boosting location and scale models	2017
446	Singh J;Vittal H;Karmakar S;Ghosh S;Niyogi D	Geophysical Research Letters	Urbanization causes nonstationarity in indian summer monsoon rainfall extremes	2016
447	Antunes A;Maia J;Stasinopoulos D M;Gouveia R;Thomis M;Lefevre J;Teixeira A;Freitas D	American Journal Of Human Biology	Gross motor coordination and weight status of portuguese children aged 614 years	2015
448	Blackwell A;Urlacher S;Beheim B;Von R C;Jaeggi A;Stieglitz J;Trumble B;Gurven M;Kaplan H	American Journal Of Physical Anthropology	Growth references for tsimane foragerhorticulturalists of the bolivian amazon	2017
449	Juutilainen I;Tamminen S;Röning J	Communications In Statistics - Theory And Methods	Visualizing predicted and observed densities jointly with beanplot	2015

450	Tinggaard J;Aksglaede L;Sørensen K;Mouritsen A;Wohlfahrt-Veje C;Hagen C;Mieritz M;Jørgensen N;Wolthers O;Heuck C;Petersen J;Main K;Juul A	Acta Paediatrica, International Journal Of Paediatrics	The 2014 danish references from birth to 20 years for height weight and body mass index	2014
451	Stanojevic S;Graham B;Cooper B;Thompson B;Carter K;Francis R;Hall G	European Respiratory Journal	Official ers technical standards global lung function initiative reference values for the carbon monoxide transfer factor for caucasians	2017
452	Li H;Zhang Q;Singh V;Shi P;Sun P	Journal Of Hydrology	Hydrological effects of cropland and climatic changes in arid and semiarid river basins a case study from the yellow river basin china	2017
453	Wahl S;Fenske N;Zeilinger S;Suhre K;Gieger C;Waldenberger M;Grallert H;Schmid M	Bmc Bioinformatics	On the potential of models for location and scale for genomewide dna methylation data	2014
454	Lausten-Thomsen U;Christiansen M;Louise H P;Esmann F C;Stjernholm T;Pedersen O;Hansen T;Holm J	Scandinavian Journal Of Clinical And Laboratory Investigation	Reference values for serum leptin in healthy nonobese children and adolescents	2016
455	Wood S;Fasiolo M	Biometrics	A generalized fellnerschall method for smoothing parameter optimization with application to tweedie location scale and shape models	2017
456	Thomas A;Hanel B;Marott J;Buchvald F;Mortensen J;Nielsen K	Plos One	The singlebreath diffusing capacity of co and no in healthy children of european descent	2014
457	Nielsen T;Lausten-Thomsen U;Fonvig C;Bøjsøe C;Pedersen L;Bratholm P;Hansen T;Pedersen O;Holm J	Bmc Pediatrics	Dyslipidemia and reference values for fasting plasma lipid concentrations in danishnortheuropean white children and adolescents	2017
458	Hossain A;Rigby R A;Stasinopoulos D M;Enea M	Statistics In Medicine	Centile estimation for a proportion response variable	2016
459	Kainu A;Timonenb K	Scandinavian Journal Of Clinical And Laboratory Investigation	Polynomial estimation of the smoothing splines for the new finnish reference values for spirometry	2016
460	Öztürk A;Çiçek B;Mazicioğlu M;Zararsiz G;Kurtoglu S	Jcrpe Journal Of Clinical Research In Pediatric Endocrinology	Wrist circumference and frame size percentiles in 617yearold turkish children and adolescents in kayseri	2017
461	Rita A;Borghetti M;Todaro L;Saracino A	Frontiers In Plant Science	Interpreting the climatic effects on xylem functional traits in two mediterranean oak species the role of extreme climatic events	2016
462	Campitelli G;Macbeth G;Ospina R;Marmolejo-Ramos F	Educational And Psychological Measurement	Three strategies for the critical use of statistical methods in psychological research	2017
463	Addo O;Himes J;Zemel B	American Journal Of Clinical Nutrition	Reference ranges for midupper arm circumference upper arm muscle area and upper arm fat area in us children and adolescents aged 120 y	2017
464	Wang J;Luo S	Biometrical Journal	Augmented beta rectangular regression models a bayesian perspective	2016
465	Colloca F;Enea M;Ragonese S;Di Lorenzo M	Aquatic Conservation: Marine And Freshwater Ecosystems Statistical Modelling	A century of fishery data documenting the collapse of smooth-hounds (Mustelus spp.) in the Mediterranean Sea	2017

466	Rigby R A;Stasinopoulos D M;Lane P	Journal Of The Royal Statistical Society. Series C: Applied Statistics	Generalized additive models for location scale and shape	2005
467	Beyerlein A;Fahrmeir L;Mansmann U;Toschke A	Bmc Medical Research Methodology	Alternative regression models to assess increase in childhood bmi	2008
468	Karvonen M;Hannila M;Saari A;Dunkel L	Annals Of Medicine	New finnish reference for head circumference from birth to 7 years	2012
469	Verschuren O;Bloemen M;Kruitwagen C;Takken T	Developmental Medicine And Child Neurology	Reference values for anaerobic performance and agility in ambulatory children and adolescents with cerebral palsy	2010
470	Villarini G;Serinaldi F;Smith J;Krajewski W	Water Resources Research	On the stationarity of annual flood peaks in the continental united states during the 20th century	2009
471	Kumar V;Jeyaseelan L;Sebastian T;Regi A;Mathew J;Jose R	Bmc Pregnancy And Childbirth	New birth weight reference standards customised to birth order and sex of babies from south india	2013
472	Sandercock G;Voss C;Cohen D;Taylor M;Stasinopoulos D M	Journal Of Sports Sciences	Centile curves and normative values for the twenty metre shuttlerun test in english schoolchildren	2012
473	Koopman M;Zanen P;Kruitwagen C;Van D E C;Arets H	Respiratory Medicine	Reference values for paediatric pulmonary function testing the utrecht dataset	2011
474	Verschuren O;Bloemen M;Kruitwagen C;Takken T	Physical Therapy	Reference values for aerobic fitness in children adolescents and young adults who have cerebral palsy and are ambulatory	2010
475	Rudge J;Gilchrist R	Energy And Buildings	Measuring the health impact of temperatures in dwellings investigating excess winter morbidity and cold homes in the london borough of newham	2007
476	Serinaldi F	Energy Economics	Distributional modeling and shortterm forecasting of electricity prices by generalized additive models for location scale and shape	2011
477	Voudouris V;Gilchrist R;Rigby R A;Sedgwick J;Stasinopoulos D M	Journal Of Applied Statistics	Modelling skewness and kurtosis with the BCPE density in gamlss	2012
478	Cohen D;Voss C;Taylor M;Stasinopoulos D M;Delextrat A;Sandercock G	Acta Paediatrica, International Journal Of Paediatrics	Handgrip strength in english schoolchildren	2010
479	Majer I;Mackenbach J;Van B P	Statistics In Medicine	Time trends and forecasts of body mass index from repeated crosssectional data a different approach	2013
480	Quanjer P;Stocks J;Cole T;Hall G;Stanojevic S	European Respiratory Journal	Influence of secular trends and sample size on reference equations for lung function tests	2011
481	Villarini G;Smith J;Serinaldi F;Ntelekos A;Schwarz U	International Journal Of Climatology	Analyses of extreme flooding in Austria over the period 19512006	2012
482	Rochat M;Laubender R;Kuster D;Braendli O;Moeller A;Mansmann U;Von M E;Wildhaber J	Plos One	Spirometry reference equations for central european populations from school age to old age	2013
483	Gleiss A;Lassi M;Blümel P;Borkenstein M;Kapelari K;Mayer M;Schemper M;Häusler G	Annals Of Human Biology	Austrian height and body proportion references for children aged 4 to under 19 years	2013
484	Bonafide C;Brady P;Keren R;Conway P;Marsolo K;Daymont C	Pediatrics	Development of heart and respiratory rate percentile curves for hospitalized children	2013
485	Komlos J;Brabec M	Economics And Human Biology	The trend of bmi values of us adults by deciles birth cohorts 18821986 stratified by gender and ethnicity	2011

486	Villarini G;Smith J;Serinaldi F;Bales J;Bates P;Krajewski W	Advances In Water Resources	Flood frequency analysis for nonstationary annual peak records in an urban drainage basin	2009
487	Kenny R;Coen R;Frewen J;Donoghue O;Cronin H;Savva G	Journal Of The American Geriatrics Society	Normative values of cognitive and physical function in older adults findings from the irish longitudinal study on ageing	2013
488	Villarini G;Smith J;Napolitano F	Advances In Water Resources	Nonstationary modeling of a long record of rainfall and temperature over rome	2010
489	Saari A;Sankilampi U;Hannila M;Kiviniemi V;Kesseli K;Dunkel L	Annals Of Medicine	New finnish growth references for children and adolescents aged 0 to 20 years length height for age weight for length height and body mass index forage	2011
490	Vinh-Hung V;Joseph S;Coutty N;Ly B;Vlastos G;Nguyen N	Oncologist	Age and axillary lymph node ratio in postmenopausal women with t1t2 node positive breast cancer	2010
491	Bohl A;Blough D;Fishman P;Harris J;Phelan E	Health Services And Outcomes Research Methodology	Are generalized additive models for location scale and shape an improvement on existing models for estimating skewed and heteroskedastic cost data	2013
492	Villarini G;Serinaldi F	International Journal Of Climatology	Development of statistical models for atsite probabilistic seasonal rainfall forecast	2012
493	De C M;Cancho V;Rodrigues J	Computer Methods And Programs In Biomedicine	A handson approach for fitting longterm survival models under the gamlss framework	2010
494	Nielsen A;Olsen E;Juul A	Acta Paediatrica, International Journal Of Paediatrics	New danish reference values for height weight and body mass index of children aged 05 years	2010
495	Brooks J;Day S;Shavelle R;Strauss D	Pediatrics	Low weight morbidity and mortality in children with cerebral palsy new clinical growth charts	2011
496	Verbeek S;Eilers P;Lawrence K;Hennekam R;Versteegh F	European Journal Of Pediatrics	Growth charts for children with ellisvan creveld syndrome	2011
497	Wan X;Qu Y;Huang Y;Zhang X;Song H;Jiang H	Contemporary Clinical Trials	Nonparametric estimation of agespecific reference percentile curves with radial smoothing	2012
498	Kirstetter P;Viltard N;Gosset M	Quarterly Journal Of The Royal Meteorological Society	An error model for instantaneous satellite rainfall estimates evaluation of braintmi over west africa	2013
499	Van O F;Vervoort R;Heller G;Stasinopoulos D M;Rigby R A	Hydrology And Earth System Sciences	Longrange forecasting of intermittent streamflow	2011
500	Khondoker M;Glasbey C;Worten B	Biometrical Journal	A comparison of parametric and nonparametric methods for normalising cDNA microarray data	2007
501	Ganegoda A;Evans J	Annals Of Actuarial Science	A scaling model for severity of operational losses using generalized additive models for location scale and shape gamlss	2012
502	Douma-Van R D;Verschuren O;Jelsma D;Kruitwagen C;Smits-Engelsman B;Takken T	Pediatric Physical Therapy	Reference values for the muscle power sprint test in 6 to 12yearold children	2012
503	Rigby R A;Stasinopoulos D M;Voudouris V	Statistical Modelling	Discussion a comparison of gamlss with quantile regression	2013
504	Ibache-Pulgar G;Paula G;Cysneiros F	Test	Semiparametric additive models under symmetric distributions	2013
505	Serinaldi F;Cuomo G	Coastal Engineering	Characterizing impulsive waveindeck loads on coastal bridges by probabilistic models of impact maxima and rise times	2011
506	Newell M	Pediatric Infectious Disease Journal	Agerelated standards for total lymphocyte cd4 and cd8 t cell counts in children born in Europe	2005

507	Serinaldi F;Kilsby C	Journal Of Hydrology	A modular class of multisite monthly rainfall generators for water resource management and impact studies	2012
508	Boutselis P;Ringrose T	Expert Systems With Applications	Gamlss and neural networks in combat simulation metamodelling a case study	2013
509	Hernandez F;Hare J;Fey D	Fisheries Oceanography	Evaluating diel ontogenetic and environmental effects on larval fish vertical distribution using generalized additive models for location scale and shape	2009
510	Stasinopoulos D M;Rigby R A	Journal Of Statistical Software	Generalized additive models for location scale and shape gamlss in r	2007
511	Kneib T	Statistical Modelling	Beyond mean regression	2013
512	Giraldo O J;García G S	Journal Of Hydrology	Nonstationary analysis of dry spells in monsoon season of Senegal river basin using data from regional climate models rcms	2012
513	Cole T;Stanojevic S;Stocks J;Coates A;Hankinson J;Wade A	Statistics In Medicine	Age and sizerelated reference ranges a case study of spirometry through childhood and adulthood	2009
514	Van B S;Hayes D;Stasinopoulos D M;Rigby R A;Ter K F;Terlouw D	Statistics In Medicine	Estimating regional centile curves from mixed data sources and countries	2009
515	Correa J;Kneib T;Ospina R;Tejada J;Marmolejo-Ramos F	Quantitative Methods For Psychology	Assessing potential heteroscedasticity in psychological data a gamlss approach	2023
516	Salfran D;Spiess M	R Journal	Generalized additive model multiple imputation by chained equations with package impute robust	2018
517	Mayr A;Fenske N;Hofner B;Kneib T;Schmid M	Journal Of The Royal Statistical Society Series C- Applied Statistics	Generalized additive models for location scale and shape for high dimensional data a flexible approach based on boosting	2012
518	Ulbl M;Stembal R;Smodis M	Geodetski Vestnik	Preliminary model of rent mass appraisal for offices	2016
519	Fantazzini T;Da R S;Carvalho G;Liska G;De C M;Coelho S;Cirillo M;Ribeiro F	Seed Science And Technology	Correlation between historical data of the germination test and of the tetrazolium test in coffee seeds by gamlss	2020
520	Ramires T;Nakamura L;Righetto A;Konrath A;Pereira C	Stats	Incorporating clustering techniques into gamlss	2021
521	Costa E;Teixeira G;Freire F;Dias J;Fransozo A	Journal Of Sea Research	Effects of biological and environmental factors on the variability of iparalanchurus brasiliensis sciaenidae density an gamlss application	2022
522	Salarijazi M;Ghorbani K;Mohammadi M;Ahmadianfar I;Mohammadrezapour O;Naser M;Yaseen Z	Urban Climate	Spatialtemporal estimation of maximum temperature high returns periods for annual time series considering stationary nonstationary approaches in iran urban area	2023
523	Ramires T;Hens N;Cordeiro G;Ortega E	Computational Statistics	Estimating nonlinear effects in the presence of cure fraction using a semiparametric regression model	2018
524	Ramires T;Nakamura L;Righetto A;Ortega E;Cordeiro G	Cadernos De Saude Publica	Predicting survival function and identifying associated factors in patients with renal insufficiency in the metropolitan area of Maringa Parana State Brazil	2018
525	Dantas L;Dos S C;De O R;De B J;Santos C;Martins E;De O G;Brunsell N	Water	Rainfall prediction in the state of Paraiba northeastern brazil using generalized additive models	2020

526	Cajias M	Journal of European Real Estate Research	Is there room for another hedonic model? The advantages of the GAMLSS approach in real estate research	2018
527	Whittier D;Burt L;Hanley D;Boyd S	Journal Of Bone And Mineral Research	Sex and sitespecific reference data for bone microarchitecture in adults measured using second generation hrpqt	2020
528	Kiffner C;Lödige C;Alings M;Vor T;Rühe F	Medical And Veterinary Entomology	Body mass or sex biased tick parasitism in roe deer <i>icapreolus capreolus</i> a gamlss approach	2011
529	Klein N;Kneib T;Lang S	Journal Of The American Statistical Association	Bayesian generalized additive models for location scale and shape for zero inflated and overdispersed count data	2015
530	Pulcherio R;Rocha M;De B R;Brito E;Santos D;Mendonça A;Strüssmann C;Ferraz R	Arquivo Brasileiro De Medicina Veterinaria E Zootecnia	The influence of season environment sex and body mass on biochemical profile of the freshwater turtle <i>iphrynops geoffroanus</i> schweigger 1812	2022
531	Katsanevakis S;Issaris Y;Poursanidis D;Thessalou-Legaki M	Marine Environmental Research	Vulnerability of marine habitats to the invasive green alga <i>icaulerpa racemosai</i> var <i>icylindraceai</i> within a marine protected area	2010
532	Piekarska-Boniecka H;Siatkowski I;Zyprych J	Acta Scientiarum Polonorum-Hortorum Cultus	Model for bionomy of privet sawfly <i>imacrophya punctum</i> albulmi I hymenoptera tenthredinidae	2010
533	Granneman J;Baxley C;Cross Z;Levine E;Murphy R;Sharkey S;Williams K	Journal Of Shellfish Research	The association of bay scallops <i>iargopecten irradians concentricus</i> with seagrass and macroalgae along the florida gulf coast	2022
534	Colloca F;Enea M;Ragonese S;Di L M	Aquatic Conservation-Marine And Freshwater Ecosystems	A century of fishery data documenting the collapse of smoothhounds <i>imustelus</i> spp in the mediterranean sea	2017
535	Campoy L;Arroyo L;Ramos A;Berra T;Crispim J;Yamamura M;Pinto I;Monroe A;Scholze A;Andrade R;Alexander K;Fiorati R;De F G;Arcencio R	Journal Of Infection In Developing Countries	The complexity of <i>tbhiv</i> coinfection an analysis of the social and health services context in the state of sao paulo brazil	2020
536	Newman L;O'connor J;Nolan H;Reilly R;Kenny R	Experimental Gerontology	Age and sex related differences in orthostatic cerebral oxygenation findings from 2764 older adults in the irish longitudinal study on ageing tilda	2022
537	El B H;Fa J;Bowler M;Valsecchi J;Bodmer R;Mayor P	Journal Of Mammalogy	Breeding seasonality in the lowland paca <i>icuniculus pacai</i> in amazonia interactions with rainfall fruiting and sustainable hunting	2018
538	Umlauf N;Klein N;Simon T;Zeileis A	Journal Of Statistical Software	Bamlss a lego toolbox for flexible bayesian regression and beyond	2021
539	Yu F;Xu Y;Hou Y;Lin Y;Jiajue R;Jiang Y;Wang O;Li M;Xing X;Zhang L;Qin L;Hsieh E;Xia W	Journal Of Bone And Mineral Research	Age site and sex specific normative centile curves for hrpqt derived microarchitectural and bone strength parameters in a chinese mainland population	2020
540	Stefanidis K;Varlas G;Vourka A;Papadopoulos A;Dimitriou E	Water Research	Delineating the relative contribution of climate related variables to chlorophylla and phytoplankton biomass in lakes using the erasland climate reanalysis data	2021
541	Cantafaro A;Ardizzone G;Enea M;Ligas A;Colloca F	Ecological Indicators	Assessing the importance of nursery areas of european hake <i>imerluccius merluccius</i> using a body condition index	2017
542	Kac G;Carilho T;Rasmussen K;Reichenheim M;Farias D;Hutcheon J	American Journal of Clinical Nutrition	Gestational weight gain charts: results from the brazilian maternal and child nutrition consortium	2021

543	Giaquinto C;Rampon O;Giacomet V;De R A;Grosch-Worner I;Mok J;Bates I;De J I;Hawkins F;Garcia-Rodriguez M;De G L;Pena J;Garcia J;Lopez J;Asensi-Botet F;Otero M;Pérez-Tamarit D;Orti A;San M M;Scherpbier H;Kreyenbroek M;Boer K;Bohlin A;Belfrage E;Navér L;Levy J;Hainaut M;Peltier A;Goetghebuer T;Barlow P;Ferrazin A;De M A;Gotta C;Mur A;López-Ílichez M;Payà A;Carreras R;Valerius N;Niemeç T;Marczynska M	Pediatric Infectious Disease Journal	Agerelated standards for total lymphocyte cd4supsup and cd8supsup t cell counts in children born in europe	2005
544	Kirshenbaum M;Ziv-Baran T;Katorza E	Journal Of Maternal-Fetal & Neonatal Medicine	Amniotic fluid index measurements in the second and third trimester and correlation to fetal biometric parameters new reference based on a big retrospective data	2022
545	Almughlaq S;Al-Laith A;Al-Thawadi S	Revista Paulista De Pediatria	Comparison between the world health organization and bahraini children growth standards	2023
546	Hao Y;Yang R;Li J;Zhou Z;Qian W;Zhang J;Wu Z;Jin L;Wu X;Zhang C;Zheng B;Tan J;Hao G;Li S;Li Q;Zheng D;Wang Y;Li R;Liu P;Qiao J	American Journal Of Obstetrics And Gynecology	Agespecific random day serum antimullerian hormone reference values for women of reproductive age in the general population a large chinese nationwide populationbased survey	2022
547	Fairclough S;Rowlands A;Cruz B;Crotti M;Foweather L;Graves L;Hurter L;Jones O;Macdonald M;Mccann D;Miller C;Noonan R;Owen M;Rudd J;Taylor S;Tyler R;Boddy L	International Journal Of Behavioral Nutrition And Physical Activity	Reference values for wristworn accelerometer physical activity metrics in england children and adolescents	2023
548	Flatley C;Kumar S;Greer R	Journal Of Maternal-Fetal & Neonatal Medicine	Reference centiles for the middle cerebral artery and umbilical artery pulsatility index and cerebroplacental ratio from a lowrisk population a generalised additive model for location shape and scale gamlss approach	2019
549	Gilchrist R;Kamara A;Rudge J	Revstat-Statistical Journal	An insurance type model for the health cost of cold housing an application of gamlss	2009
550	Rigby R A;Stasinopoulos D M	Statistical Modelling	Using the boxcox iti distribution in gamlss to model skewness and kurtosis	2006
551	Nogueira D;Safadi T;De L R;Da M A;Graciano M;Barcante J;Barcante T;Dourado S	Cadernos De Saude Publica	Spatiotemporal gamlss modeling of the incidence of schistosomiasis in the central region of the state of minas gerais brazil	2023
552	Pereira E;Janeiro V;Guedes T;Cameiro J	Acta Scientiarum-Agronomy	Istevia rebaudianai bert berton regression models with mixed effects for investigating seed germination data	2019
553	Xu Y;Lu F;Zhou Y;Ruan B;Dai Y;Wang K	Frontiers In Environmental Science	Drynesswetness encounter probabilities analysis for lake ecological water replenishment considering nonstationarity effects	2022

554	Delrieu G;Bonnifait L;Kirstetter P;Boudevillain B	Hydrological Sciences Journal-Journal Des Sciences Hydrologiques	Dependence of radar quantitative precipitation estimation error on the rain intensity in the cevennes region france	2014
555	Postuma F;Gasalla M	Fisheries Research	Reproductive activity of the tropical arrow squid <i>idoryteuthis pleii</i> around sao sebastiao island se brazil based on a 10year fisheries monitoring	2014
556	Fernández I;Enriquez-De-Salamanca A;Portero A;García-Vázquez C;Calonge M;Herreras J	Applied Sciences-Basel	Age and sexadjusted reference intervals in tear cytokine levels in healthy subjects	2021
557	Albrecht B;Stalling I;Bammann K	Bmc Geriatrics	Sex and agespecific normative values for handgrip strength and components of the senior fitness test in communitydwelling older adults aged 6575years in germany results from the outdoor active study	2021
558	Sinnige A;Kittelson A;Van D W P;Tejjink J;Hoogeboom T	European Journal Of Vascular And Endovascular Surgery	Personalised outcomes forecasts of supervised exercise therapy in intermittent claudication an application of neighbours based prediction methods with routinely collected clinical data	2022
559	Rodríguez-González A;May-Tec A;Herrera-Silveira J;Puch-Hau C;Quintanilla-Mena M;Villafuerte J;Velazquez-Abunader I;Aguirre-Macedo M;Vidal-Martánez V	Ecological Indicators	Fluctuating asymmetry of sclerotized structures of <i>ihaliotrematoidesi</i> spp monogenea dactylogyridae as bioindicators of aquatic contamination	2020
560	Mouksassi M;Marier J;Cyran J;Vinks A	Clinical Pharmacology & Therapeutics	Clinical trial simulations in pediatric patients using realistic covariates application to teduglutide a glucagonlike peptide2 analog in neonates and infants with shortbowel syndrome	2009
561	Lanvin P;Goronflot T;Isidor B;Nizon M;Durand B;El C S;Genevieve D;Ruault V;Fradin M;Pasquier L;Thevenon J;Delobel B;Burglen L;Afenjar A;Faivre L;Francannet C;Guerrot A;Goldenberg A;Mercier S;Heron D;Lehalle D;Mignot C;Marey I;Charles P;Moutton S;Bezieau S;Bayat A;Piton A;Willems M;Vincent M	American Journal Of Medical Genetics Part A	Growth charts in dyrk1a syndrome	2023
562	Prataviera F;Cordeiro G;Ortega E;Suzuki A	Advances In Data Science And Adaptive Analysis	The odd loglogistic geometric normal regression model with applications	2019
563	Zhang L;Zhang T;Liu H;Xing D;Zhao Y;Zhang Y;Li Y	Nutrition	Body composition in healthy singleton term infants using the three dimensional photonic scanning method a multicenter crossectional study	2023
564	Samraoui B;Nedjah R;Boucheker A;Bouzid A;El-Serehy H;Samraoui F	Ecology And Evolution	Blowin in the wind dispersal of glossy ibisi <i>plegadisii falcinellusi</i> in the west mediterranean basin	2023

565	Oyhenart E;Castro L;Garraza M;Cesani M;Torres M;Quintero F;Dahinten S;Alfaro E;Bejarano I;Carrillo R;Dip N;Lomaglio D;Marrodan M;Menecier N;Navazo B;Roman E;Zonta M;Dipierri J	Nutricion Hospitalaria	Comparative study of midupper arm circumference arm muscle area and arm fat area percentiles in argentinean and us children aged 4 14 years	2019
566	Barber L;Abbott C;Nakhate V;Do A;Blissett A;Marini J	Genetics In Medicine	Longitudinal growth curves for children with classical osteogenesis imperfecta types iii and iv caused by structural pathogenic variants in type i collagen	2019
567	Schmoll T;Schurr F;Winkel W;Epplen J;Lubjuhn T	Journal Of Evolutionary Biology	Polyandry in coal tits iparus ater fitness consequences of putting eggs into multiple genetic baskets	2007
568	Santana-Piñeros A;Cruz-Quintana Y;May-Tec A;Mera-Loor G;Aguirre-Macedo M;Suárez-Morales E;González-Solis D	Plos One	The 20152016 el nino increased infection parameters of copepods on eastern tropical pacific dolphinfish populations	2020
569	Tan X;Gan T	Journal Of Climate	Nonstationary analysis of annual maximum streamflow of canada	2015
570	Seaton S;Yadav K;Field D;Khunti K;Manktelow B	Neonatology	Birthweight centile charts for south asian infants born in the uk	2011
571	Umlauf N;Kneib T	Statistical Modelling	A primer on bayesian distributional regression	2018
572	Prasannakumar I;Kodandaramaiah U	Arthropod-Plant Interactions	Adaptive phenotypic plasticity of mandibles with respect to host plants	2023
573	Ferrari S;Fumes G	Asta-Advances In Statistical Analysis	Boxcox symmetric distributions and applications to nutritional data	2017
574	Vogel M;Kirsten T;Kratzsch J;Engel C;Kiess W	Journal Of Pediatric Endocrinology & Metabolism	A combined approach to generate laboratory reference intervals using unbalanced longitudinal data	2017
575	Villarini G;Vecchi G	Journal Of Climate	North atlantic power dissipation index pdi and accumulated cyclone energy ace statistical modeling and sensitivity to sea surface temperature changes	2012
576	Villarini G;Vecchi G;Smith J	Monthly Weather Review	Modeling the dependence of tropical storm counts in the north atlantic basin on climate indices	2010
577	Mayr A;Hofner B	Statistical Modelling	Boosting for statistical modelling a nontechnical introduction	2018
578	Apurv T;Cai X	Water Resources Research	Evaluation of the stationarity assumption for meteorological drought risk estimation at the multidecadal scale in contiguous united states	2019
579	Villarini G;Seo B;Serinaldi F;Krajewski W	Atmospheric Research	Spatial and temporal modeling of radar rainfall uncertainties	2014
580	Villarini G;Vecchi G;Knutson T;Smith J	Journal Of Geophysical Research-Atmospheres	Is the recorded increase in shortduration north atlantic tropical storms spurious	2011
581	Villarini G;Vecchi G;Smith J	Monthly Weather Review	Us landfalling and north atlantic hurricanes statistical modeling of their frequencies and ratios	2012
582	Villarini G;Strong A	Agriculture Ecosystems & Environment	Roles of climate and agricultural practices in discharge changes in an agricultural watershed in iowa	2014
583	Nishii R;Tanaka S	Environmental And Ecological Statistics	Modeling and inference of forest coverage ratio using zeroone inflated distributions with spatial dependence	2013

584	Villarini G;Jones C;Schilling K	Journal Of Environmental Quality	Soybean area and baseflow driving nitrate in iowas raccoon river	2016
585	Villarini G;Smith J;Baeck M;Vitolo R;Stephenson D;Krajewski W	Journal Of Hydrology	On the frequency of heavy rainfall for the midwest of the united states	2011
586	Leroy B;Peatman T;Usu T;Caillot S;Moore B;Williams A;Nicol S	Marine Policy	Interactions between artisanal and industrial tuna fisheries insights from a decade of tagging experiments	2016
587	Villarini G;Schilling K;Jones C	Journal Of Environmental Management	Assessing the relation of usda conservation expenditures to suspended sediment reductions in an iowa watershed	2016
588	Villarini G;Smith J;Napolitano F;Baeck M	Hydrological Sciences Journal	Hydrometeorological analysis of the december 2008 flood in rome	2011
589	Kirstetter P;Gourley J;Hong Y;Zhang J;Moazamigoodarzi S;Langston C;Arthur A	Water Resources Research	Probabilistic precipitation rate estimates with groundbased radar networks	2015
590	Villarini G;Vecchi G	Journal Of Climate	Projected increases in north atlantic tropical cyclone intensity from cmip5 models	2013
591	Yumnam T;Banerjee B;Kodandaramaiah U	Biological Journal Of The Linnean Society	Pupal colour plasticity in the butterfly icatopsilia pomonai lepidoptera pieridae	2021
592	Villarini G;Scoccimarro E;White K;Arnold J;Schilling K;Ghosh J	Journal Of The American Water Resources Association	Projected changes in discharge in an agricultural watershed in iowa	2015
593	Rigby R A; Stasinopoulos D M	Statistics In Medicine	Smooth centile curves for skew and kurtotic data modelled using the boxcox power exponential distribution	2004

Supplementary Material 2
Trend Topics

Trend topic: *Reference curves*

1. Addo, O. Y., Himes, J. H., and Zemel, B. S. (2017). Reference ranges for midupper arm circumference, upper arm muscle area, and upper arm fat area in US children and adolescents aged 1-20 y. *The American Journal of Clinical Nutrition*, 105(1), 111–120.
2. Balder, J. W., Lansberg, P. J., Hof, M. H., Wiegman, A., Hutten, B. A., and Kuivenhoven, J. A. (2018). Pediatric lipid reference values in the general population: The Dutch lifelines cohort study. *Journal of Clinical Lipidology*, 12(5), 1208–1216.
3. Blagus, R., Jurak, G., Starc, G., and Leskošek, B. (2023). Centile reference curves of the SLOfit physical fitness tests for school-aged children and adolescents. *The Journal of Strength & Conditioning Research*, 37(2), 328–336.
4. Bongers, B. C., de Vries, S. I., Obeid, J., van Buuren, S., Helders, P. J. M., and Takken, T. (2013). The Steep Ramp Test in Dutch white children and adolescents: age- and sex-related normative values. *Physical Therapy*, 93(11), 1530–1539.
5. Bozek, J., Griffanti, L., Lau, S., and Jenkinson, M. (2023). Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria. *Neuroimage*, 268, 119864.
6. Burt, L. A., Liang, Z., Sajobi, T. T., Hanley, D. A., and Boyd, S. K. (2016). Sex- and site-specific normative data curves for HR-pQCT. *Journal of Bone and Mineral Research*, 31(11), 2041–2047.
7. Cromwell, E. A., Dube, Q., Cole, S. R., Chirambo, C., Dow, A. E., Heyderman, R. S., and Van Rie, A. (2014). Validity of US norms for the Bayley Scales of Infant Development-III in Malawian children. *European Journal of Paediatric Neurology*, 18(2), 223–230.
8. De Henauw, S., Michels, N., Vyncke, K., Hebestreit, A., Russo, P., Intemann, T., Peplies, J., Fraterman, A., Eiben, G., de Lorgeril, M., Tornaritis, M., Molnar, D., Veidebaum, T., Ahrens, W., and Moreno, L. A. (2014). Blood lipids among young children in Europe: results from the European IDEFICS study. *International Journal of Obesity*, 38, S67–S75.
9. Douma-van Riet, D., Verschuren, O., Jelsma, D., Kruitwagen, C., Smits-Engelsman, B., and Takken, T. (2012). Reference values for the muscle power sprint test in 6- to 12-year-old children. *Pediatric Physical Therapy*, 24(4), 327–332.
10. Ertl, D. A., Gleiss, A., Sagmeister, S., and Haeusler, G. (2014). Determining the normal range for IGF-I, IGFBP-3, and ALS: new reference data based on current internal standards. *Wiener Medizinische Wochenschrift*, 164, 343–352.
11. Flatley, C., Sole-Navais, P., Vaudel, M., Helgeland, Ø., Modzelewska, D., Johansson, S., Jacobsson, B., and Njølstad, P. (2022). Placental weight centiles adjusted for age, parity and fetal sex. *Placenta*, 117, 87–94.
12. Gjessing, H. K., Grøttum, P., Økland, I., and Eik-Nes, S. H. (2017). Fetal size monitoring and birth-weight prediction: a new population-based approach. *Ultrasound in Obstetrics & Gynecology*, 49(4), 500–507.

13. Gleiss, A., Lassi, M., Blümel, P., Borkenstein, M., Kapelari, K., Mayer, M., Schemper, M., and Häusler, G. (2013). Austrian height and body proportion references for children aged 4 to under 19 years. *Annals of Human Biology*, 40(4), 324–332.
14. He, J. R., Xia, H. M., Xia, X. Y., Mo, W. J., Wang, P., Cheng, K. K., Leung, G. M., Feng, Q., Schooling, C. M., and Qiu, X. (2014). A new birthweight reference in Guangzhou, southern China, and its comparison with the global reference. *Archives of Disease in Childhood*, 99(12), 1091–1097.
15. Herrmann, D., Intemann, T., Lauria, F., Mårild, S., Molnár, D., Moreno, L. A., Sioen, I., Tornaritis, M., Veidebaum, T., Pigeot, I., Ahrens, W., and IDEFICS consortium (2014). Reference values of bone stiffness index and c-terminal telopeptide in healthy european children. *International Journal of Obesity*, 38, S76–85.
16. Mattiello, R., Mundstock, E., and Ziegelmann, P. K. (2022). Brazilian reference percentiles for bioimpedance phase angle of healthy individuals. *Frontiers in Nutrition*, 9, 912840.
17. Nielsen, T. R. H., Lausten-Thomsen, U., Fonvig, C. E., Bøjsøe, C., Pedersen, L., Bratholm, P. S., Hansen, T., Pedersen, O., and Holm, J. C. (2017). Dyslipidemia and reference values for fasting plasma lipid concentrations in Danish/North-European white children and adolescents. *BMC Pediatrics*, 17, 116.
18. Oyhenart, E. E., Castro, L. E., Garraza, M., Cesani, M. F., Torres, M. F., Quintero, F. A., Dahinten, S. L., Alfaro, E. L., Bejanaro, I. F., Carrillo, R., Dip, N. B., Lomaglio, D. B., Marrodán, M. D., Menezier, N., Navazo, B., Román, E. M., Zonta, M. L., and Dipierri, J. E. (2019). Comparative study of mid-upper arm circumference, arm muscle area and arm fat area percentiles in Argentinean and US children aged 4-14 years. *Nutrición Hospitalaria*, 36(3), 552–562.
19. Rand, M. S., Diemar, S. S., Møllehave, L. T., Heidemann, M., Thuesen, B. H., Petersen, J. H., Johannesen, J., Schou, A. J., Wedderkopp, N., Mølgaard, C., and Jørgensen, N. R. (2023). Z-scores of bone turnover markers calculated from new established sex- and age-specific reference curves are associated to future change in BMD in children and adolescents. *Bone*, 167, 116641.
20. Rashidi, A. A., Kiani, O., Heidarzadeh, M., Imani, B., Nematy, M., Taghipour, A., Sadr, M., and Norouzy, A. (2018). Reference curves of birth weight, length, and head circumference for gestational age in Iranian singleton births. *Iranian Journal of Pediatrics*, 28(5), e66291.
21. Raut, D. S., Raje, D. V., Dandge, V. P., and Singh, D. (2018). Percentile reference curves for normal pancreatic dimensions in Indian children. *Indian Journal of Radiology and Imaging*, 28(4), 442–447.
22. Rigby, R. A. and Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the box-cox power exponential distribution. *Statistics in Medicine*, 23(19), 3053–3076.
23. Rigby, R. A. and Stasinopoulos, D. M. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6(3), 209–229.
24. Smith, E., Wilkinson, K., Wyatt, S., Vaish, P. K., and Miller, S. (2019b). Reference curves for a fitness battery developed for children ages 5-12 years in England. *Journal of Physical Education and Sport*, 19(4), 1481–1495.

25. Tinggaard, J., Aksglaede, L., Sørensen, K., Mouritsen, A., Wohlfahrt-Veje, C., Hagen, C. P., Mieritz, M. G., Jørgensen, N., Wolthers, O. D., Heuck, C., Petersen, J. H., Main, K. M., and Juul, A. (2014). The 2014 Danish references from birth to 20 years for height, weight and body mass index. *Acta Paediatrica*, 103(2), 214–224.
26. Truckenbrod, C., Meigen, C., Brandt, M., Vogel, M., Wahl, S., Jurkutat, A., and Kiess, W. (2020). Reference curves for refraction in a German cohort of healthy children and adolescents. *Plos One*, 15(3), e0230291.
27. Van Eyck, A., Eerens, S., Trouet, D., Lauwers, E., Wouters, K., De Winter, B. Y., van der Lee, J., Van Hoeck, K., and Ledeganck, K. J. (2021). Body composition monitoring in children and adolescents: reproducibility and reference values. *European Journal of Pediatrics*, 180(6), 1721–1732.
28. Vendula, J., Miroslava, P., and Tereza, K. (2021). Reference curves of selected circumferential parameters for Czech children aged 6 to 11 years. *Children*, 8(10), 908.
29. Verschuren, O., Bloemen, M., Kruitwagen, C., and Takken, T. (2010a). Reference values for aerobic fitness in children, adolescents, and young adults who have cerebral palsy and are ambulatory. *Physical Therapy*, 90(8), 1148–1156.
30. Verschuren, O., Bloemen, M., Kruitwagen, C., and Takken, T. (2010b). Reference values for anaerobic performance and agility in ambulatory children and adolescents with cerebral palsy. *Developmental Medicine & Child Neurology*, 52(10), e222–8.
31. Whittier, D. E., Burt, L. A., Hanley, D. A., and Boyd, S. K. (2020). Sex- and site-specific reference data for bone microarchitecture in adults measured using second-generation HR-pQCT. *Journal of Bone and Mineral Research*, 35(11), 2151–2158.
32. Winkler, C., Linden, K., Mayr, A., Schultz, T., Welchowski, T., Breuer, J., and Herberg, U. (2020). Refcurv: A software for the construction of pediatric reference curves. *Software Impacts*, 6, 100040.
33. Zbořilová, V., Přidalová, P., and Kaplanová, T. (2021). Body fat mass, percent body fat, fat-free mass, and skeletal muscle mass reference curves for Czech children aged 6–11 years. *Children*, 2021(8), 366.

Trend topic: *Overdispersion*

1. Coupé, C. (2018). Modeling linguistic variables with regression models: Addressing non-gaussian distributions, non-independent observations, and non-linear predictors with random effects and generalized additive models for location, scale, and shape. *Frontiers in Psychology*, 9, 513.
2. Freitas, L. A. C., Taconeli, C. A., Silva, J. L. P., Tamioso, P. R., and Molento, C. F. M. (2021). A case study on animal behavior analysis using GAMLSS. *Brazilian Journal of Biometrics*, 39(3), 412–433.
3. Gonçalves, J. N. and Barreto-Souza, W. (2020). Flexible regression models for counts with high-inflation of zeros. *Metron*, 78, 71–95.

4. Klein, N., Kneib, T., and Lang, S. (2015). Bayesian generalized additive models for location, scale, and shape for zero-inflated and overdispersed count data. *Journal of the American Statistical Association*, 110(509), 405–419.
5. Marletta, A. and Sciandra, M. (2020). GAMLSS for high-variability data: an application to liver fibrosis case. *The International Journal of Biostatistics*, 16(2), 20190113.
6. Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3), 507–554.
7. Smith, A., Hofner, B., Lamb, J. S., Osenkowski, J., Allison, T., Sadoti, G., McWilliams, S. R., and Paton, P. (2019a). Modeling spatiotemporal abundance of mobile wildlife in highly variable environments using boosted GAMLSS hurdle models. *Ecology and Evolution*, 9(5), 2346–2364.
8. Stasinopoulos, D. M. and Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, 23(7), 1–46.
9. Thomas, J., Mayr, A., Bischl, B., Schmid, M., Smith, A., and Hofner, B. (2018). Gradient boosting for distributional regression: faster tuning and improved variable selection via noncyclical updates. *Statistics and Computing*, 28, 673–687.
10. Ziel, F. (2022). M5 competition uncertainty: Overdispersion, distributional forecasting, GAMLSS and beyond. *International Journal of Forecasting*, 38(4), 1546–1554.

Trend topic: *Machine learning*

1. Bozek, J., Griffanti, L., Lau, S., and Jenkinson, M. (2023). Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria. *Neuroimage*, 268, 119864.
2. Marmolejo-Ramos, F., Tejo, M., Brabec, M., Kuzilek, J., Joksimovic, S., Kovanovic, V., González, J., Kneib, T., Bühlmann, P., Kook, L., Briseño Sánchez, G., and Ospina, R. (2023). Distributional regression modeling via generalized additive models for location, scale, and shape: An overview through a data set from learning analytics. *WIREs Data Mining and Knowledge Discovery*, 13(1), e1479.
3. Mayr, A. and Hofner, B. (2018). Boosting for statistical modelling - a non-technical introduction. *Statistical Modelling*, 18(3–4), 365–384.
4. Murray, J., Ayers, J., and Brookfield, A. (2023). The impact of climate change on monthly baseflow trends across Canada. *Journal of Hydrology*, 618, 129254.
5. Restrepo-Mesa, S. L., Rincón, M. V. B., Restrepo, A. E., Carrilho, T. R. B., Kac, G., Pulgarín, J. S. C., Cano-Pulgarín, K., Severi, C., Sinisterra, O., Sarmiento, M. C. Z., Ocampos, M. I. L., Bannout, M. A., Chico-Barba, G., Arteaga, N. P., Grandi, C., Samur, E. A., and Escobar, C. D. S. (2023). Gestational weight gain charts for Latin American adolescents. *Plos One*, 18(11), e0292070.
6. Rojas, F., Leiva, V., Huerta, M., and Martin-Barreiro, C. (2021). Lot-size models with uncertain demand considering its skewness/kurtosis and stochastic programming applied to hospital pharmacy with sensor-related COVID-19 data. *Sensors*, 21(15), 5198.

7. Sonnenschein, B. and Ziel, F. (2023). Probabilistic intraday wastewater treatment plant inflow forecast utilizing rain forecast data and sewer network sensor data. *Water Resources Research*, 59(11), e2022WR033826.
8. Ulrich, M., Jahnke, H., Langrock, R., Pesch, R., and Senge, R. (2021). Distributional regression for demand forecasting in e-grocery. *European Journal of Operational Research*, 294(3), 831–842.
9. Yagli, G. M., Yang, D., and Srinivasan, D. (2020). Ensemble solar forecasting using data-driven models with probabilistic post-processing through GAMLSS. *Solar Energy*, 208, 612–622.

Trend topic: Other Medical Science terms

1. Aasim, M. and Chand, S. (2023). Height-for-age and weight-for-age growth charts for Pakistani infants under six months: derived from a novel case selection method using multiple indicator cluster survey data. *BMC Medical Research Methodology*, 23, 289.
2. Almughlaq, S., Al-Laith, A. A., and Al-Thawadi, S. (2023). Comparison between the World Health Organization and Bahraini children growth standards. *Revista Paulista de Pediatria*, 41, e2022050.
3. Arnold, N., Ascherl, R. G., and Thome, U. H. (2022). Charts and LMS tables of transfontanelar and transversal ear-to-ear distances for gestational age. *Frontiers in Pediatrics*, 10, 838333.
4. Baer, S., Tuzin, N., Kang, P. B., Mohammed, S., Kubota, M., van Ierland, Y., Busa, T., Rossi, M., Morel, G., Michot, C., Baujat, G., Durand, M., Obringer, C., May, N. L., Calmels, N., and Laugel, V. (2021). Growth charts in Cockayne syndrome type 1 and type 2. *European Journal of Medical Genetics*, 64(1), 104105.
5. Barber, L. A., Abbott, C., Nakhate, V., Do, A. N. D., Blissett, A. R., and Marini, J. C. (2019). Longitudinal growth curves for children with classical osteogenesis imperfecta (types iii and iv) caused by structural pathogenic variants in type I collagen. *Genetics in Medicine*, 21(5), 1233–1239.
6. Bergerat, M., Heude, B., Taine, M., Tich, S. N. T., Werner, A., Frandji, B., Blauwblomme, T., Sumanaru, D., Charles, M. A., Chalumeau, M., and Scherdel, P. (2021). Head circumference from birth to five years in France: New national reference charts and comparison to WHO standards. *The Lancet Regional Health - Europe*, 5, 100114.
7. Bertapelli, F., Agiovlaitis, S., Machado, M. R., Roso, R. V., and Guerra-Junior, G. (2017). Growth charts for Brazilian children with Down syndrome: Birth to 20 years of age. *Journal of Epidemiology*, 27(6), 265–273.
8. Brooks, J., Day, S., Shavelle, R., and Strauss, D. (2011). Low weight, morbidity, and mortality in children with cerebral palsy: new clinical growth charts. *Pediatrics*, 128(2), e299–307.
9. Cheung, M. S., Cole, T. J., Araundel, P., Bridges, N., Burren, C. P., Cole, T., Davies, J. H., Hagenäs, L., Högl, W., Hulse, A., Mason, A., McDonnell, C., Merker, A., Mohnike, K., Sabir, A., Skae, M., Rothenbuhler, A., Warner, J., and Irving, M. (2024). Growth reference charts for children with hypochondroplasia. *American Journal of Medical Genetics – Part A*, 194(2), 243–252.

10. Chhavi, N., Ojha, S., Awasthi, A., Shalimar, and Goel, A. (2022). Serum level of alanine- and aspartate-aminotransferase levels in newborns in India. *Journal of Clinical and Experimental Hepatology*, 12(2), 306–311.
11. Cole, T. J. (2021). Sample size and sample composition for constructing growth reference centiles. *Statistical Methods in Medical Research*, 30(2), 488–507.
12. Cole, T. J., Stanojevic, S., Stocks, J., Coates, A. L., Hankinson, J. L., and Wade, A. M. (2009). Age- and size-related reference ranges: A case study of spirometry through childhood and adulthood. *Statistics in Medicine*, 28(5), 880–898.
13. Dangat, K., Gupte, S., Wagh, G., Lalwani, S., Randhir, K., Madiwale, S., Pisal, H., Kadam, V., Gundu, S., Chandhiok, N., Kulkarni, B., Joshi, S., Fall, C., and Sachdev, H. S. (2022). Gestational weight gain in the REVAMP pregnancy cohort in Western India: Comparison with international and national references. *Frontiers in Medicine*, 9, 1022990.
14. Durán, P., Merker, A., Briceño, G., Colón, E., Line, D., Abad, V., Del Toro, K., Chahín, S., Matallana, A. M., Lema, A., Llano, M., Céspedes, J., and Haenäs, L. (2016). Colombian reference growth curves for height, weight, body mass index and head circumference. *Acta Paediatrica*, 105(3), e116–125.
15. Funk, S. M., Guerra, B. P., Zamora, A. B., Ickowitz, A., Poni, N. A., Abdou, M. A., Sibama, Y. H., Penda, R., Brull, G. R., Abossolo, M., Martín, E. A., Okale, R., Ze, B. A., Carrión, A. M., Sebastián, C. G., García, C. R. L., Salazar, F. L. R., Amazia, H., Reyes, I. A., Expósito, R. S., and Fa, J. E. (2020). Understanding growth and malnutrition in Baka Pygmy children. *Human Ecology*, 48, 293–306.
16. Genowska, A., Strukcinskiene, B., Bochenko-Luczyńska, J., Motkowski, R., Jamiolkowski, J., Abramowicz, P., and Konstantynowicz, J. (2023). Reference values for birth weight in relation to gestational age in Poland and comparison with the global percentile standards. *Clinical medicine*, 12(17), 5736.
17. Georgiadis, L., Keski-Nisula, L., Harju, M., Räisänen, S., Georgiadis, S., Hannila, M. L., and Heinon, S. (2014). Umbilical cord length in singleton gestations: a Finnish population-based retrospective register study. *Placenta*, 35(4), 275–280.
18. Hans, N., Klein, N., Faschingbauer, F., Schneider, M., and Mayr, A. (2023). Boosting distributional copula regression. *Biometrics*, 79(3), 2298–2310.
19. Heude, B., Scherdel, P., Werner, A., Guern, M. L., Gelbert, N., Walther, D., Arnould, M., Bellaïche, M., Chevallier, B., Cheymol, J., Jobez, E., N’Guyen, S., Pietrement, C., Reynaud, R., Salaün, J. F., Khoshnood, B., Zeitlin, J., Maccario, J., Breart, G., Thalabard, J. C., Charles, M. A., Botton, J., Frandji, B., and Chalumeau, M. (2019). A big-data approach to producing descriptive anthropometric references: a feasibility and validation study of paediatric growth charts. *The Lancet Digital Health*, 1(8), e413–e423.
20. Ivanoviski, I., Djuric, O., Broccoli, S., Caraffi, S. G., Accorsi, P., Adam, M. P., Avela, K., Badura-Stronka, M., Bayat, A., Clayton-Smith, J., Cocco, I., Cordelli, D. M., Cuturilo, G., Di Pisa, V., Garcia, J. D., Gastaldi, R., Giordano, L., Guala, A., Hoei-Hansen, C., Inaba, M., Iodice, A., Nielsen, J E K, K. V., Lazalde-Medina, B., Malbora, B., Mizuno, S., Moldovan, O., Møller, R. S., Muschke, P., Otelli, V., Pantaleoni, C., Piscopo, C., Poch-Olive, M. L., Prpic, I., Reina, P. M., Raviglione, F., Ricci, E., Scarano, E., Simonte, G., Smigiel, R., Tanteles, G., Tarani, L., Trimouille,

- A., Valera, E. T., Vergano, S. S., Writzl, K., Callewaert, B., Savasta, S., Street, M. E., Iughetti, L., Bernasconi, S., Rossi, P. G., and Garavelli, L. (2020). Mowat-Wilson syndrome: growth charts. *Orphanet Journal of Rare Diseases*, 15, 151.
21. Karvonen, M., Hannila, M. L., Saari, A., and Dunkel, L. (2012). New Finnish reference for head circumference from birth to 7 years. *Annals of Medicine*, 44(4), 369–374.
22. Kocevar, G., Rioland, M., Laxalde, J., Mugnier, A., Adib-Lesaux, A., Gaillard, V., and Bodin, J. (2023). Growth charts for small sample sizes using unsupervised clustering: Application to canine early growth. *Veterinary Research Communications*, 47, 693–706.
23. Lanvin, P. L., Goronflot, T., Isidor, B., Nizon, M., Durand, B., El Chehadeh, S., Geneviève, D., Ruault, V., Fradin, M., Pasquier, L., Thévenon, J., Delobel, B., Burglen, L., Afenjar, A., Faivre, L., Francannet, C., Guerrot, A. M., Goldenberg, A., Mercier, S., Héron, D., Lehalle, D., Mignot, C., Marey, I., Charles, P., Moutton, S., Bézieau, S., Bayat, A., Piton, A., Willems, M., and Vincent, M. (2024). Growth charts in DYRK1A syndrome. *American Journal of Medical Genetics - Part A*, 194(1), 9–16.
24. Lauffer, P., Pals, G., Zwinderman, A. H., Postema, F. A. M., Baars, M. J. H., Dulfer, E., Hilhorst-Hofstee, Y., Houweling, A. C., Kempers, M., Krapels, I. P. C., van de Laar, I. M. B. H., Loeys, B., Spaans, A. M. J., Warnink-Kavelaars, J., de Waard, V., Wit, J. M., and Menke, L. A. (2023). Growth charts for Marfan syndrome in the Netherlands and analysis of genotype-phenotype relationships. *American Journal of Medical Genetics – Part A*, 191(2), 479–489.
25. Levin, D., Marryat, L., Cole, T. J., McColl, J., Harjunmaa, U., Ashorn, P., and Wright, C. (2016). Fit to WHO weight standard of European infants over time. *Archives of Disease in Childhood*, 101(5), 455–460.
26. Merker, A., Neumeyer, L., Hertel, N. T., Grigelioniene, G., Mohnike, K., and Hagenäs, L. (2018). Development of body proportions in achondroplasia: Sitting height, leg length, arm span, and foot length. *American Journal of Medical Genetics – Part A*, 176(9), 1819–1829.
27. Morkuniene, R., Tutkuvienė, J., Cole, T. J., Jakimaviciene, E. M., Isakova, J., Bankauskiene, A., Drazdiene, N., and Basys, V. (2022). Neonatal head circumference by gestation reflects adaptation to maternal body size: comparison of different standards. *Scientific Reports*, 12, 11057.
28. Morkuniene, R., Cole, T. J., Jakimaviciene, E. M., Bankauskiene, A., Isakova, J., Drazdiene, N., Basys, V., and Tutkuvienė, J. (2023). Regional references vs. international standards for assessing weight and length by gestational age in Lithuanian neonates. *Frontiers in Pediatrics*, 14(11), 1173685.
29. Oliveira, M. H., Araújo, J., Ramos, E., and Conde, W. L. (2023). MULT: New height references and their efficiency in multi-ethnic populations. *American Journal of Human Biology*, 35(5), e23859.
30. Öztürk, A., Çiçek, B., Mazicioğlu, M. M., Zararsiz, G., and Kurtoğlu, S. (2017). Wrist circumference and frame size percentiles in 6-17-year-old Turkish children and adolescents in Kayseri. *Journal of Clinical Research in Pediatric Endocrinology*, 15(9), 329–336.
31. Premkumar, P., Antonisamy, B., Mathews, J., Benjamin, S., Regi, A., Jose, R., Kuruvilla, A., and Mathai, M. (2016). Birth weight centiles by gestational age for twins born in south India. *BMC Pregnancy and Childbirth*, 16, 64.

32. Rashidi, A. A., Kiani, O., Heidarzadeh, M., Imani, B., Nematy, M., Taghipour, A., Sadr, M., and Norouzy, A. (2018). Reference curves of birth weight, length, and head circumference for gestational age in Iranian singleton births. *Iranian Journal of Pediatrics*, 28(5), e66291.
33. Restrepo-Mesa, S. L., Rincón, M. V. B., Restrepo, A. E., Carrilho, T. R. B., Kac, G., Pulgarín, J. S. C., Cano-Pulgarín, K., Severi, C., Sinisterra, O., Sarmiento, M. C. Z., Ocampos, M. I. L., Bannout, M. A., Chico-Barba, G., Arteaga, N. P., Grandi, C., Samur, E. A., and Escobar, C. D. S. (2023). Gestational weight gain charts for Latin American adolescents. *Plos One*, 18(11), e0292070.
34. Saari, A., Sankilampi, U., Hannila, M. L., Kiniviemi, V., Kesseli, K., and Dunkel, L. (2011). New Finnish growth references for children and adolescents aged 0 to 20 years: Length/height-for-age, weight-for-length/height, and body mass index-for-age. *Annals of Medicine*, 43(3), 235–248.
35. Truckenbrod, C., Meigen, C., Brandt, M., Vogel, M., Wahl, S., Jurkutat, A., and Kiess, W. (2020). Reference curves for refraction in a German cohort of healthy children and adolescents. *Plos One*, 15(3), e0230291.
36. Valencia-Sosa, E., Chávez-Palencia, C., Vallarta-Robledo, J. R., Romedo-Velarde, E., Larrosa-Haro, A., Vásquez-Garibay, E. M., and Ramos-García, C. O. (2021). Percentile reference values for the neck circumference of mexican children. *Children*, 8(5), 407.
37. van Beijsterveldt, I. A. L. P., van der Steen, M., de Fluiter, K. S., Spaans, S. A. M. J., and Hokken-Koelega, A. C. S. (2022). Body composition and bone mineral density by dual energy x-ray absorptiometry: Reference values for young children. *Clinical Nutrition*, 41(1), 71–79.
38. Verbeek, S., Eilers, P. H. C., Lawrence, K., Hennekam, R. C. M., and Versteegh, F. G. A. (2011). Growth charts for children with Ellis–van Creveld syndrome. *European Journal of Pediatrics*, 170, 207–211.
39. Voigt, M., Meyer-Kahrweg, L. M., Landau-Crangle, E., So, H. Y., Däbritz, J., Rochow, M., Kunze, M., and Rochow, N. (2020a). Individualized birth length and head circumference percentile charts based on maternal body weight and height. *Journal of Perinatal Medicine*, 48(7), 656–664.
40. Voigt, M., Rochow, N., Landau-Crangle, E., Meyer-Kahrweg, L. M., Olbertz, D. M., Kunze, M., Nikischin, W., Wittwer-Backofen, U., Rochow, M., Däbritz, J., and Hentschel, R. (2020b). Individualized sex-specific birth weight percentiles for gestational age based on maternal height and weight. *Journal of Perinatal Medicine*, 49(1), 94–103.
41. Wu, Q., Zhang, H. Y., Zhang, L., Xu, Y. Q., Sun, J., Gao, N. N., Qiao, X. Y., and Li, Y. (2022). A new birthweight reference by gestational age: A population study based on the generalized additive model for location, scale, and shape method. *Frontiers in Pediatrics*, 10, 810203.
42. Zhang, L., Lin, J. G., Liang, S., Sun, J., Gao, N. N., Wu, Q., Zhang, H. Y., Liu, H. J., Cheng, X. D., Cao, Y., and Li, Y. (2021). Comparison of postnatal growth charts of singleton preterm and term infants using World Health Organization standards at 40–160 weeks postmenstrual age: A Chinese single-center retrospective cohort study. *Frontiers in Pediatrics*, 9, 595882.
43. Zhang, T., Su, X., Wu, L., and Chu, J. (2023). Identification of dynamic drought propagation from a nonstationary perspective and its application to drought warnings. *Journal of Hydrology*, 626, 130372.

Supplementary Material 3

Articles explicitly addressing scientific research gaps

The major goal of Almughlaq et al. (2023) was to investigate the growth patterns of male and female infants in Bahrain aged 0 to 24 months. Their findings revealed considerable discrepancies when compared with the WHO Multicentre Growth Reference Study (WHO-MGRS) reference charts, wherein Bahraini children aged 0 to 24 months, of both sexes, exhibited greater height and weight than their counterparts. Authors emphasise that future longitudinal studies are needed to monitor children's growth patterns utilising body composition methods, adiposity markers, and growth determinants to investigate this deviation from the WHO-MGRS.

Reyes et al. (2022) assessed the impact of the initial mass vaccination programme using a two-dose schedule on reducing adverse epidemiological outcomes due to SARS-CoV-2 infection. Authors emphasise the importance of completing the full vaccine regimen to mitigate the impacts of COVID-19 and suggest that future research should use epidemiological data to analyse the impact of vaccines, including booster doses, as the pandemic advances and new variants emerge.

The study by Sinnige et al. (2022) aimed to develop and evaluate personalised outcomes forecasts for functional claudication distance over six months of supervised exercise therapy for individuals with intermittent claudication using neighbours-based prediction methods. The authors propose that future study investigate the impact of personalised outcome forecasts on clinical decision-making and patient outcomes.

Albrecht et al. (2021) aimed to calculate specific normative values by sex and age for handgrip strength of the Senior Fitness Test (SFT) in German individuals aged 65 to 75. According to the research findings, both men and women's normative values decrease with age. Further investigation on the relationship between normative physical fitness values, health outcomes, and quality of life in adult is required. Authors attest that future studies should provide gender and age-specific normative values for handgrip strength and components of the SFT in older adults in the country.

Voncken et al. (2021) investigated how the statistical model's flexibility impacts bias, variance and total variability in estimates of z-scores under empirically relevant conditions. Based on the GAMLSS framework, the authors found that models with overly strict distributional assumptions produced biased results, whereas models with excessive flexibility result in higher variance. In their findings, the skew Student t distribution appeared to be more difficult to estimate than the Box-Cox power exponential distribution for normally distributed data. Authors conclude their article with recommendations for future empirical normalisation studies.

Zouiouich et al. (2021) explored the relationships between serological measures of metabolic health and the gut micro biome in two population cohorts, trying to find potential associations and microbiological markers of these metabolic variables. The authors point out that among the existing gaps, there is a need for a better understanding of the underlying mechanisms behind the observed associations, as well as the replication of these results in other populations. They further claim that the study is mostly based on observational relationships, and it is necessary a longitudinal and

experimental research to establish causality and evaluate the clinical relevance of these associations.

Funk et al. (2020) focused on the study of the prevalence of growth issues, such as short stature, low weight, and obesity, among Baka Pygmy children ages 2 to 12 in the Djoum-Mintom region of southwestern Cameroon. The authors argue that WHO child growth Standards are not appropriate for Pygmies, and to help achieve the United Nations Sustainable Development goals, future studies should focus on defining unique growth standards for the various Pygmy tribes, taking into consideration their genetic differences.

Tadphale et al. (2020) aimed to construct nomograms for the diameters of femoral arteries and femoral veins in children aged 0 to 4 years, as well as to create probability curves for the incidence of difficulties in arterial access based on the femoral artery diameter. The study involved ultrasonographic examination of both characteristics in 400 children over a three-year period. The authors emphasise the importance of future studies with larger samples, including children of various ages, to fill gaps and increase the applicability of the nomograms and probability curves.

Truckenbrod et al. (2020) created centile curves of refractive development for German children and compared their findings with data from children in Central Asia. The study gives the first reference curve for refraction in Central Europe, based on non-cycloplegic refraction measured in children. In comparison to data from China and Korea, only a small difference is detected at 5 years of age in all percentiles, which then increases continually, in addition to the progression of myopia being much higher in China and Korea than in Germany. The most noticeable differences can be observed in the lower centiles. Authors state that future research should determine whether the onset of preschool activities with prolonged near-work causes divergence in refractive development.

Chang et al. (2019) developed spirometric reference equations for healthy Taiwanese children and assessed the applicability of the Global Lung Function Initiative (GLI)-2012 equations to the data. According to the authors, future research should look into characteristics other than those examined in their study, such as ethnicity and environmental exposures, to provide a more comprehensive understanding of spirometry in healthy Taiwanese children. Since the study provides evidence that the GLI-2012 reference equations do not adequately correspond to their spirometric data, authors identify an urgent need for an update of GLI reference values, including more data from non-Caucasian descendants as a research gap.

Dunn et al. (2019) established reference ranges for maternal levels of placental growth factor (PLGF) in a cohort of normotensive pregnant women after 36 weeks of gestation, suggesting that these levels decrease throughout this period. The authors argue that there is a gap in understanding the relationship between PLGF levels and perinatal outcomes in normotensive pregnant women after 36 weeks of gestation, and that the utility of PLGF as a predictor of adverse term perinatal outcomes should be investigated further through clinical trials.

References

1. Albrecht, B. M., Stalling, I., and Bammann, K. (2021). Sex- and age-specific normative values for handgrip strength and components of the Senior Fitness Test in community-dwelling older adults aged 65–75 years in Germany: results from the OUTDOOR ACTIVE study. *BMC Geriatrics*, 21, 273.
2. Almughlaq, S., Al-Laith, A. A., and Al-Thawadi, S. (2023). Comparison between the World Health Organization and Bahraini children growth standards. *Revista Paulista de Pediatria*, 41, e2022050.
3. Chang, S. M., Tsai, H. J., Tzeng, J. Y., Yeh, K. W., Chen, L. C., Lai, S. H., Liao, S. L., Hua, M. C., Tsai, M. H., Huang, J. L., and Yao, T. C. (2019). Reference equations for spirometry in healthy Asian children aged 5 to 18 years in Taiwan. *World Allergy Organization Journal*, 12(11), 100074.
4. Dunn, L., Sherrell, H., Bligh, L., Alsolai, A., Flatley, C., and S, K. (2019). Reference centiles for maternal placental growth factor levels at term from a low-risk population. *Placenta*, 86, 15–19.
5. Funk, S. M., Guerra, B. P., Zamora, A. B., Ickowitz, A., Poni, N. A., Abdou, M. A., Sibama, Y. H., Penda, R., Brull, G. R., Abossolo, M., Martín, E. A., Okale, R., Ze, B. A., Carrión, A. M., Sebastián, C. G., García, C. R. L., Salazar, F. L. R., Amazia, H., Reyes, I. A., Expósito, R. S., and Fa, J. E. (2020). Understanding growth and malnutrition in Baka Pygmy children. *Human Ecology*, 48, 293–306.
6. Reyes, H., Diethelm-Varela, B., Méndez, C., Rebolledo-Zelada, D., Lillo-Dapremont, B., Muñoz, S. R., Bueno, S. M., González, P. A., and Kalergis, A. M. (2022). Contribution of two-dose vaccination toward the reduction of COVID-19 cases, ICU hospitalizations and deaths in Chile assessed through explanatory generalized additive models for location, scale, and shape. *Frontiers in Public Health*, 10, 815036.
7. Sinnige, A., Kittelson, A., Van der Wees, P. J., Teijing, J. A. W., and Hoogeboom, T. J. (2022). Personalised outcomes forecasts of supervised exercise therapy in intermittent claudication: An application of neighbours based prediction methods with routinely collected clinical data. *European Journal of Vascular and Endovascular Surgery*, 63(4), 594–601.
8. Tadphale, S. D., Zurakowski, D., Bird, L. E., Yohannan, T. M., Agrawal, V. K., Lloyd, H. G., Allen, K. J., Waller III, B. R., Hall, A. M., and Sathanandam, S. K. (2020). Construction of femoral vessel nomograms for planning cardiac interventional procedures in children 0–4 years old. *Pediatric Cardiology*, 41, 1135–1144.
9. Truckenbrod, C., Meigen, C., Brandt, M., Vogel, M., Wahl, S., Jurkutat, A., and Kiess, W. (2020). Reference curves for refraction in a German cohort of healthy children and adolescents. *Plos One*, 15(3), e0230291.
10. Voncken, L., Albers, C. J., and Timmerman, M. E. (2021). Bias-variance trade-off in continuous test norming. *Assessment*, 28(8), 1932–1948.

11. Zouiouich, S., Loftfield, E., Huybrechts, I., Viallon, V., Louca, P., Vogtmann, E., Wells, P. M., Steves, C. J., Herzig, K. H., Menni, C., Jarvelin, M. R., Sinha, R., and Gunter, M. J. (2021). Markers of metabolic health and gut microbiome diversity: findings from two population-based cohort studies. *Diabetologia*, 64, 1749–1759.

**ARTIGO 2 - Novas distribuições zero-ajustadas baseada na
família Box-Cox**

Novas distribuições zero-ajustadas baseada na família Box-Cox

Elias Manensa Sabe¹, Luiz R. Nakamura², Andréa Cristina Konrath³, Luiz O. O. Pala²

¹Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras, Lavras, Brazil.

²Departamento de Estatística, Universidade Federal de Lavras, Lavras, Brazil.

³Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil.

RESUMO

A modelagem de dados é um objeto central em estatística e os modelos de regressão fornecem uma estrutura essencial para essa tarefa. Diferentes distribuições estatísticas são utilizadas para explicar variados fenômenos na natureza, desde os mais simples até os mais complexos. Neste artigo, propomos versões modificadas de duas distribuições pertencentes à família Box-Cox, baseada nas distribuições Box-Cox Cole e Green (BCCG) e Box-Cox exponencial potência (BCPE), capazes de capturar comportamentos assimétricos e com diferentes graus de curtose em variáveis resposta estritamente positivas. Nesta versão, um componente adicional que incorpora uma probabilidade de ocorrência de valores exatamente iguais a zero é introduzido em ambas distribuições, originando as distribuições BCCG zero-ajustada (zBCCG) e BCPE zero-ajustada (zBCPE). Ambas as distribuições são integradas à estrutura dos modelos aditivos generalizados para localização, escala e forma (GAMLSS), com o objetivo de aplicações em um contexto preditivo. Desenvolvemos artefato matemático, conduzimos um estudo de simulação de Monte Carlo e, posteriormente, aplicamos os modelos zero-ajustados em um conjunto de dados real referente a seguros agrícolas para a cultura de soja. Os resultados destacam a capacidade das distribuições propostas de capturar padrões assimétricos e caudas pesadas, frequentes em dados reais com excesso de zeros.

Palavras-chave: Distribuições zero-ajustadas; GAMLSS; Regressão distribucional.

1 Introdução

A modelagem de dados é um objetivo central em estatística, e os modelos de regressão fornecem uma estrutura essencial para essa tarefa. Modelos lineares (LM) assumem uma relação linear entre as variáveis preditoras (ou explicativas) e a resposta Y , onde os valores das variáveis preditoras são fixos, a média do erro é nula e a variância do erro é constante (Rencher; Schaalje, 2008). No entanto, a simplicidade e os pressupostos rigorosos dos modelos lineares (LM) podem ser limitantes em muitas aplicações, motivando o desenvolvimento dos modelos lineares generalizados (GLM) (Nelder; Wedderburn, 1972) que possibilitam considerar outras distribuições de probabilidade para capturar o comportamento da resposta, desde que pertencentes à família exponencial, e modelam uma função da média da resposta a partir das covariáveis disponíveis. Subsequentemente, os modelos aditivos generalizados (GAM), desenvolvidos por Hastie e Tibshirani (1990), introduziram funções de suavização não-paramétricas à estrutura dos GLM para capturar relações não-lineares complexas entre as variáveis preditoras e uma função da média da resposta, aumentando assim a flexibilidade da modelagem.

Entretanto, apesar dos avanços proporcionados pelos GAM, ainda podemos observar duas importantes limitações: apenas o parâmetro da média é modelado explicitamente e a distribuição da variável resposta deve, necessariamente, pertencer à família exponencial. Para superar essas limitações, Rigby e Stasinopoulos (2005) propuseram os modelos aditivos generalizados para localização, escala e forma (GAMLSS), atualmente também denominados como modelos de regressão distribucional (Heller et al., 2022). Os GAMLSS formam uma classe de modelos de regressão univariados semi-paramétricos, nos quais qualquer distribuição \mathcal{D} pode ser assumida para representar a variável resposta, e todo e qualquer parâmetro desta distribuição pode ser modelado diretamente por meio de funções lineares e de suavização não-paramétricas, isto é, estruturas de regressão distintas são construídas para cada um dos parâmetros (Nakamura et al., 2022b).

As distribuições que podem ser assumidas para a resposta são divididas em três grandes grupos: distribuições contínuas, discretas e mistas, ou seja, uma mistura das duas primeiras sendo uma parte discreta e uma parte contínua (Rigby et al., 2019). Dentre elas, destacam-se aquelas capazes de modelar dados extremamente assimétricos e com diferentes graus de curtose, particularmente as pertencentes à família de distribuições Box-Cox (Nakamura et al., 2022a), tendo como casos particulares, por exemplo, as distribuições Box-Cox Cole e Green (*Box-Cox*

Cole and Green, BCCG) (Cole; Green, 1992) e a Box-Cox exponencial potência (*Box-Cox Power Exponential*, BCPE) (Rigby; Stasinopoulos, 2004).

Neste artigo temos interesse particular nas distribuições que são capazes de modelar respostas positivas no contexto dos modelos de regressão, como as da família de distribuições Box-Cox, mas que também incorporam uma probabilidade extra de não ocorrência do evento de interesse, problema comum em diversas áreas aplicadas como, por exemplo, nas ciências atuariais (Bortoluzzo et al., 2011), microbiologia (Rocha et al., 2017), e meteorologia (Zamani; Bazrafshan, 2020). Assim, o objetivo principal deste trabalho é propor a construção de duas distribuição zero-ajustadas, isto é, que são adequadas para os casos em que temos como resposta $0 \leq Y < \infty$, com base na família de distribuições Box-Cox, mais especificamente das distribuições BCCG e BCPE, implementando-as no contexto dos GAMLSS.

2 Estrutura dos GAMLSS

Matematicamente, consideremos uma variável resposta \mathbf{Y} qualquer. Os GAMLSS podem ser definidos como

$$\begin{aligned} \mathbf{Y} &\stackrel{\text{ind}}{\sim} \mathcal{D}(\boldsymbol{\theta}_k) \\ g_k(\boldsymbol{\theta}_k) &= \mathbf{X}_k \boldsymbol{\beta}_k + \sum_{j=1}^{J_k} s_{kj}(\mathbf{x}_{kj}), \end{aligned} \quad (1)$$

em que \mathcal{D} é a distribuição assumida para a variável resposta com vetor de parâmetros $\boldsymbol{\theta}_k = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p)^\top$, $g_k(\cdot)$, $k = 1, \dots, p$, denotam as funções de ligação relacionadas aos k parâmetros da distribuição - usualmente escolhidas com base no espaço paramétrico (De Bastiani et al., 2018), \mathbf{X}_k é uma matriz de delineamento, $\boldsymbol{\beta}_k = (\beta_{1k}, \dots, \beta_{J_k k})^\top$ é um vetor de parâmetros associados à parte linear da estrutura de regressão e $s_{kj}(\cdot)$ é uma função de suavização - neste artigo considerado como um spline cúbico (Wood, 2017) - que explica relacionamentos não-lineares com a covariável \mathbf{x}_{kj} .

Sob essa perspectiva, percebe-se que, nos GAMLSS, além da modelagem da média (ou um parâmetro de locação), qualquer outro parâmetro, como, por exemplo, relacionado à escala, assimetria ou curtose da distribuição também podem ser modelados como funções das covariáveis, proporcionando uma abordagem mais rica para capturar diferentes graus de complexidade da distribuição da resposta. Na equação (1) assume-se que qualquer distribuição \mathcal{D} (contínua,

discreta ou mista) pode ser utilizada para representar a variável resposta, independentemente dela pertencer ou não à família exponencial (Rigby et al., 2019).

2.1 Família de distribuição Box-Cox: uma breve revisão

As distribuições pertencentes à família Box-Cox ocupam posição proeminente na modelagem de respostas positivas contínuas no contexto dos GAMLSS (Nakamura et al., 2022a) e, segundo o Rigby et al. (2019), podem ser definidas a partir da transformação da variável aleatória Z dada por

$$Z = \begin{cases} \frac{1}{\sigma\nu} \left[\left(\frac{Y_1}{\mu} \right)^\nu - 1 \right] & , \text{ se } \nu \neq 0, \\ \frac{1}{\sigma} \log \left(\frac{Y_1}{\mu} \right) & , \text{ se } \nu = 0, \end{cases} \quad (2)$$

para $Y_1 > 0$, em que $\mu > 0$ corresponde exatamente à mediana, $\sigma > 0$ é o coeficiente de variação e $-\infty < \nu < \infty$ é um parâmetro associado à assimetria. Se Z segue uma distribuição normal padrão, então Y_1 seguirá uma distribuição BCCG, representada por $BCCG(\mu, \sigma, \nu)$. Da mesma maneira, se Z segue uma distribuição exponencial potência padrão, com parâmetro de potência $\tau > 0$, então Y_1 segue uma distribuição BCPE, denotada por $BCPE(\mu, \sigma, \nu, \tau)$. Na distribuição BCPE, o parâmetro τ está associado ao grau de curtose, tornando possível a modelagem tanto de respostas platicúrticas como leptocúrticas.

A função densidade de probabilidade (pdf) para $Y_1 \sim BCCG(\mu, \sigma, \nu)$, é então dada por

$$f_{Y_1}(y|\mu, \sigma, \nu) = \frac{y^{\nu-1} \exp(-0.5z^2)}{\mu^\nu \sigma \sqrt{2\pi} \Phi \left[(\sigma|\nu|)^{-1} \right]},$$

em que o termo $\Phi \left[(\sigma|\nu|)^{-1} \right]$ é a probabilidade truncada.

Por outro lado, a distribuição Box-Cox exponencial potência (*Box-Cox Power Exponential*, BCPE) (Rigby; Stasinopoulos, 2004) é uma generalização que adiciona um parâmetro adicional τ para modelar a curtose. Assim, $Y_1 \sim BCPE(\mu, \sigma, \nu, \tau)$ é definido por uma transformação da variável aleatória Z como em 2, com pdf dada por

$$f_{Y_1}(y | \mu, \sigma, \nu, \tau) = \frac{y^{\nu-1} f_T(z)}{\mu^\nu \sigma F_T \left[(\sigma|\nu|)^{-1} \right]},$$

em que

$$f_T(z) = \frac{\tau \exp\left(-|zc^{-1}|^\tau\right)}{2c\Gamma(\tau^{-1})},$$

$$c^2 = \Gamma(\tau^{-1}) \left[\Gamma(3/\tau)^{-1} \right] e$$

$$F_T(\cdot) = 0.5 \left[1 + \frac{\gamma\left(\tau^{-1}, |zc^{-1}|^\tau\right)}{\Gamma(\tau^{-1})} \text{sinal}(z) \right],$$

em que $\gamma(\alpha, x) = \int_0^x w^{\alpha-1} e^{-w} dw$ é a função gama incompleta, são a pdf e a função acumulada da distribuição exponencial potência padronizada, isto é, $T \sim PE(0, 1, \tau)$. Note que a distribuição BCCG é um caso particular da distribuição BCPE quando $\tau = 2$ (Rigby; Stasinopoulos, 2004), sendo adequada nos casos em que podemos assumir que a resposta é mesocúrtica. Maiores informações acerca de ambas as distribuições podem ser encontradas em Rigby et al. (2019).

Assim, no contexto dos GAMLSS, podemos reescrever o modelo (1) com base na distribuição BCPE, considerando funções de ligação adequadas, como

$$\begin{aligned} \log(\mu) &= X_1\beta_1 + \sum_{j=1}^{J_1} s_{1j}(x_{1j}), \\ \log(\sigma) &= X_2\beta_2 + \sum_{j=1}^{J_2} s_{2j}(x_{2j}) \\ \mathbf{v} &= X_3\beta_3 + \sum_{j=1}^{J_3} s_{3j}(x_{3j}) \\ \log(\tau) &= X_4\beta_4 + \sum_{j=1}^{J_4} s_{4j}(x_{4j}). \end{aligned} \tag{3}$$

No caso do modelo baseado na distribuição BCCG, as mesmas estruturas de (3) são consideradas, omitindo-se o modelo associado ao parâmetro relacionado à curtose τ .

2.2 Distribuições BCCG e BCPE zero-ajustadas

Qualquer distribuição contínua, cujo suporte da resposta esteja definido em $(0, \infty)$ pode ser estendida para uma distribuição zero-ajustada, ou seja, $0 \leq Y < \infty$, ao incorporar uma probabilidade adicional de não-ocorrência do evento de interesse (*ibid*). Conforme explicitado por Stasinopoulos et al. (2023), duas abordagens podem ser utilizados nestas situações: (i) a mode-

lagem baseada em modelos de duas partes, nos quais a probabilidade de zero e o componente contínuo são estimados separadamente; e (ii) os modelos Tobit (Tobin, 1958), que modelam simultaneamente ambas as características, mas têm como principal limitação a impossibilidade de modelar explicitamente a probabilidade de ocorrência de zeros.

Neste artigo, adotamos a primeira estratégia onde adicionamos uma probabilidade extra associada a ocorrência de zeros a um componente positivo, ou seja,

$$f_Y(y|\pi_0, \boldsymbol{\theta}_k) = \begin{cases} \pi_0 & \text{se } y = 0 \\ (1 - \pi_0)f_{Y_1}(y|\boldsymbol{\theta}_k) & \text{se } y > 0 \end{cases} \quad (4)$$

em que $\pi_0 = P(Y = 0)$, $0 < \pi_0 < 1$, é a probabilidade do evento de interesse não ocorrer e $f_{Y_1}(y|\boldsymbol{\theta}_k)$ é qualquer pdf com suporte na linha real positiva, com parâmetros $\boldsymbol{\theta}_k^\top = (\theta_1, \dots, \theta_p)$.

Se considerarmos que $f_{Y_1}(y|\boldsymbol{\theta}_k)$ é a densidade da distribuição BCCG ou da BCPE, teremos as novas distribuições BCCG zero-ajustada, denotada por $zBCCG(\pi_0, \mu, \sigma, \nu)$, e BCPE zero-ajustada, denotada por $zBCPE(\pi_0, \mu, \sigma, \nu, \tau)$, respectivamente. Assim, na estrutura dos GAMLSS, podemos adicionar ao modelo (3) a estrutura de regressão específica relacionada ao parâmetro associado à probabilidade π_0 de ocorrência de zero, dada por

$$\text{logit}(\pi_0) = \log\left(\frac{\pi_0}{1 - \pi_0}\right) = \mathbf{X}_\pi \boldsymbol{\beta}_\pi + \sum_{j=1}^{J_\pi} s_{\pi_j}(\mathbf{x}_{\pi_j}). \quad (5)$$

No que tange ao processo de estimação dos GAMLSS, em geral, maximizamos o logaritmo da função de verossimilhança penalizada (Rigby; Stasinopoulos, 2005). Na estrutura dos modelos baseados em distribuições inflacionadas de zero, podemos observar pela equação (4) que o parâmetro π_0 é ortogonal aos parâmetros θ_k e, desta maneira, a estimação do modelo final pode ser realizada a partir da combinação de duas etapas distintas (Stasinopoulos et al., 2023): (i) a estimação do componente binário a partir de um conjunto de variáveis explicativas, isto é, o ajuste da equação (5); e (ii) a estimação dos parâmetros associados à parte positiva dos dados, isto é, o ajuste da equação (3).

2.3 Seleção de covariáveis e diagnóstico do modelo

Diferentes processos de seleção de covariáveis em cada uma das estruturas de regressão podem ser utilizadas. Neste artigo, o método *stepwise* foi aplicado para selecionar as covariáveis que afetam a probabilidade extra de zero, isto é, para o componente binário do modelo.

Em relação à parte contínua positiva, utilizamos a Estratégia A, baseada no *stepwise*, um dos procedimentos mais utilizados para este fim nos GAMLSS (Ramires et al., 2021). No caso da distribuição BCPE, os seguintes passos, baseados no critério de informação de Akaike (AIC) (Akaike, 1974), são empregados:

- a) um procedimento *forward* é realizado para selecionar as covariáveis que irão compor a estrutura de μ , considerando σ , ν e τ como constantes;
- b) considerando o modelo obtido no passo (1) e ν e τ constantes, um procedimento *forward* é realizado para selecionar covariáveis para σ ;
- c) considerando os modelos obtidos nos passos (1) e (2) e τ constante, um novo *forward* é realizado para identificar as covariáveis para ν ;
- d) considerando os passos (1), (2) e (3), realiza-se um procedimento *stepwise* para selecionar as variáveis para τ ;
- e) com base nos modelos obtidos nos passos (1), (2) e (4), realiza-se um procedimento *backward* no modelo obtido em (3) para ν ;
- f) utilizando os modelos obtidos em (1), (4) e (5), um procedimento *backward* no modelo obtido em (2) é realizado para σ ;
- g) no último passo, a partir dos modelos obtidos em (4), (5) e (6), realiza-se um último procedimento *backward* no modelo obtido em (1), para selecionar o modelo final para μ .

No final do procedimento, o subconjunto de covariáveis adicionadas em cada uma das estruturas de regressão não necessariamente será o mesmo (Nakamura et al., 2017).

Para verificar a adequação dos modelos ajustados, conduzimos uma análise de diagnóstico baseada nos resíduos quantílicos normalizados (Dunn; Smyth, 1996), definidos para uma distribuição zero-ajustada como

$$\hat{r} = \Phi^{-1} [F_Y(y|\hat{\theta})], \quad (6)$$

em que $\Phi^{-1}(\cdot)$ é a inversa da função distribuição acumulada (cdf) da distribuição normal padrão, $F_Y(y|\hat{\theta}) = \pi_0 + (1 - \pi_0)F_{Y_1}(y|\hat{\theta})$ é a cdf da distribuição zero-ajustada e $F_{Y_1}(y|\theta)$ é a cdf

da distribuição relacionada à parte contínua positiva do modelo. A grande vantagem da utilização deste tipo de resíduo é que, caso o modelo ajustado seja adequado, independentemente da distribuição atribuída à resposta, os resíduos seguirão uma distribuição normal padrão (Stasinopoulos et al., 2017). Usualmente, esses resíduos são verificados a partir de um worm plot (Buuren; Fredriks, 2001) e, caso não apresentem quaisquer tendências (linear, quadrática ou cúbica), o modelo ajustado é adequado para explicar o conjunto de dados. Ademais, a seleção do modelo final pode ser realizada com base no AIC ou no critério de informação bayesiano (BIC) (Schwarz, 1978).

3 Estudo de simulação

Com o intuito de investigar o comportamento dos estimadores de máxima verossimilhança dos parâmetros para diferentes tamanhos de amostra ($n = 30$ e $n = 100$), nesta seção, apresentamos estudos de simulação de Monte Carlo para as distribuições zBCCG e zBCPE, considerando diferentes cenários, gerando amostras aleatórias provenientes de variáveis aleatórias seguindo ambas as distribuições, obtidas a partir de suas respectivas funções quantílicas, isto é, a partir do inverso de suas cdfs.

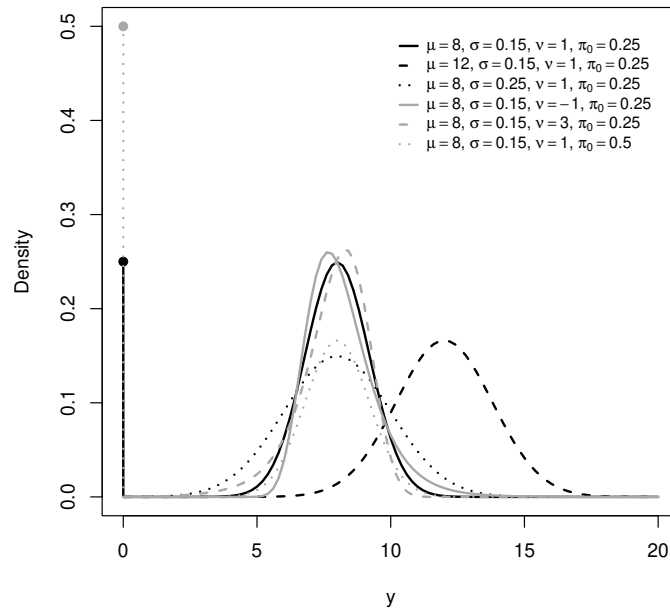
3.1 Distribuição zBCCG

Os cenários considerados para esta simulação contemplam os casos em que a distribuição zBCCG apresenta diferentes valores para a mediana ($\mu = 8$ e $\mu = 12$), diferentes valores para o coeficiente de variação ($\sigma = 0,15$ e $\sigma = 0,25$), simetria ($\nu = 1$), assimetria positiva ($\nu = -1 < 1$), assimetria negativa ($\nu = 3 > 1$) e diferentes probabilidades da não ocorrência do evento ($\pi_0 = 0,25$ e $\pi_0 = 0,50$), conforme apresentado na Figura 3.1.

Com base nas diferentes combinações de valores de parâmetros possíveis, a Tabela 3.1 apresenta as médias das estimativas de máxima verossimilhança, vieses e erros quadráticos médios (EQM), baseado em 1.000 simulações, da distribuição zBCCG, considerando os seguintes cenários:

- Cenário 1: $\mu = 8, \sigma = 0.15, \nu = 1, \pi_0 = 0.25$
- Cenário 2: $\mu = 12, \sigma = 0.15, \nu = 1, \pi_0 = 0.25$
- Cenário 3: $\mu = 8, \sigma = 0.25, \nu = 1, \pi_0 = 0.25$

Figura 3.1 – Distribuição zBCCG para diferentes combinações de valores de parâmetros



Fonte: Autores (2024)

- Cenário 4: $\mu = 8, \sigma = 0.15, \nu = -1, \pi_0 = 0.25$
- Cenário 5: $\mu = 8, \sigma = 0.15, \nu = 3, \pi_0 = 0.25$
- Cenário 6: $\mu = 8, \sigma = 0.15, \nu = 1, \pi_0 = 0.50$

Os resultados apresentados na Tabela 3.1 indicam que, conforme esperado, à medida que o tamanho amostral aumenta, as estimativas de máxima verossimilhança obtidas tendem a se aproximar cada vez mais do verdadeiro valor do parâmetro, além dos EQM obtidos diminuir, indicando assim o bom desempenho dos resultados em termos de consistência e eficiência.

3.2 Distribuição zBCPE

Para a simulação da distribuição zBCPE, uma vez que a distribuição zBCCG seu caso particular, além dos casos apresentados na Seção 3.1, apresentamos cenários em que a distribuição é leptocúrtica ($\tau = 1, 5 < 2$), isto é, possui caudas pesadas, bem como quando a distribuição é platicúrtica ($\tau = 3 > 2$), isto é, possui caudas leves, conforme apresentado na Figura 3.2.

Os cenários utilizados neste caso, foram divididos entre as combinações de valores de parâmetros que culminam em uma forma leptocúrtica ($\tau = 1, 5$) e aquelas cujas formas são platicúrticas ($\tau = 3$). Os resultados (estimativas de máxima verossimilhança, vieses e erros

Tabela 3.1 – Estimativas médias, vieses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCCG, considerando diferentes cenários.

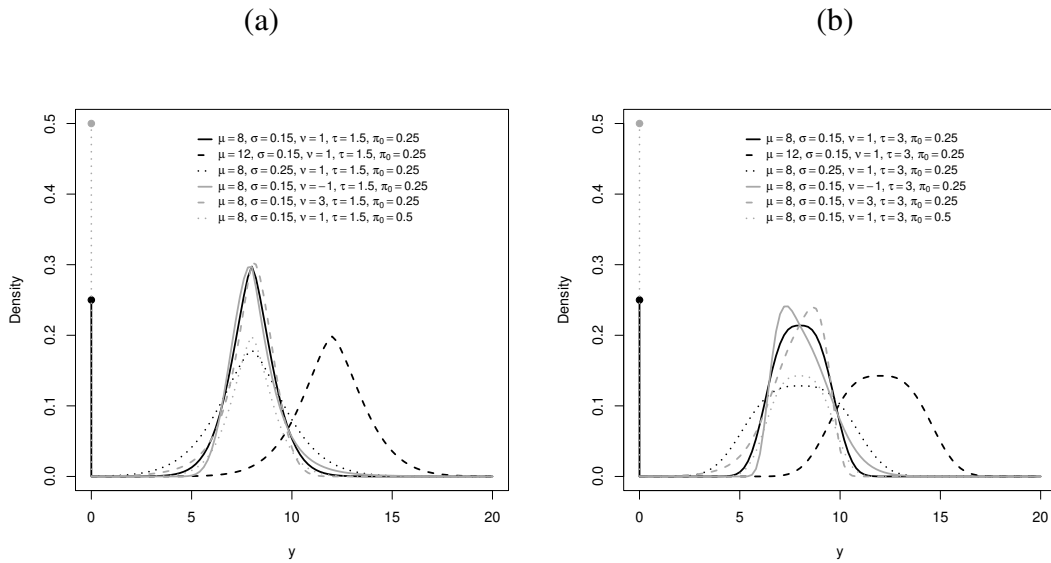
	Valor real	$n = 30$			$n = 100$		
		Estimativa	Viés	EQM	Estimativa	Viés	EQM
Cenário 1	$\mu = 8.000$	7.966	0.034	0.070	7.986	0.014	0.023
	$\sigma = 0.150$	0.142	0.008	0.001	0.147	0.003	0.000
	$\nu = 1.000$	1.167	0.167	3.639	0.984	0.016	0.492
	$\pi_0 = 0.250$	0.240	0.010	0.006	0.247	0.003	0.002
Cenário 2	$\mu = 12.000$	11.925	0.075	0.170	12.007	0.007	0.052
	$\sigma = 0.150$	0.143	0.007	0.001	0.146	0.004	0.000
	$\nu = 1.000$	1.121	0.121	4.082	0.997	0.003	0.526
	$\pi_0 = 0.250$	0.246	0.004	0.006	0.248	0.002	0.002
Cenário 3	$\mu = 8.000$	7.899	0.101	0.218	7.973	0.027	0.061
	$\sigma = 0.250$	0.251	0.001	0.003	0.249	0.001	0.001
	$\nu = 1.000$	1.220	0.220	1.592	1.075	0.075	0.356
	$\pi_0 = 0.250$	0.238	0.012	0.006	0.248	0.002	0.002
Cenário 4	$\mu = 8.000$	8.048	0.048	0.076	8.003	0.003	0.025
	$\sigma = 0.150$	0.143	0.007	0.001	0.146	0.004	0.000
	$\nu = -1.000$	-1.060	0.060	3.650	-0.984	0.016	0.496
	$\pi_0 = 0.250$	0.238	0.012	0.007	0.247	0.003	0.002
Cenário 5	$\mu = 8.000$	7.876	0.124	0.102	7.922	0.078	0.035
	$\sigma = 0.150$	0.158	0.008	0.002	0.161	0.011	0.001
	$\nu = 3.000$	3.131	0.131	3.760	3.277	0.277	1.396
	$\pi_0 = 0.250$	0.237	0.013	0.007	0.248	0.002	0.002
Cenário 6	$\mu = 8.000$	7.953	0.047	0.105	7.986	0.014	0.036
	$\sigma = 0.150$	0.141	0.009	0.001	0.145	0.005	0.000
	$\nu = 1.000$	1.216	0.216	7.359	0.997	0.003	0.914
	$\pi_0 = 0.500$	0.503	0.003	0.008	0.501	0.001	0.002

Fonte: Autores (2024)

quadráticos médios - EQM - baseado em 1.000 simulações) para a forma leptocúrtica são apresentados na Tabela 3.2 e contemplam os seguintes cenários:

- Cenário 1: $\mu = 8, \sigma = 0.15, \nu = 1, \tau = 1.5, \pi_0 = 0.25$
- Cenário 2: $\mu = 12, \sigma = 0.15, \nu = 1, \tau = 1.5, \pi_0 = 0.25$
- Cenário 3: $\mu = 8, \sigma = 0.25, \nu = 1, \tau = 1.5, \pi_0 = 0.25$
- Cenário 4: $\mu = 8, \sigma = 0.15, \nu = -1, \tau = 1.5, \pi_0 = 0.25$
- Cenário 5: $\mu = 8, \sigma = 0.15, \nu = 3, \tau = 1.5, \pi_0 = 0.25$
- Cenário 6: $\mu = 8, \sigma = 0.15, \nu = 1, \tau = 1.5, \pi_0 = 0.50$

Figura 3.2 – Distribuição zBCPE para diferentes combinações de valores de parâmetros: (a) forma leptocúrtica; e (b) forma platicúrtica



Fonte: Autores (2024)

A Tabela 3.3 apresenta as estimativas de máxima verossimilhança, vieses e erros quadráticos médios (EQM), baseado em 1.000 simulações, para os cenários que contemplam a forma platicúrtica, a saber:

- Cenário 1: $\mu = 8, \sigma = 0.15, \nu = 1, \tau = 3, \pi_0 = 0.25$
- Cenário 2: $\mu = 12, \sigma = 0.15, \nu = 1, \tau = 3, \pi_0 = 0.25$
- Cenário 3: $\mu = 8, \sigma = 0.25, \nu = 1, \tau = 3, \pi_0 = 0.25$
- Cenário 4: $\mu = 8, \sigma = 0.15, \nu = -1, \tau = 3, \pi_0 = 0.25$
- Cenário 5: $\mu = 8, \sigma = 0.15, \nu = 3, \tau = 3, \pi_0 = 0.25$
- Cenário 6: $\mu = 8, \sigma = 0.15, \nu = 1, \tau = 3, \pi_0 = 0.50$

Assim como o observado em relação à distribuição zBCCG na Subseção 3.1, os resultados obtidos para a distribuição zBCPE com forma leptocúrtica (Tabela 3.2) e com forma platicúrtica (Tabela 3.3) indicam boa performance do ponto de vista de consistência e eficiência.

4 Aplicação: seguro rural

Para esta aplicação, foram considerados dados de seguro agrícola para a cultura de soja, dentro do sub-setor de produtividade, direcionando a análise para o estado de Minas Gerais,

Tabela 3.2 – Estimativas médias, vieses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCPE, considerando diferentes cenários em que a forma da distribuição é leptocúrtica.

	Valor real	$n = 30$			$n = 100$		
		Estimativa	Viés	EQM	Estimativa	Viés	EQM
Cenário 1	$\mu = 8.000$	8.007	0.007	0.066	8.001	0.001	0.020
	$\sigma = 0.150$	0.158	0.008	0.045	0.146	0.004	0.000
	$v = 1.000$	1.217	0.217	4.589	1.034	0.034	0.688
	$\tau = 1.500$	1.504	0.004	0.745	1.669	0.169	0.325
	$\pi_0 = 0.250$	0.233	0.017	0.006	0.249	0.001	0.002
Cenário 2	$\mu = 12.000$	11.978	0.022	0.148	11.998	0.002	0.044
	$\sigma = 0.150$	0.157	0.007	0.040	0.147	0.003	0.000
	$v = 1.000$	1.133	0.133	4.398	1.108	0.108	0.809
	$\tau = 1.500$	1.522	0.022	0.719	1.651	0.151	0.291
	$\pi_0 = 0.250$	0.237	0.013	0.007	0.251	0.001	0.002
Cenário 3	$\mu = 8.000$	7.952	0.048	0.185	7.978	0.022	0.057
	$\sigma = 0.250$	0.268	0.018	0.249	0.252	0.002	0.002
	$v = 1.000$	1.229	0.229	1.676	1.094	0.094	0.398
	$\tau = 1.500$	1.460	0.040	0.776	1.627	0.127	0.303
	$\pi_0 = 0.250$	0.234	0.016	0.007	0.247	0.003	0.002
Cenário 4	$\mu = 8.000$	8.029	0.029	0.073	8.015	0.015	0.021
	$\sigma = 0.150$	0.157	0.007	0.027	0.149	0.001	0.001
	$v = -1.000$	-1.340	0.340	5.220	-1.100	0.100	1.151
	$\tau = 1.500$	1.502	0.002	0.741	1.654	0.154	0.315
	$\pi_0 = 0.250$	0.237	0.013	0.006	0.247	0.003	0.002
Cenário 5	$\mu = 8.000$	7.928	0.072	0.079	7.968	0.032	0.021
	$\sigma = 0.150$	0.162	0.012	0.019	0.153	0.003	0.001
	$v = 3.000$	2.814	0.186	3.042	3.033	0.033	1.284
	$\tau = 1.500$	1.534	0.034	0.803	1.635	0.135	0.335
	$\pi_0 = 0.250$	0.239	0.011	0.006	0.249	0.001	0.002
Cenário 6	$\mu = 8.000$	7.992	0.008	0.097	8.004	0.004	0.031
	$\sigma = 0.150$	0.175	0.025	1.094	0.146	0.004	0.001
	$v = 1.000$	0.249	0.751	6.849	1.053	0.053	1.143
	$\tau = 1.500$	1.373	0.127	0.929	1.731	0.231	0.509
	$\pi_0 = 0.500$	0.480	0.020	0.008	0.499	0.001	0.002

Fonte: Autores (2024)

Brasil, obtidos do Ministério da Agricultura, Pecuária e Abastecimento (MAPA) em junho de 2023, abrangendo o período de maio de 2016 a dezembro de 2021. Trata-se de uma base de dados secundária de acesso público, disponível em Brasil (2024), onde os contratos analisados estão relacionados a apólices de seguro subsidiadas pelo governo federal por meio do Programa de Subvenção ao Prêmio do Seguro Rural. Minas Gerais possui um território de 586.513,983 km², com uma área de 2.022.682 ha de cultivo de soja em 2022 (IBGE, 2022b) e, em termos

Tabela 3.3 – Estimativas médias, vieses e erros quadráticos médios (EQM) baseado nas 1.000 simulações para a distribuição zBCPE, considerando diferentes cenários em que a forma da distribuição é platicúrtica.

	Valor real	$n = 30$			$n = 100$		
		Estimativa	Viés	EQM	Estimativa	Viés	EQM
Cenário 1	$\mu = 8.000$	7.981	0.019	0.089	7.995	0.005	0.028
	$\sigma = 0.150$	0.141	0.009	0.004	0.147	0.003	0.000
	$v = 1.000$	1.134	0.134	3.099	0.985	0.015	0.387
	$\tau = 3.000$	2.704	0.296	1.714	3.253	0.253	1.048
	$\pi_0 = 0.250$	0.230	0.020	0.006	0.246	0.004	0.002
Cenário 2	$\mu = 12.000$	12.003	0.003	0.192	11.994	0.006	0.061
	$\sigma = 0.150$	0.143	0.007	0.001	0.147	0.003	0.000
	$v = 1.000$	1.256	0.256	3.208	0.984	0.016	0.422
	$\tau = 3.000$	2.691	0.309	1.722	3.280	0.280	1.053
	$\pi_0 = 0.250$	0.230	0.020	0.007	0.246	0.004	0.002
Cenário 3	$\mu = 8.000$	7.950	0.050	0.243	7.991	0.009	0.072
	$\sigma = 0.250$	0.247	0.003	0.016	0.245	0.005	0.000
	$v = 1.000$	1.227	0.227	1.353	1.009	0.009	0.194
	$\tau = 3.000$	2.518	0.482	2.009	3.226	0.226	1.072
	$\pi_0 = 0.250$	0.233	0.017	0.006	0.250	0.000	0.002
Cenário 4	$\mu = 8.000$	8.018	0.018	0.086	8.011	0.011	0.026
	$\sigma = 0.150$	0.143	0.007	0.003	0.147	0.003	0.000
	$v = -1.000$	-1.190	0.190	3.140	-0.972	0.028	0.392
	$\tau = 3.000$	2.671	0.329	1.816	3.286	0.286	1.125
	$\pi_0 = 0.250$	0.232	0.018	0.007	0.245	0.005	0.002
Cenário 5	$\mu = 8.000$	7.883	0.117	0.107	7.922	0.078	0.035
	$\sigma = 0.150$	0.164	0.014	0.006	0.161	0.011	0.001
	$v = 3.000$	3.434	0.434	2.895	3.404	0.404	1.123
	$\tau = 3.000$	2.441	0.559	2.022	3.017	0.017	1.083
	$\pi_0 = 0.250$	0.239	0.011	0.006	0.246	0.004	0.002
Cenário 6	$\mu = 8.000$	7.974	0.026	0.128	7.985	0.015	0.038
	$\sigma = 0.150$	0.158	0.008	0.706	0.145	0.005	0.000
	$v = 1.000$	1.028	0.028	5.691	0.996	0.004	0.753
	$\tau = 3.000$	1.744	1.256	2.567	2.744	0.256	0.495
	$\pi_0 = 0.500$	0.483	0.017	0.008	0.499	0.001	0.003

Fonte: Autores (2024)

de produção, foi o sexto maior produtor de soja no Brasil, com um total de 7.639.648 toneladas (IBGE, 2022a).

O conjunto de dados contém 3.431 observações, das quais 3.268 (95,25%) indicam ausência de sinistro e tem como resposta a indenização em caso de sinistro (em reais). As potenciais covariáveis, provenientes do MAPA, Instituto Nacional de Meteorologia (INMET, 2023) e da Agência Nacional de Águas e Saneamento Básico (ANA, 2023), compreendem: localização geográfica das lavouras (coordenadas de latitude e longitude), área da lavoura (em ha), valor

do prêmio do seguro (em reais), nível de cobertura, produtividade estimada (em kg/ha), evento predominante e duração do contrato, número médio mensal de dias com precipitação, precipitação média mensal (em mm), pressão atmosférica média mensal (em milibares), temperatura média mensal (°C), velocidade média e máxima do vento mensal (em m/s), número máximo de dias consecutivos sem chuva, temperatura mínima observada (°C), temperatura média mensal do ponto de orvalho (°C), ocorrência de geada (considerada quando a temperatura mínima foi inferior a 4°C) e capacidade de água disponível (mm/m).

4.1 Modelo para $Y = 0$

Conforme descrito na Seção 2.2, iniciamos o processo de construção do modelo zero-ajustado, a partir do ajuste da estrutura de regressão para π_0 , com base em uma distribuição binomial, utilizando como método de seleção de covariáveis o procedimento *stepwise*. O modelo final para a probabilidade de não ocorrência do evento é dado por

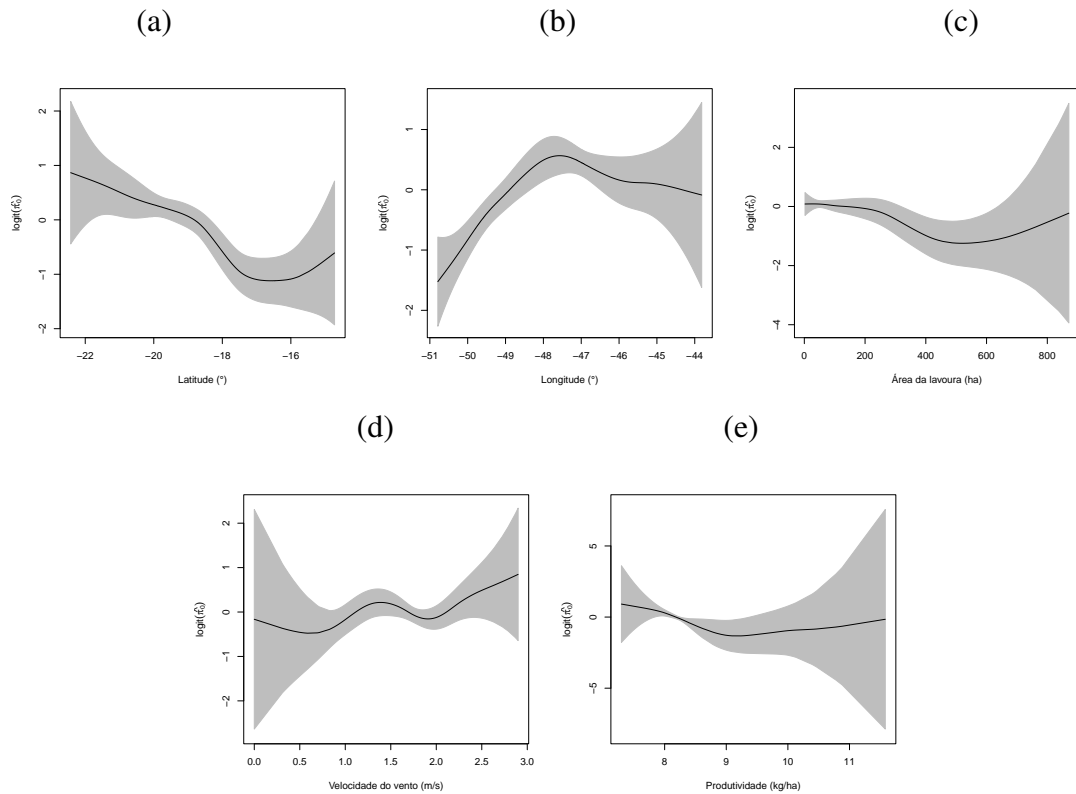
$$\log\left(\frac{\hat{\pi}_0}{1 - \hat{\pi}_0}\right) = 5,853 - 0,132 \text{ mínima} + 0,305 \text{ orvalho} + s_{\pi 1} (\text{latitude}) + s_{\pi 2} (\text{área}) \\ + s_{\pi 3} (\text{longitude}) + s_{\pi 4} (\text{velocidade}) + s_{\pi 5} (\text{produtividade}).$$

O resultado é similar ao obtido por Pala et al. (2024), que considerou apenas funções de suavização na estrutura de regressão, uma vez que o intuito de seu trabalho era puramente a predição da resposta. Aqui, temos, além do interesse em predizer adequadamente a resposta, simplificar, quando possível, a interpretação dos resultados, isto é, não considerar funções de suavização nos casos em que o relacionamento entre a covariável e o parâmetro modelado é aproximadamente linear.

As duas covariáveis que não necessitaram de funções de suavização para seu ajuste, temperatura mínima e orvalho, são significativas ao nível de 5% de significância. A cada incremento de 1 °C na temperatura mínima, há um decréscimo de, aproximadamente, 12,37% na chance de não-ocorrência de sinistro. Ao passo que, a cada aumento de 1 °C na temperatura de orvalho, há um aumento de, aproximadamente, 35,66% nesta mesma chance. Isto é lógico, uma vez, temperaturas mínimas mais baixas podem estar associadas a geadas, aumentando os riscos, enquanto temperaturas de orvalho mais altas indicam maior umidade, o que pode reduzir o risco de perda na lavoura. Em relação às outras covariáveis consideradas para explicar π_0 , as funções de suavização ajustadas são apresentadas na Figura 3.3. Conforme elucidado por Ramires et al.

(2019), usualmente não realizamos testes de hipóteses para verificar a significância da variável na estrutura de regressão, mas sim, verificamos seu comportamento graficamente.

Figura 3.3 – Splines cúbicas ajustadas para explicar a relação entre a probabilidade de não ocorrência de um sinistro, em que: (a) latitude (°); (b) longitude (°); (c) área da lavoura (ha); (d) velocidade média mensal do vento (m/s); e (e) logaritmo da produtividade estimada (kg/ha)



Fonte: Autores (2024)

A partir da Figura 3.3, podemos observar que determinadas áreas no estado de Minas Gerais apresentam maior risco de sinistros, possivelmente devido à concentração de contratos nessas regiões. Este padrão geográfico é corroborado pelos resultados de latitude (Painel (a)) e longitude (Painel (b)). A relação entre a área cultivada e a probabilidade de sinistros (Painel (c)) exibe um comportamento não linear: a probabilidade de ausência de sinistros diminui até 500 ha e, posteriormente, aumenta. Um padrão semelhante é observado no Painel (e), para a produtividade estimada, indicando que sistemas produtivos mais intensivos podem apresentar maior vulnerabilidade a riscos climáticos.

Finalmente, no Painel (d), a relação sinusoidal entre a velocidade do vento e a probabilidade de ausência de sinistros sugere que ventos moderados podem ser benéficos, enquanto ventos extremos representam um risco significativo às plantas e a produtividade da cultura de milho, como discutido por Bhandari et al. (2023). E estratégias como o uso de híbridos resis-

tentes ao vento (*ibid*) e gerenciamento cuidadoso de recursos como solo, água e biodiversidade (Bhattacharya, 2019) são recomendadas para mitigar os impactos.

4.2 Modelo para $Y > 0$

Na Seção 4.1, abordamos o modelo logit para a probabilidade de ocorrência de zeros (π_0), considerado uma etapa inicial para o ajuste de outros modelos voltados aos valores positivos. O objetivo inicial foi identificar as covariáveis associadas à π_0 . Nesta segunda etapa, focamos no modelo para os valores positivos de indenização.

A Estratégia A, descrita na Seção 2.3, foi aplicada para a seleção do melhor modelo baseado em ambas as distribuições zBCCG e zBCPE. A Tabela 3.4 apresenta os valores das medidas de AIC e BIC. Conforme podemos observar, o modelo que considera a distribuição zBCPE para a resposta é o melhor uma vez que apresenta os menores valores de ambas as métricas (AIC e BIC iguais a 3.870,21 e 3.922,80, respectivamente).

Tabela 3.4 – Critério de Informação de Akaike (AIC) para os modelos GAMLSS ajustados com as distribuições zBCPE e zBCCG

Distribuição	AIC	BIC
zBCPE	3.870,21	3.922,80
zBCCG	3.942,36	3.998,05

O modelo final baseado na distribuição zBCPE é dado por

$$\log(\hat{\mu}) = 15,684 + 0,208 \text{ latitude} + s_{11} (\text{área}) + s_{12} (\text{precipitação})$$

$$\log(\hat{\sigma}) = 18,498 + s_{21} (\text{CAD})$$

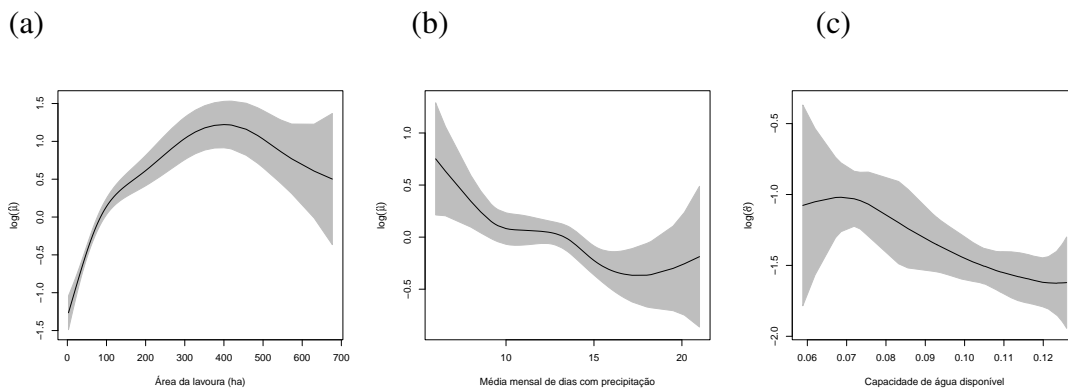
$$\hat{\nu} = -1,281$$

$$\log(\hat{\tau}) = -1,568.$$

A única covariável selecionada sem a necessidade de inserção de uma função de suavização foi a latitude na estrutura de regressão para μ , sendo significativa ao nível de 5%. Uma vez que $\exp(0,208) = 1,231$, a cada 1° na latitude, espera-se um aumento de, aproximadamente, 23,1% no valor mediano da indenização. Além da latitude, outras duas covariáveis foram selecionadas para explicar o parâmetro μ com base em uma função de suavização, a saber: a área e a precipitação. Com base na Figura 3.4 Painel (a), podemos verificar que o efeito do tamanho da área é positivo na indenização mediana para lavouras de até, aproximadamente,

400 ha e, após este limiar, o relacionamento passa a ser negativo. Este resultado é consistente, em termos paramétricos, com os achados de Pala e Sáfadi (2022). Ainda, o efeito da média mensal de dias com precipitação é negativo na indenização mediana até, aproximadamente, 16 dias e, posteriormente, permanece quase constante, com uma tendência positiva. Esse padrão é compatível com estudos como o de Seamon et al. (2023), que destacam a influência complexa da precipitação nas perdas seguradas.

Figura 3.4 – Splines cúbicas ajustadas para explicar a relação entre o valor mediano e variabilidade das indenizações (dado que um sinistro ocorreu) em que: (a) área da lavoura (ha); (b) média mensal de dias com precipitação; e (c) a capacidade de água disponível (CAD).



Fonte: Autores (2024)

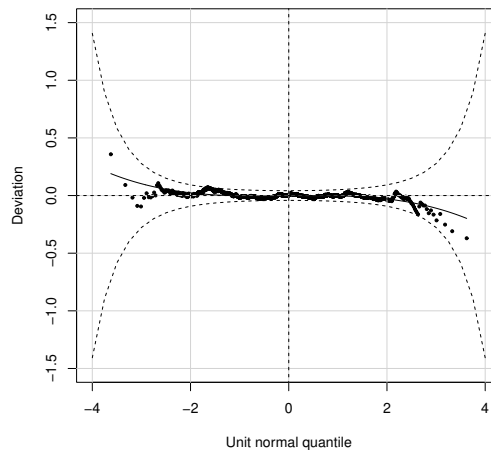
O Painel (c) apresenta o relacionamento da única covariável - CAD - selecionada para a estrutura do coeficiente de variação σ . A variabilidade é praticamente constante até, aproximadamente, 0,075 mm/m e, em seguida, o relacionamento é negativo. Este resultado é particularmente relevante para avaliar a incerteza no valor pago em caso de sinistro e para aplicações em seguros paramétricos, como discutido por Gasparetto et al. (2023).

Os demais parâmetros da distribuição, ν e τ , foram modelados a partir de uma constante. Os valores ajustados demonstram, de fato, a necessidade da utilização de uma distribuição flexível como a BCPE para a modelagem da parte positiva dos dados. O valor de indenização tem uma distribuição positiva acentuada, uma vez que $\hat{\nu} = -1,281 < 1$, além de possuir uma forma leptocúrtica ($\tau = \exp(-1,568) = 0,208 < 2$), isto é, apresenta caudas da distribuição pesadas (Rigby; Stasinopoulos, 2004).

A Figura 3.5 apresenta o worm plot do modelo zBCPE ajustado, isto é, considerando todas as estruturas de regressão da distribuição BCPE, bem como do componente binomial associado à probabilidade de não ocorrência de sinistro, conforme. Como podemos observar, o

modelo ajustado é razoável para explicar o conjunto de dados em estudo, uma vez que todos os resíduos encontram-se dentro das bandas de 95% de confiança.

Figura 3.5 – Worm plot obtido a partir do modelo GAMLSS zero-ajustado com base na distribuição BCPE



Fonte: Autores (2024)

5 Considerações finais

Neste artigo, propomos duas novas distribuições zero-ajustadas - Box-Cox Cole e Green zero-ajustada (zBCCG) e Box-Cox exponencial potência zero-ajustada (zBCPE) -, isto é, que consideram além de uma resposta estritamente positiva, uma probabilidade extra $P(Y = 0) = \pi_0$ de não ocorrência do evento. As duas novas distribuições são apresentadas no contexto dos modelos aditivos generalizados para locação, escala e forma (GAMLSS), possibilitando assim, a sua utilização na construção de modelos preditivos. Por meio de estudos de simulação de Monte Carlo, verificamos que os estimadores de máxima verossimilhança para os parâmetros de ambas as distribuições fornecem resultados consistentes, considerando diferentes tamanhos amostrais e cenários: distribuição simétrica, assimétrica positiva e negativa e, no caso da distribuição zBCPE, formas platicúrticas e leptocúrticas, isto é, com caudas pesadas e leves, respectivamente. A potencialidade dos modelos de regressão baseados nessas novas distribuições foi apresentada em uma aplicação prática em dados relacionados ao seguro rural, especialmente sobre a distribuição zBCPE, devido à sua capacidade de capturar adequadamente tanto o padrão assimétrico positivo, quanto o excesso de curtose presente nos valores de indenização.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. LR Nakamura recebeu apoio parcial da FAPEMIG/Brasil, RED-00133-21.

REFERÊNCIAS

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716-723, 1974.

ANA. **Capacidade de água disponível - CAD ou AWC dos solos no Brasil**. Acessado em: jul. 03, 2023. 2023.

BHANDARI, A.; KHAREL, T.; REDDY, K. Assessing wind damage and potential yield loss in mid-season corn using a geospatial approach. **Frontiers in Agronomy**, v. 5, p. 1–7, 2023.

BHATTACHARYA, A. Chapter 1 - global climate change and its impact on agriculture. In: BHATTACHARYA, A. (Ed.). **Changing Climate and Resource Use Efficiency in Plants**. Academic Press, 2019. p. 1–50. ISBN 978-0-12-816209-5. URL: <<https://www.sciencedirect.com/science/article/pii/B9780128162095000015>>.

BORTOLUZZO, A. B.; CLARO, D. P.; CAETANO, M. A. L.; ARTES, R. Estimating total claim size in the auto insurance industry: a comparison between Tweedie and zero-adjusted inverse Gaussian distribution. **BAR-Brazilian Administration Review**, v. 8, p. 37–47, 2011.

BRASIL. **Dados do PSR 2024**. URL:

<<https://www.gov.br/agricultura/pt-br/assuntos/riscos-seguro/seguro-rural/dados>. Acessado em: jun.27,2023>. 2024.

BUUREN, S. v.; FREDRIKS, M. Worm plot: a simple diagnostic device for modelling growth reference curves. **Statistics in Medicine**, v. 20, p. 1259–1277, 2001.

COLE, T. J.; GREEN, P. J. Smoothing reference centile curves: the LMS method and penalized likelihood. **Statistics in Medicine**, v. 11, p. 1305–1319, 1992.

DE BASTIANI, F.; RIGBY, R. A.; STASINOPOULOS, D. M.; CYSNEIROS, A. H.; URIBE-OPAZO, M. A. Gaussian Markov random field spatial models in GAMLSS. **Journal of Applied Statistics**, v. 45, p. 168–186, 2018.

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, v. 5, p. 236–244, 1996.

GASBARETTO, S. C.; DAL POZZO, B. S.; OZAKI, V. A. Modelagens atuariais para o seguro agrícola. **P. 106**, 69, 2023.

HASTIE, T. J.; TIBSHIRANI, R. J. in **Statistical Models in S**. p. 249–307. Boca Raton: CRC Press, 1990.

HELLER, G. Z.; ROBLEDO, K. P.; MARSCHNER, I. C. Distributional regression in clinical trials: treatment effects on parameters other than the mean. **BMC Medical Research Methodology**, v. 22, p. 56, 2022.

IBGE. **Economia**. URL:

<<https://cidades.ibge.gov.br/brasil/mg/pesquisa/14/10239?tipo=ranking&indicador=10368>>.

Acessado em: fev. 16, 2024. 2022.

IBGE. **Produção Agrícola**. URL: <<https://cidades.ibge.gov.br/brasil/mg/pesquisa/14/10366>>.

Acessado em: fev. 16, 2024. 2022.

INMET. **Banco de Dados Meteorológicos do INMET**. URL:

<<https://bdmep.inmet.gov.br/#>>. Acessado em: jul. 02, 2023. 2023.

NAKAMURA, L. R.; RAMIRES, T.; RIGHETTO, A.; SILVA, V.; KONRATH, A. Using the Box-Cox family of distributions to model censored data: a distributional regression approach.

Brazilian Journal of Biometrics, v. 40, p. 407–414, 2022.

NAKAMURA, L. R.; RAMIRES, T. G.; RIGHETTO, A. J.; PESCIM, R. R.; ROQUIM, F. V.; SAVIAN, T. V.; STASINOPOULOS, D. M. Cattle reference growth curves based on centile estimation: a GAMLSS approach. **Computers and Electronics in Agriculture**, v. 192, p. 106572, 2022.

NAKAMURA, L. R.; RIGBY, R. A.; STASINOPOULOS, D. M.; LEANDRO, R. A.; VILLEGAS, C.; PESCIM, R. R. Modelling location, scale and shape parameters of the Birnbaum-Saunders generalized t distribution. **Journal of Data Science**, v. 15, p. 221–237, 2017.

NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. **Journal of the Royal Statistical Society Series A: Statistics in Society**, v. 135, p. 370–384, 1972.

PALA, L. O. O.; NAKAMURA, L. R.; SABE, E. M.; RAMIRES, T. G. Agricultural insurance and climate variability: analyzing soybean yield risks in Brazil using distributional regression models. Submetido para publicação. 2024.

PALA, L. O.; SÁFADI, T. Uma abordagem estocástica para precificação e análise de indenizações de seguros agrícolas no Brasil. 2022.

RAMIRES, T. G.; NAKAMURA, L. R.; RIGHETTO, A. J.; PESCIM, R. R.; MAZUCHELI, J.; CORDEIRO, G. M. A new semiparametric Weibull cure rate model: fitting different behaviors within GAMLSS. **Journal of Applied Statistics**, v. 46, p. 2744–2760, 2019.

RAMIRES, T. G.; NAKAMURA, L. R.; RIGHETTO, A. J.; PESCIM, R. R.; MAZUCHELI, J.; RIGBY, R. A.; STASINOPOULOS, D. M. Validation of stepwise-based procedure in GAMLSS. **Journal of Data Science**, v. 19, p. 96–110, 2021.

RENCHER, A. C.; SCHAALJE, G. B. **Linear models in statistics**. John Wiley & Sons, 2008.

RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. **Journal of the Royal Statistical Society Series C: Applied Statistics**, v. 54, p. 507–554, 2005.

RIGBY, R. A.; STASINOPOULOS, D. M. Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. **Statistics in Medicine**, v. 23, p. 3053–3076, 2004.

RIGBY, R. A.; STASINOPOULOS, M. D.; HELLER, G. Z.; DE BASTIANI, F. **Distributions for modeling location, scale, and shape: Using GAMLSS in R**. Chapman and Hall/CRC, 2019.

ROCHA, L. O.; REIS, G. M.; FONTES, L. C.; PIACENTINI, K. C.; BARROSO, V. M.; REIS, T. A.; PEREIRA, A. A.; CORRÊA, B. Association between FUM expression and fumonisin contamination in maize from silking to harvest. **Crop Protection**, v. 94, p. 77–82, 2017.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, p. 461–464, 1978.

SEAMON, E.; GESSLER, P. E.; ABATZOGLOU, J. T.; MOTE, P. W.; LEE, S. S. Climatic damage cause variations of agricultural insurance loss for the Pacific Northwest region of the United States. **Agriculture**, v. 13, p. 2214, 2023.

STASINOPOULOS, D. M.; RIGBY, R. A.; HELLER, G. Z.; DE BASTIANI, F. P-splines and GAMLSS: a powerful combination, with an application to zero-adjusted distributions. **Statistical Modelling**, v. 23, p. 510–524, 2023.

STASINOPOULOS, M. D.; RIGBY, R. A.; HELLER, G. Z.; VOUDOURIS, V.; DE BASTIANI, F. **Flexible regression and smoothing: using GAMLSS in R**. CRC Press, 2017.

TOBIN, J. Estimation of relationships for limited dependent variables. **Econometrica: Journal of the Econometric Society**, p. 24–36, 1958.

WOOD, S. N. **Generalized Additive Models: An Introduction with R**. CRC Press, 2017.

ZAMANI, H.; BAZRAFSHAN, O. Modeling monthly rainfall data using zero-adjusted models in the semi-arid, arid, and extra-arid regions. **Meteorology and Atmospheric Physics**, v. 132, p. 239–253, 2020.

TERCEIRA PARTE - CONCLUSÃO

4 Conclusão

Os Modelos Aditivos Generalizados para Localização, Escala e Forma (GAMLSS), são reconhecidos como uma das abordagens de regressão mais flexíveis na literatura dada sua versatilidade para modelar respostas com diferentes características. Neste sentido o objetivo geral deste estudo foi construir distribuições contínuas com probabilidade extra de ocorrência de zeros, utilizando os GAMLSS. Inicialmente foi realizada um estudo bibliométrico utilizando as bases *Web of Science* e *Scopus*, no qual foram selecionados 593 artigos entre o período de 2004 e 2023, seguindo o protocolo PRISMA. Por meio do estudo bibliométrico verificou-se o impacto significativo dos GAMLSS em diversas áreas do conhecimento, como evidenciado pela crescente produção científica e pelo reconhecimento de autores, revistas influentes na área e tópicos de tendência.

O desenvolvimento das novas distribuições zero-ajustadas – Box-Cox Cole e Green zero-ajustada (zBCCG) e Box-Cox exponencial potência (zBCPE) – foi validado por meio de estudos de simulação de Monte Carlo. Os estudos de simulação mostraram que os estimadores de máxima verossimilhança para os parâmetros das distribuições zBCCG e zBCPE fornecem resultados consistentes para diferentes tamanhos amostrais e cenários. No estudo de caso, a distribuição zBCPE demonstrou superioridade em relação à zBCCG, devido à sua capacidade de capturar caudas pesadas e padrões leptocúrticos nos valores de indenização.

Portanto, este trabalho representa uma contribuição importante para a evolução dos GAMLSS e para o desenvolvimento de novas distribuições zero-ajustadas, fornecendo uma base sólida para futuras pesquisas.

Para estudos futuros, sugere-se expandir o escopo bibliométrico, incluindo mais bases de dados e ferramentas alternativas para análise; investigar diferentes tipos de publicações, como livros e artigos de revisão em outros idiomas; alterar palavras-chave; realizar uma revisão sistemática; aprimorar o mapeamento bibliométrico, na inclusão de novas bases de dados na diversificação de tipos de dados, implementadas no pacote *Bibliometrix*, disponível no *software* R. Além disso, desenvolver estratégias computacionais para otimizar o ajuste de modelos zero-ajustados em grandes volumes de dados; e ampliar as aplicações das distribuições propostas para áreas como saúde pública e finanças, onde o excesso de zeros é uma característica frequente.

Além disso, as lacunas no mapeamento bibliométrico também foram identificadas, incluindo a necessidade de explorar novas bases de dados, diferentemente de bases de dados

bibliográficas suportadas por *Bibliometrix* e *Biblioshiny*, como, por exemplo, AGRIS (*International Information System for Agricultural Sciences and Technology* – FAO), CABI CAB DIRECT e FSTA (*Food Science and Technology Abstracts* – EBSCO) e design da rede de colaboração de coautoria mais amplas. Por outro lado, os GAMLSS, devido ao grande interesse e à ampla aplicabilidade em diversas áreas do conhecimento, ainda apresentam lacunas que podem ser investigadas e abordadas dentro dessa estrutura. Maiores detalhes podem ser encontrados no Artigo 1 desta dissertação, intitulado: “*Historical developments and research trends in GAMLSS*” em “*Supplementary Material 3*”.

REFERÊNCIAS

- AHMED, K. M.; DHUBAIB, B. E. A. Zotero: A bibliographic assistant to researcher. **Journal of Pharmacology and Pharmacotherapeutics**, SAGE Publications Sage India: New Delhi, India, v. 2, n. 4, p. 304–305, 2011.
- AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, Ieee, v. 19, n. 6, p. 716–723, 1974.
- ARIA, M.; CUCCURULLO, C. bibliometrix: An r-tool for comprehensive science mapping analysis. **Journal of informetrics**, Elsevier, v. 11, n. 4, p. 959–975, 2017.
- ARIAS, O. V. et al. Homogenisation of a soil properties map by principal component analysis to define index agricultural insurance policies. **Geoderma**, Elsevier, v. 311, p. 149–158, 2018.
- ASSELDONK, M. van et al. Agricultural risk management in the european union: A proposal to facilitate precautionary savings. **EuroChoices**, Wiley Online Library, v. 18, n. 2, p. 40–46, 2019.
- BANK, W. Agriculture e rural development. **Washington, DC: Development Data Group, The World Bank**, 2024. Accessed: 2024-07-04.
- BILGIN, N. A. et al. Agricultural insurance preferences of apricot farmers: the case of malatya. **Ciência Rural**, SciELO Brasil, v. 53, p. e20210785, 2022.
- BULTE, E.; LENSINK, R. Why agricultural insurance may slow down agricultural development. **American Journal of Agricultural Economics**, Wiley Online Library, v. 105, n. 4, p. 1197–1220, 2023.
- BUUREN, S. v.; FREDRIKS, M. Worm plot: a simple diagnostic device for modelling growth reference curves. **Statistics in medicine**, Wiley Online Library, v. 20, n. 8, p. 1259–1277, 2001.
- CAPES, P. P. et al. Portal periódicos capes. **Disponível em: <http://www.periodicos.capes.gov.br/ez88.periodicos.capes.gov.br/index.php>**, 2024.
- CARRER, M. J. et al. Determinants of agricultural insurance adoption: evidence from farmers in the state of são paulo, brazil. **RAUSP Management Journal**, SciELO Brasil, v. 55, n. 4, p. 547–566, 2020.
- CARTER, M. R.; CHENG, L.; SARRIS, A. Where and how index insurance can boost the adoption of improved agricultural technologies. **Journal of Development Economics**, Elsevier, v. 118, p. 59–71, 2016.
- CHEN, C. Citespace ii: Detecting and visualizing emerging trends and transient patterns in scientific literature. **Journal of the American Society for information Science and Technology**, Wiley Online Library, v. 57, n. 3, p. 359–377, 2006.
- COBO, M. J. et al. Scimat: A new science mapping analysis software tool. **Journal of the American Society for information Science and Technology**, Wiley Online Library, v. 63, n. 8, p. 1609–1630, 2012.

COLE, T. J.; GREEN, P. J. Smoothing reference centile curves: the lms method and penalized likelihood. **Statistics in medicine**, Wiley Online Library, v. 11, n. 10, p. 1305–1319, 1992.

CORDEIRO, G. M.; DEMÉTRIO, C. G. B.; MORAL, R. A. **Modelos Lineares Generalizados e Aplicações**. [S.l.]: Blucher, 2024.

COSTA, E. F. S. et al. Effects of biological and environmental factors on the variability of *paralonchurus brasiliensis* (sciaenidae) density: An gamlss application. **Journal of Sea Research**, Elsevier, v. 183, p. 102203, 2022.

CUCCURULLO, C.; ARIA, M.; SARTO, F. Foundations and trends in performance management. a twenty-five years bibliometric analysis in business and public administration domains. **Scientometrics**, Springer, v. 108, p. 595–611, 2016.

DE BASTIANI, F. et al. Gaussian markov random field spatial models in gamlss. **Journal of Applied Statistics**, Taylor & Francis, v. 45, n. 1, p. 168–186, 2018.

DE BOOR, C. **A Practical Guide to Splines**. [S.l.]: Springer, 1978.

DONTHU, N. et al. How to conduct a bibliometric analysis: An overview and guidelines. **Journal of Business Research**, v. 133, p. 285–296, 2021. ISSN 0148-2963. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0148296321003155>>.

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and graphical statistics**, Taylor & Francis, v. 5, n. 3, p. 236–244, 1996.

ECK, N. J. V.; WALTMAN, L. Citnetexplorer: A new software tool for analyzing and visualizing citation networks. **Journal of informetrics**, Elsevier, v. 8, n. 4, p. 802–823, 2014.

ECK, N. V.; WALTMAN, L. Software survey: Vosviewer, a computer program for bibliometric mapping. **scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV . . . , v. 84, n. 2, p. 523–538, 2010.

EILERS, P. H.; MARX, B. D. Flexible smoothing with b-splines and penalties. **Statistical science**, Institute of Mathematical Statistics, v. 11, n. 2, p. 89–121, 1996.

FERENHOF, H. A.; FERNANDES, R. F. Demystifying the literature review as basis for scientific writing: Ssf method. **Revista ACB**, v. 21, n. 3, p. 550–563, 2016.

GASPARETTO, S. C.; POZZO, B. S. D.; OZAKI, V. A. Modelagens atuariais para o seguro agrícola. **Pág. 106**, p. 69, 2023.

GOMPERTZ, B. Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. **Philosophical transactions of the Royal Society of London**, The Royal Society London, n. 115, p. 513–583, 1825.

HASTIE, T.; TIBSHIRANI, R. **Generalized additive models**. London: Chapman and Hall, 1990.

HE, C. et al. Frequency analysis of snowmelt flood based on gamlss model in manas river basin, china. **Water**, Multidisciplinary Digital Publishing Institute, v. 2021, n. 13, p. 2007, 2021.

- HELLER, G. Z.; ROBLEDO, K. P.; MARSCHNER, I. C. Distributional regression in clinical trials: treatment effects on parameters other than the mean. **BMC Medical Research Methodology**, BioMed Central, v. 22, n. 1, p. 1–12, 2022.
- IVEY, C.; CRUM, J. Choosing the right citation management tool: Endnote, mendeley, refworks, or zotero. **Journal of the Medical Library Association: JMLA**, Medical Library Association, v. 106, n. 3, p. 399, 2018.
- KUTLAR, I.; AKCAOZ, H. Inclinations of greenhouse farmers in turkey toward agricultural insurance. **Ciência Rural**, SciELO Brasil, v. 52, p. e20200704, 2021.
- LEITE, F. N. de S.; CASTRO, E. R. d.; TATEISHI, H. R. Regional impacts of rural credit and rural insurance policies on crop area and productivity: evidence from são paulo state, brazil (2008 and 2017). **Agricultural Finance Review**, Emerald Publishing Limited, v. 83, n. 2, p. 352–374, 2023.
- LIU, L. et al. Statistical analysis of zero-inflated nonnegative continuous data. **Statistical Science**, JSTOR, v. 34, n. 2, p. 253–279, 2019.
- LOM, I. V.; CAPRIAN, I. Mutual insurance funds to reducing agricultural risks-a theoretical approach. **Scientific papers. Series “Management, Economic Engineering in Agriculture and Rural Development”**, p. 399–408, 2023.
- MAPA. **PSR - Programa de Subvenção ao Prêmio do Seguro Rural**. 2024. Available at: <<https://mapa-indicadores.agricultura.gov.br/publico/extensions/SISSER/SISSER.html>>. Accessed: 2024-07-04.
- MOSCHINI, G.; HENNESSY, D. A. Uncertainty, risk aversion, and risk management for agricultural producers. **Handbook of agricultural economics**, Elsevier, v. 1, p. 87–153, 2001.
- NAKAMURA, L. R. et al. A new continuous distribution on the unit interval applied to modelling the points ratio of football teams. **Journal of Applied Statistics**, Taylor & Francis, v. 46, n. 3, p. 416–431, 2019.
- NAKAMURA, L. R. et al. Using the box-cox family of distributions to model censored data: a distributional regression approach. **Brazilian Journal of Biometrics**, v. 40, n. 4, p. 407–414, 2022.
- NAKAMURA, L. R. et al. Cattle reference growth curves based on centile estimation: A gamlss approach. **Computers and Electronics in Agriculture**, v. 192, p. 106572, 2022.
- NAKAMURA, L. R. et al. Modelling location, scale and shape parameters of the Birnbaum-Saunders generalized t distribution. **Journal of Data Science**, v. 15, p. 221–238, 2017.
- NEELON, B.; O’MALLEY, A. J.; SMITH, V. A. Modeling zero-modified count and semicontinuous data in health services research part 1: background and overview. **Statistics in Medicine**, Wiley Online Library, v. 35, n. 27, p. 5070–5093, 2016.
- NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. **Journal of the Royal Statistical Society: Series A (General)**, Wiley Online Library, v. 135, n. 3, p. 370–384, 1972.

- OLIVEIRA, T. A. et al. An application of generalized additive models of location, scale, and shape (GAMLSS) to estimate the eucalyptus height. **Ciência e Natura**, v. 42, p. 1–10, 2019.
- O’SULLIVAN, F. A statistical perspective on ill-posed inverse problems. **Statistical science**, JSTOR, p. 502–518, 1986.
- PAULA, G. A. **Modelos de regressão: com apoio computacional**. [S.l.]: IME-USP São Paulo, 2004.
- PEARCY, J.; SMITH, V. H. The tangled web of agricultural insurance: evaluating the impacts of government policy. **Journal of Agricultural and Resource Economics**, JSTOR, p. 80–111, 2015.
- PERSSON, O.; DANELL, R.; SCHNEIDER, J. W. How to use bibexcel for various types of bibliometric analysis. **Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday**, v. 5, n. 2009, p. 9–24, 2009.
- R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2024. Disponível em: <<https://www.R-project.org/>>.
- RAMIRES, T. G. et al. Predicting survival function and identifying associated factors in patients with renal insufficiency in the metropolitan area of maringá, paraná state, brazil. **Cadernos de Saúde Pública**, v. 34, n. 1, p. e00075517, 2018.
- RAMIRES, T. G. et al. A new semiparametric weibull cure rate model: fitting different behaviors within gamlss. **Journal of Applied Statistics**, Taylor & Francis, v. 46, n. 15, p. 2744–2760, 2019.
- RAMIRES, T. G. et al. A new semiparametric weibull cure rate model: fitting different behaviors within gamlss. **Journal of Applied Statistics**, Taylor & Francis, v. 46, n. 15, p. 2744–2760, 2019.
- RAMIRES, T. G. et al. Validation of stepwise-based procedure in gamlss. **Journal of Data Science**, , v. 19, n. 1, p. 96–110, 2021.
- RENCHER, A. C.; SCHAALJE, G. B. **Linear models in statistics**. [S.l.]: John Wiley & Sons, 2008.
- RIGBY, R. A.; STASINOPOULOS, D. M. Smooth centile curves for skew and kurtotic data modelled using the box–cox power exponential distribution. **Statistics in medicine**, Wiley Online Library, v. 23, n. 19, p. 3053–3076, 2004.
- RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, Wiley Online Library, v. 54, n. 3, p. 507–554, 2005.
- RIGBY, R. A.; STASINOPOULOS, M. D. Mean and dispersion additive models. In: SPRINGER. **Statistical Theory and Computational Aspects of Smoothing: Proceedings of the COMPSTAT’94 Satellite Meeting held in Semmering, Austria, 27–28 August 1994**. [S.l.], 1996. p. 215–230.
- RIGBY, R. A. et al. **Distributions for modeling location, scale, and shape: Using GAMLSS in R**. [S.l.]: CRC press, 2019.

- RIGHETTO, A. J. et al. Predicting weed invasion in a sugarcane cultivar using multispectral image. **Journal of Applied Statistics**, Taylor & Francis, v. 46, n. 1, p. 1–12, 2019.
- ROQUIM, F. V. et al. Building flexible regression models: including the birnbaum-saunders distribution in the gamlss package. **Semina: Ciências Exatas e Tecnológicas**, v. 42, n. 2, p. 163–168, 2021.
- STASINOPOULOS, D. M.; RIGBY, R. A. Generalized additive models for location scale and shape (gamlss) in r. **Journal of Statistical Software**, v. 23, p. 1–46, 2008.
- STASINOPOULOS, D. M. et al. P-splines and gamlss: a powerful combination, with an application to zero-adjusted distributions. **Statistical Modelling**, SAGE Publications Sage India: New Delhi, India, v. 23, n. 5-6, p. 510–524, 2023.
- STASINOPOULOS, M. D. et al. **Flexible regression and smoothing: using GAMLSS in R**. [S.l.]: CRC Press, 2017.
- TEAM, S. Science of science (sci2) tool. **Indiana University and SciTech Strategies**, Bloomington, IN, USA:, v. 379, 2009.
- TIMMERMAN, M. E.; VONCKEN, L.; ALBERS, C. J. A tutorial on regression-based norming of psychological tests with gamlss. **Psychological Methods**, v. 26, n. 3, p. 357–373, 2021.
- TOBIN, J. Estimation of relationships for limited dependent variables. **Econometrica: journal of the Econometric Society**, JSTOR, p. 24–36, 1958.
- TÜRKTEN, H.; YILDIRIM, Ç.; CEYHAN, V. Evaluating willingness to pay for agricultural insurance and determinants of agricultural insurance purchasing in samsun province of turkiye. **NEW MEDIT**, v. 2023, n. 2, 2023.
- WEN, S. et al. The impact of agricultural insurance on urban–rural income gap: Empirical evidence from china. **Agriculture**, MDPI, v. 13, n. 10, p. 1950, 2023.
- ZOU, B. et al. The role of agricultural insurance in boosting agricultural output: An aggregate analysis from chinese provinces. **Agribusiness**, Wiley Online Library, v. 38, n. 4, p. 923–945, 2022.