Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Estatística

JOSIMAR MENDES DE VASCONCELOS

Essays on New Models and the Mellin Transform: Experiments in Survival and Radar Data

Recife

2018

JOSIMAR MENDES DE VASCONCELOS

ESSAYS ON NEW MODELS AND THE MELLIN TRANSFORM: EXPERIMENTS IN SURVIVAL AND RADAR DATA

Doctoral thesis submitted to the Graduate Program in Statistics, Department of Statistics, Federal University of Pernambuco as a partial requirement for obtaining a doctorate in Statistics.

Advisor: DR. ABRAÃO D. C. NASCIMENTO
CO-ADVISOR: DR. RENATO J. CINTRA

Recife
2018

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

V331e Vasconcelos, Josimar Mendes de

Essays on New Models and the Mellin Transform: Experiments in Survival and Radar Data / Josimar Mendes de Vasconcelos. −2018.

221 f.: il., fig., tab.

Orientador: Abraão David Costa do Nascimento.

Tese (Doutorado) - Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2018.

Inclui referências e apêndices.

1. Estatística. 2. Análise de sobrevivência. I. Nascimento, Abraão David Costa do (orientador). ll. Título.

310 CDD (23. ed.) UFPE-MEI 2018-27

JOSIMAR MENDES DE VASCONCELOS

Essays on New Models and the Mellin Transform: Experiments in Survival and Radar Data

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Aprovada em: 05 de fevereiro de 2018.

BANCA EXAMINADORA

Prof. Abraão David Costa do Nascimento
UFPE

Prof. Francisco José de Azevêdo Cysneiros
UFPE

Prof.ª Audrey Helen Maria de Aquino Cysneiros
UFPE

Prof.ª Isis Didier Lins
UFPE

Prof. Eufrásio de Andrade Lima Neto
UFPB

Dedico esta tese a Deus, por ser o meu guia em tudo que faço. Consequentemente, sem distinção, dedico ao meu filho, Theófilo Paiva de Vasconcelos, à minha esposa Marília O. P. de Vasconcelos e ao complemento da minha família, como irmãs, Mikarla Mendes de Vasconcelos & Mirlândia Mendes de Vasconcelos, e de meus pais, José Kerginaldo de Vasconcelos & Marli Mendes de Vasconcelos.

Agradecimentos

A tese de doutorado faz parte do pacote de ferramentas para obter o título de doutor. Sua finalidade acadêmica é ser um trabalho individual para o amadurecimento do discente na carreira científica. Contudo, há contribuições em diversas proporções nesse trabalho que não podem deixar de ser destacadas. Por isso, cabe-me a minha gratidão através da menção a algumas pessoas e instituições que contribuíram para a realização desta tese:

Agradeço aos meus orientadores, Dr. Abraão David Costa do Nascimento e Dr. Renato J. Cintra, pela paciência, disponibilidade, companheirismo, compreensivos com minhas dificuldades, por todos os ensinamentos científicos e de vida. Em particular, dar-me a oportunidade de estudar um assunto interessante, tanto no aspecto de desafios matemáticos quanto estatísticos/computacionais. Além disso, deu-me a liberdade para buscar as minhas próprias ideias e interesses, sob sua orientação quando necessário.

Especialmente, lembro com todo carinho os momentos de confraternizações organizados pelo Dr. Abraão nos finais de ano; momentos de diálogos sociais, discussões acadêmicas, conversas sobre pesquisas atuais e futuras. Sempre serei muito grato por tudo que me fez nesses quatro anos de estudos, pesquisas e ensinamentos. Que Deus continue iluminando sua caminhada científica para que faça o melhor na vida de seus orientandos.

À minha esposa, Marília Vasconcelos, pelo amor, carinho, paciência, dedicação e apoio psicológico nessa fase final do doutorado na qual me forneceu um ambiente agradável para a elaboração desta tese. Além de tudo por dar o maior presente da vida, um filho abençoado por Deus. A meu filho, Theófilo Paiva de Vasconcelos, pela sua vinda durante o doutorado. Um amor de pessoa que me trouxe amor, carinho e paciência. Sempre te amarei e cuidarei de ti. Aos meus pais pelo interesse em me ver concluir o curso e porque sempre me deram o melhor que poderiam dar. Em especial a minha mãe, Marli Mendes de Vasconcelos, por cuidar de mim com carinho até hoje,

só amor de mãe para fazer isso. As minhas irmãs, Mikarla Mendes de Vasconcelos & Mirlândia Mendes de Vasconcelos, pelos momentos de descontrações, carinho, ensinamentos e por sempre cuidarem de tudo na minha ausência. Aos respectivos sogros, Elza Soares e João Maria, manifesto todo o carinho. Aos queridos cunhados, Marielly Oliveira e João Ewerton, externo minha gratidão. Também, a demais familiares que me ajudaram de forma indireta.

Quero registrar os meus agradecimentos aos amigos e companheiros do doutorado, mestrado e graduação pelo clima cordial que sempre reinou em nosso convívio social e intelectual, em especial aos amigos Rita (sempre em alto astral; agradeço pelo apoio, amor e carinho fraterno), Wenia (sempre bem humorada e tem historias populares engraçadas; sou grato pela sua cordial amizade), Thiago (sempre pronto a ajudar no que for preciso; sou grato pela sua amizade), Giannini (sempre tem histórias engraçadas que ocorreram consigo por onde passou; excelente amigo para tudo, pessoa que a gente gosta de estar ao lado). Além disso, tem os demais companheiros que não posso deixar de agradecer que de alguma forma contribuíram para esta tese: Elton Saulo, Jeremias Leão, Lutemberg Florêncio, Manoel Neto, Marcelo Bourguignon, Priscila Gonçalves, Claudio Tablada, Enai Taveira, Jéssica Rivas, Luciana Cecília, Marley Apolinário e Renilma Pereira.

A Universidade Federal do Piauí, campus Senador Helvídio Nunes de Barros (CSHNB) por conceder uma licença de trabalho por cerca de quatro anos para a realização deste doutorado. Aos colegas do departamento de matemática da UFPI-CSHNB, manifesto minha gratidão pelas palavras de incentivo durante a caminhada. Entre eles, destaco o João Santos, Francisco Gilberto e Daniel da Costa.

À Valéria Bittencourt, secretária do programa de pós-graduação em estatística, pela imensa competência e o cuidado no momento de me orientar em burocracias acadêmicas.

A todos os professores do departamento de estatística da UFPE, pelos conhecimentos que me foram passados.

Aos participantes da banca examinadora, agradeço antecipadamente pelos comentários e sugestões.

Agradeço a existência da linguagem LaTeX, softwares `R` e `wxMaxima`.

A arte de não submergir a face diante das dificuldades é uma dádiva de Deus que herdamos com o tempo, paciência, humildade e acima de tudo a força do querer ter o possível perante a nossa fraqueza.
(Autoria própria).

Abstract

Survival data have been applied in several contexts, such as survival time of mechanical components, the failure times of electrical insulator films, and in censored data from head-and-neck-cancer clinical trials. The resulting data are positive-valued and are often censored of heavy tails. This latter fact suggests that tailored tools are necessary for modelling survival data behavior. In particular, there is a need for flexible models; inferential methods, such as estimation and goodness-of-fit (GoF); and conditional representation (e.g., regression and time series models). Several models have been proposed to describe survival data based on distribution families derived from transformations of reference distributions (called baselines). One of the most important derived distribution families is the beta-G family introduced by Eugene *et al.* [Beta-normal distribution and its applications. *Communication in Statistics-Theory and Methods*, 31, 497-512]. Although the beta-G class is capable of producing even distributions for bimodal data, it requires both efficient estimation methods and GoF criteria. GoF methodology proposals are sought because not rarely it is hard to distinguish models within the beta-G class using: (i) criteria without a cut-off point rule or (ii) criteria originally suitable for nested models (e.g., the Akaike Information Criteria). Further, the likelihood function for beta-G models in real and synthetic experiments have suggested the proposal of estimation criteria which do not involve such function. In this thesis, the synthetic aperture radar (SAR) imagery is taken as a concrete context for data modelling. SAR is widely regarded as an important tool for remote sensing, partly because of its ability to operate independently of atmospheric conditions and producing images in high spatial resolution. However, features from SAR images are corrupted by a multiplicative noise that imposes the use of specifically designed probabilistic models. An important SAR feature is the SAR *intensity* image, which is defined as the norm of a complex return. Further, experiments with real SAR intensities often produce multimodal data. Several works aimed at modeling SAR intensity data by means of dis-

tribution mixtures, but such strategy may impose a large number of parameters. In this thesis, we adopt the Mellin transform as a way to derive new tools for the understanding of survival analysis data. With this we propose: (i) new qualitative and quantitative GoF measures suitable for survival analysis data and (ii) a unique estimation method not based on the likelihood function. In the context of SAR imagery analysis, we introduce: (i) two new probabilistic models: the compound Poisson-truncated Cauchy and the G-$\mathcal{G}$ family with three and four parameters, respectively; and (ii) a regression model at the $\mathcal{G}_I^0$ distribution for speckled data.

Keywords: Survival Analysis. SAR Imagery. Beta-G class. Mellin Transform. $\mathcal{G}_I^0$ Regression.

Resumo

Dados de sobrevivência foram aplicadas em vários contextos, tais como o tempo de sobrevivência dos componentes mecânicos, os tempos de falha dos filmes isolantes elétricos e em dados censurados de ensaios clínicos de câncer de cabeça e pescoço. Os dados resultantes têm natureza positiva, frequentemente marcadas por censura e caudas pesadas. Este último fato sugere o uso de modelos flexíveis; métodos de inferência tal como estimação e bondade de ajuste; e ajuste condicional (por exemplo, regressão e séries temporais). Vários modelos têm sido propostos para descrever dados de sobrevivência na forma de famílias definidas pela transformação de distribuições de referência (chamadas de *baselines*). Uma delas é a família beta-G proposta por Eugene *et al.* [Beta-normal distribution and its applications. *Communication in Statistics-Theory and Methods*, 31, 497-512]. Embora a classe beta-G seja capaz de produzir até mesmo distribuições para dados bimodais, ela carece tanto de métodos mais eficientes de bondade de ajuste como de estimação. A problemática que suscita a proposta de métodos de comparação de ajustes é distinguir a diferença entre dois modelos definidos na classe beta-G usando critérios: (i) sem uma regra envolvendo um ponto de corte associado ou (ii) próprios de modelos encaixados (a exemplo do critério de Informação de Akaike). Além disso, a função de verossimilhança para modelos beta-G em experimentos reais e sintéticos têm sugerido a proposta de critérios de estimação que não trabalhe diretamente com esta função. Nesta tese, imagens de radar de abertura sintética (*Synthetic Aperture Radar*-SAR) são consideradas como um contexto concreto para modelagem de dados. O SAR tem sido indicado como uma importante ferramenta para resolver problemas de sensoriamento remoto, isso se deve a sua capacidade de operar independente de condições atmosféricas e produzir imagens em alta resolução espacial. Entretanto, as imagens SAR têm seus atributos corrompidos por um ruído multiplicativo que impõe o uso de um modelo adaptado a sua presença. Uma importante característica em imagens SAR é a imagem em *intensidade* SAR, que é definida como a

norma de um retorno complexo. Além disso, experimentos com dados reais de intensidades SAR produzem dados multimodais. Vários trabalhos destinados a modelar dados de intensidade SAR usam misturas de distribuições, mas essa estratégia pode impor um grande número de parâmetros. Nesta tese, adotamos a transformada de Mellin como um caminho para obter novas ferramentas para a compreensão de dados de análise de sobrevivência. Com isso nós propomos: (i) uma nova medida de bondade de ajuste no aspecto qualitativo e quantitativo para dados de análise de sobrevivência e (ii) um método de estimação independente da verossimilhança. No contexto da análise de imagens SAR, apresentamos: (i) dois novos modelos de probabilidade: a composta Poisson-truncado Cauchy e a família G-$\mathcal{G}$ com três e quatro parâmetros, respectivamente; e (ii) um modelo de regressão na distribuição $\mathcal{G}_I^0$ para dados *speckled*.

Palavras-chave: Análise de Sobrevivência. Imagens SAR. Classe beta-G. Transformada de Mellin. Regressão $\mathcal{G}_I^0$.

List of Figures

List of Tables

List of Symbols

$\mathbb{N}$	Natural numbers or positive integers
$\mathbb{R}$	Real numbers
$\mathbb{R}_+$	Positive real numbers
$\mathbb{Z}_+$	Integer numbers
s	Complex number with unitary norm
T^2	Hotelling's statistic
$\psi(z)$	Digamma function
$\psi^{(\nu)}(z)$	Polygamma function
	(the νth derivative of the gamma function)
$\Phi_X(s)$	Characteristic function
$f(x)$	Probability density function
$F(x)$	Cumulative distribution function
$g(\bullet)$	Baseline Probability density function
$h(x)$	Hazard rate function
$Q[\bullet]$	Quantile function
$\phi_X(s)$	Mellin transform of a probability density function
$\psi_X(s)$	Second characteristic function of the second kind
$\widetilde{m}_\nu$	Log-moments
$\widetilde{\kappa}_\nu$	Log-cumulants
$\mathcal{X}$	Sample space
$\mathcal{N}^{-1}$	Inverse normal distribution
$\Gamma(\alpha)$	Complete gamma function

Γ	Gamma distribution
Γ^{-1}	Inverse gamma distribution
GΓ	Generalized gamma distribution
$B(\bullet,\bullet)$	Beta function
$\mathrm{E}(\bullet)$	Expected value
$M_X(t)$	Moment-generating function
$K_S(t)$	Cumulant generating function
$\mathcal{M}_{l,j,k}$	Probability weighted moment
$(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$	Log-cumulants diagram of the $\widetilde{\kappa}_3$ versus $\widetilde{\kappa}_2$
$K_\alpha(\bullet)$	Denotes the modified Bessel function of the third kind
$F_{\nu,n-\nu}$	F distribution with ν and $n-\nu$ degrees of freedom
$Q_F(\bullet;\nu,n-\nu)$	is the quantile function for $F_{\nu,n-\nu}$
$Q_{\chi^2}(\bullet;\nu)$	is the quantile function for the χ^2 distribution with ν degrees of freedom

List of Abbreviations and Acronyms

ACI	Asymptotic confidence interval
AIC	Akaike information criterion
AICc	Corrected AIC
beta-G	Generalized beta
BF	Beta-Fréchet
BIC	Bayesian information criterion
BKw	Beta-Kumaraswamy
BLL	Beta-log-logistic
BW	Beta-Weibull
cdf	Cumulative distribution function
cf	Characteristic function
cgf	Cumulant generating function
CTPC	Compound truncated-Poisson Cauchy
CTPG	Compound truncated-Poisson gamma
CTP-G	Compound truncated-Poisson-G
CTPN	Compound truncated-Poisson normal
CZTP	Compound zero-truncated Poisson
EGB2	Exponential generalized beta type 2
FIM	Fisher information matrix
FPPWM	Fractional partial probability weighted moment
FPWM	Fractional probability weighted moment
GEV	Generalized extreme value

GLM	Generalized linear models
GoF	Goodness-of-fit
hrf	Hazard rate function
IPPWM	Integer-order partial probability weighted moment
IPWM	Integer-order probability weighted moment
Kw-G	Kumaraswamy-G
Kw-LL	Kumaraswamy-log-logistic
KS	Kolmogorov-Smirnov
Lambda-G	Generalized lambda distribution
LC	Log-cumulant
LCE	LC estimators
LM	Log-moment
mgf	Moment-generating function
MO-G	Marshall and Olking-G
MLC	Matrix LC
ML	Maximum likelihood
MLE	ML estimator
MM	Method of moments
MulM	multiplicative model
MT	Mellin transform
Pareto-G	Generalized Pareto distribution
pdf	Probability density function
PPWM	Partial probability weighted moment
PWM	Probability weight moment
QD	Quadratic distance
qf	Quantile function
SAR	Synthetic aperture radar
TP	Zero-truncated Poisson

Contents

1 Introduction

In this chapter, we aim to give reasons for the two applied contexts used in this thesis: lifetime and synthetic aperture radar (SAR). Further, sections about main contributions, computational support, and the structure of this thesis are presented.

1.1 Motivation

Survival analysis tools have been applied in several contexts, such as lifetime of mechanical components [1], industrial [2] areas, breaking strengths of glass fibers [3], the failure times of electrical insulator films [4], biomedical science [5], and survival times from other practical experiments [6]. Survival data impose some difficulties (e.g., censured nature and resulting empirical distributions having heavy tail and asymmetric behavior) and, therefore, require the proposal of tailored modeling and inference methods. The derivation of new probability models capable of better explaining reliability data is a central task in the field of survival analysis. In recent years, families of distributions were developed, such as the Marshall and Olking (MO)-G [7], the Kumaraswamy (Kw)-G [8], the McDonald [9], the T-X [10, 11], and the generalized Weibull [12].

Despite the significant number of new models, few Goodness-of-fit (GoF) measures have been proposed for the recent distributions; thus hindering model selection. Considering progressive type-II censored data, Pakyari and Balakrishnan [13] proposed a general GoF test [14] that encompasses the GoF test described in [15]. Chen and Balakrishnan [15] derived an approximate GoF method as well. Such methods are based on distance measures between theoretical and empirical cumulative distribution functions (cdfs).

Taking a different approach, Linhart and Zucchini [16] proposed information-theoretical measures considering the Akaike and Bayesian information criteria [17] as figures of merit for model selection.

An alternative method for GoF assessment is given by the Pearson system diagram for model

selection [18], which is based on skewness and kurtosis measures [19, p. 23]. Delignon *et al.* [20] and Vogel and Fennessey [21] applied such diagram for SAR and hydrology data, respectively.

However, in [22], Nicolas noticed that the Pearson diagram tends to be not well-suited for positive random variables. Thus, the log-cumulant (LC) diagram, which plots the third-kind LC, $\widetilde{\kappa}_3$, against second-kind LC, $\widetilde{\kappa}_2$, was introduced as an alternative. The LCs diagram offers some advantages over the Pearson diagram. Besides being suitable for positive random variables, its computational implementation is casier and it also captures the distribution flexibility in the sense of skewness and kurtosis [23]. A detailed description of the LC diagram is provided in [22, 23]. Nicolas [22] proposed (i) the application of the Mellin transform (MT) [R. H. Mellin, 1854–1933] [24, p. 50] as an alternative to the usual characteristic function (cf) and (ii) a Pearson-like diagram based on LCs estimators instead of skewness and kurtosis. Such diagram was demonstrated to be relevant for quantitative and qualitative comparison of non-nested distributions in SAR and Polarimetric SAR (PolSAR) data [25–28].

In this thesis, we have as

First goal: to propose a combination of probability weighted moments (PWMs) and the MT in order to furnish new GoF measures which consider quantitative and qualitative aspects for models in the beta-G family.

On the other hand, models proposed from the beta-G family are often fitted by maximum likelihood estimators (MLEs). MLEs often have not closed-form expressions and, as a consequence, require the use of interactive optimization sources. Dias [29] discussed about some numerical issues which come from using MLEs in extended distributions. It is known that MLEs have well-defined asymptotic properties, but they may present an expressive bias for small and moderate sample sizes. Thus, this thesis considers that:

Second goal: proposing estimation criteria based on LC, not dependent on of the likelihood function, may be a good solution to numerical issues and events of flat likelihood, in which obtain from the use of MLEs in beta-G models.

As a second applied context, PolSAR images are processed obeying the following dynamic: Polarized pulses oriented at horizontal (H) and vertical (V) directions are emitted towards to an

under-study target. Subsequently, such pulses are recorded also at those two directions. One pixel of a PolSAR image concentrates information from four polarization channels in the form of four complex numbers, say HH, HV, VH, and VV.

The intensities (squared norm of four polarization channels) extracted from SAR images are important features for understanding SAR scenarios. These components can also be understood as the result of compounding sums of kind [30]:

$$S = X_1 + X_2 + \cdots + X_N, \tag{1.1}$$

where both N and $\{X_i\}$ are independent random variables.

Probability mixture models are often used to describe data of this nature, but these assumptions impose a great number of parameters and, as consequence, the inference stage becomes hard. In order to provide a more parsimonious solution, we propose a probability distribution with only three parameters which is able to describe data with multiple modals. Our model is defined by the sum of a random number N following the truncated Poisson model of independent random variables having the Cauchy law, with μ and γ being scale and location parameters, respectively.

The Cauchy distribution has been employed on several contexts; e.g., to model experiments with circular [31, 32], finance [33], and optical [34] data. In recent years, some papers have proposed to extend the Cauchy model. Some of them are the generalized odd half-Cauchy family proposed by Cordeiro *et al.* [35], the beta-Cauchy distribution proposed by Alshawarbeh *et al.* [36], the generalized skew-Cauchy distribution proposed by Huang and Chen *et al.* [37], the Weibull-power Cauchy distribution proposed by Tahir *et al.* [38], the generalized Cauchy family proposed by Alzaatreh *et al.* [39], the half-Cauchy distribution proposed by Bosch [40], the Kumaraswamy-half Cauchy distribution proposed by Ghosh [41] and the beta-half-Cauchy distribution proposed by Cordeiro and Lemonte [42].

One of the strategies to generate probability distributions is by "compounding", pioneered by Grushka [43] and Golubev [44]. This procedure proposes to generate a new model from a scale transform of a random vector. Some examples are:

- ☞ Minimum or maximum of a random sample;
- ☞ Sum of a random vector;

☞ Minimum on which one of its parameters is random.

Golubev [44] proposed the exponentially modified Gaussian model which has physic appeal in biology and is the result of the convolution between exponential and Gaussian models. Teich and Diament [45] showed that the compound Poisson distribution such that its parameter follows the range distribution yields in the negative binomial model. These compositions were considered as special cases by Karlis and Xekalaki [30]. They defined the "compounding N" method that describes an event by the sum given in Eq. (4.1). The wide use of (4.1) can be justified by its analytic form be approximate to several natural phenomenon. Revfeim [46] proposed the compound Poisson exponential for describing total precipitation at a day, where the daily precipitation number is Poisson distributed and the precipitation amount follows the exponential distribution. Panger [47] showed that the compound Poisson and negative binomial model are extensively used in the risk economic theory. Besides, Thompson [48] made an application of Compound Poisson distribution for modeling the month rain total.

In this thesis,

Third goal: we advocate using (4.1) to model intensities obtained from SAR systems.

SAR images have been used as important tools for solving remote sensing issues. Each entry of SAR image is associated to a complex element whose squared norm is called as *intensity*. Two questions are important to take account when working with this type of images: (i) they are contaminated by speckle noise which hampers their processing and (ii) some of their scenarios result in multimodal data. Other two contributions are also made in the SAR context:

Fourth goal: Proposal of a new family of distributions which includes extensions for the $\mathcal{G}_I^0$ and $\mathcal{K}$ models and

Fifth goal: Proposal of a $\mathcal{G}_I^0$ regression model.

The above five goals are detailed subsequently.

1.2 Main goals

In general, one wishes to contribute in two applied great areas: Survival analysis and SAR image processing. With respect to the former, we propose non-likelihood estimation criteria as well as GoF measures for the beta-G class. Second we aim to offer a parsimonious model to describe multimodal SAR intensity data, an extension of the family of $\mathcal{G}$ distributions, and a regression model that is tailored for situations where the dependent variable, Y, assumes positive real values following the $\mathcal{G}_I^0$ distribution.

The technical mechanism used is based on (a) the MT as a statistical tool applied to inference, (b) random compound sum, (c) extension of the family of $\mathcal{G}$ distributions, and (d) a regression model for the $\mathcal{G}_I^0$ distribution. This work addresses the following specific objectives:

1. To derive PWMs of Fréchet and Kumaraswamy models for real powers.

2. To derive expressions for the MT of beta-Weibull (BW), beta-Fréchet (BF), beta-Kw (BKw), and beta-log-logistic (BLL) distributions.

3. To propose four Hotelling's T^2 statistics as new GoF measures.

4. To propose a new alternative methodology to the Pearson diagram for some beta-G distributions like qualitative scheme of GoF.

5. To furnish new estimators in closed-form based on LCs for BW parameters.

6. Proposal and assessment of a Hotelling's T^2 statistic at the BW model for both inputs MLEs and LCEs.

7. To propose a new three-parameter model called the compound truncated-Poisson Cauchy distribution (CTPC) for describing multimodal SAR intensities.

8. To derive some properties of the CTPC model: cf and a distance measure between cfs.

9. To provide two estimation procedures for the CTPC parameters: MLEs and quadratic distance (QD) estimators.

10. To design a GoF tool based on a distance between cfs with associated cut-off point.

11. To propose a new probability distribution capable of encompassing two $\mathcal{G}$ submodels, $\mathcal{K}$ and $\mathcal{G}^0$. In particular, they are sought to be extended with the inclusion of an additional parameter. Some mathematical properties of the proposed distributions are studied.

12. To propose a regression model for the $\mathcal{G}_I^0$ distribution and to derive some of its properties, such as, score function, MLEs, Fisher information matrix, and influential measures.

The following five essays summarize Chapters 2, 3, 4, 5, and 6.

“ **Essay 1: Goodness-of-fit Measures Based on the Mellin Transform for Beta Generalized Lifetime Data**”, submitted to `Reliability Engineering & Systems Safety`.

In this essay, we propose GoF measures for generalized beta (beta-G) models based on the MT and PWMs. We combine PWMs and the MT in order to furnish new GoF qualitative and quantitative tools for choosing models within the beta-G class. First, we derive PWMs for the Fréchet and Kumaraswamy distributions. After, we provide expressions for the MT associated with the beta-G family and some of its special cases: in particular, BW, BF, BKw, and BLL distributions. Subsequently, we make a combination between the Hotelling's T^2 statistic and the multivariate delta method to beget asymptotic confidence ellipses to test hypotheses involving second kind cumulants. Finally, we apply the proposed GoF measures on five real data sets applications in survival data analysis.

“ **Essay 2: New Mellin-based Estimation and GoF Criteria for the Beta-Weibull Model**”, under review.

In this essay, we propose a new estimation method based on LC expressions of the BW distribution, an important model in the beta-G generator. This process is called LCEs. Further, we furnish a new BW GoF measure as well as its asymptotic behavior. This latter tool combines the Hotelling's T^2 statistic with the generalized delta method. Unlike MLEs, our proposal has both expression and asymptotic standard errors in closed-form. In order to quantify the performance of proposed LCEs and to compare it with that due to MLEs, we perform Monte Carlo experiments. Finally, we apply the proposed method based on LC (LC) expressions with real data.

“ **Essay 3: The Compound Truncated Poisson Cauchy Model: A Descriptor for Multimodal Data**”, under review.

In this essay, we propose a probability distribution having only three parameters, which is able to describe multimodal data. Our model is defined by the sum of a random number, following the truncated Poisson law, of independent random variables with the Cauchy model, called CTPC distribution. We derive some of its properties: cf and a distance measure between cfs. Further, we provide two estimation procedures for the CTPC parameters: MLEs and QD estimators. Furthermore, we derive a new GoF measure stemmed from the CTPC law and based on empirical cf. To quantify the performance of both proposed estimators and GoF statistic, we make a Monte Carlo simulation study. According to three figures of merit, results indicate QDEs may present better performance than MLEs. Finally, an experiment with actual SAR data is performed. It aims to describe a segment of SAR intensities (positive real feature of SAR) with at least two types of textures. Our model can outperform six classic distributions for modeling SAR intensities: Weibull, gamma, generalized gamma, $\mathcal{K}$, $\mathcal{G}_I^0$, and beta generalized normal models.

" **Essay 4:** **An Extension for the Family of $\mathcal{G}$ Distributions**", under review.

In this essay, we extend the $\mathcal{G}$ family, proposing models more extensive than classical distributions from the $\mathcal{G}$ family: such as $\mathcal{G}_I^0$ and $\mathcal{K}$ models. Additionally, the statistical properties such as moment, Mellin-based LCs, and maximum likelihood methods concerning to the new distributions are derived. Finally, applications to actual data provide evidence that the new distributions outperform usual $\mathcal{G}$ models.

" **Essay 5:** $\mathcal{G}_I^0$ **Regression Model for Speckled Data**", under review.

Synthetic aperture radar (SAR) is an efficient and widely used remote sensing tool. However, data extracted from SAR images are contaminated with speckle noise, which precludes their modeling directly. Intensities are important SAR features which have a non-additive nature and various distributions obtained from the multiplicative approach have been proposed to describe them. The $\mathcal{G}_I^0$ model is one of the most successful among them. Several inferential procedures have been proposed for $\mathcal{G}_I^0$ parameters, but–from a literature review we made–there are not works which tackle a regression structure for this model. This paper proposes a $\mathcal{G}_I^0$ regression model in order to consider the influence of other intensities (present in the polarimetric SAR data) in the modeling of intensities due to one particular polarization channel. We derive some theoretical properties for the new model: Fisher information matrix, residual-kind measures, and influential tools. Point and interval estimation procedures via maximum likelihood are proposed and assessed by Monte Carlo experiments. Results from synthetic and real studies point out that the new model may be useful for speckled data.

1.3 Organization of the thesis

This thesis is organized as follows. This chapter presents the introduction. Next, a brief review of theoretical background of the papers developed in this thesis is provided. The five subsequent

chapters present in detail the articles of this thesis. Each essay is self-contained, having its own notation and labels for formulas and tables. Finally, the last chapter approaches main conclusions of this thesis and some proposals for future works.

1.4 Computational support

All results we presented in this thesis were obtained using the R statistical software [49]. We employ MATHEMATICA and WXMAXIMA softwares [1] to check deduced expressions when possible.The text was written in the LaTeX [2] and its references were made with BibTeX. The main packages and functions of the R software used in this thesis are presented in Table 1.1.

Table 1.1: The main packages and functions of R software used in this thesis

Packages	Functions	Maintainer
MASS	ginv( )	Brian Ripley
moments	skewness( ), kurtosis( )	Lukasz Komsta
stats	quantile( ), optim( ), ks.test( )	R Core Team
xtable	xtable( )	David Scott
maxLik	maxNR( ), maxBFGS( ), maxBFGSR( ), maxSANN( ), maxCG( ), maxNM( )	Ott Toomet
mixtools	ellipse( )	Derek Young
compiler	enableJIT( )	R Core Team
BB	BBsolve( )	Paul Gilbert
gamlss	gamlss( )	Mikis Stas.
plot3D	scatter3D( )	Karline Soetaert
tools	pdf( ), dev.off( ), compactPDF( )	R Core Team
base	summary( ), var( ), cov( ), t( ), c( ), seq( ), log( ), exp( ), cor( ), length( ), tan( ), atan( ), sin( ), cos( ), abs( ), sqrt( ), round( ), ceiling( ), head( ), tail( ), matrix( ), as.matrix( ), vector( ), as.vector( ), psigamma( ), numeric( ), sort( ), rbind( ), cbind( ), list( ), table( ), read.table( ), diag( ), cloud( ), qnorm( ), rnorm( ), pnorm( ), qchisq( ), pf( ), runif( ), rpois( ), rcauchy( ), pweibull( ), rweibull( ), dweibull( ), pbeta( ), hist( ), plot( ), lines( ), curve( ), points( ), density( ), abline( ), text( ), mtext( ), legend( ), box( ), expression( ), polygon( ), axis( ), title( ), range( ), write( ), anyNA( ), proc.time( ), cut( ), is.infinite( ), return( ), print( ), apply( ), replicate( ), factorial( ), choose( ), cumsum( ), qqplot( ), parse( ), chull( ), solve( ), RNGkind( ), seet.sed( ), eval( ), rep( ), getwd( ), setwd( ), source( ), match.arg( ), rm(list=ls(all=TRUE)), require( )	R Core Team

[1] For more information about these software, see <http://andrejv.github.io/wxmaxima/> and <https://www.wolframalpha.com/>, respectively.

[2] For more information and details on the typography system LaTeX, visit <http://www.tex.ac.uk/CTAN/latex>.

2 Goodness-of-fit Measures Based on the Mellin Transform for Beta Generalized Lifetime Data

Abstract

In recent years various probability models have been proposed for describing lifetime data. Increasing model flexibility is often sought as a means to better describe asymmetric and heavy tail distributions. Such extensions were pioneered by the he beta-G family. However, efficient goodness-of-fit (GoF) measures for the beta-G distributions are sought. In this chapter, we combine probability weighted moments (PWMs) and the Mellin transform (MT) in order to furnish new qualitative and quantitative GoF tools for model selection within the beta-G class. We derive PWMs for the Fréchet and Kumaraswamy distributions; and we provide expressions for the MT, and for the log-cumulants (LC) of the beta-Weibull, beta-Fréchet, beta-Kumaraswamy, and beta-log-logistic distributions. Subsequently, we construct LC diagrams and, based on the Hotelling's T^2 statistic, we derive confidence ellipses for the LCs. Finally, the proposed GoF measures are applied on five real data sets in order to demonstrate their applicability.

Keywords: Class beta-G. Mellin transform. Second-kind statistic. Probability weighted moments. Hotelling's T^2 statistic.

2.1 Introduction

Survival analysis tools have been applied in several contexts, such as survival time of mechanical components [1], the failure times of electrical insulator films [4], the effect of varying IL-2 concentration on T cell response [50], and in censored data from head-and-neck-cancer clinical trials [6]. Further applications were found in digital image processing, for instance in SAR

imagery analysis [51, 52]. The derivation of new probability models capable of better explaining reliability data is a central task in the field of survival analysis. In recent years, an effort to extend classical models by means of probability distribution generators has been sought [53]. As a result, the following probability models were introduced: the Marshall and Olking (MG)-G class of distributions [7], the generalized exponential distribution [54], the beta-normal distribution [53], the Kumaraswamy (Kw)-G class [8], the McDonald normal distribution [9], the generalized gamma distribution [55, 56], the T-X family of distributions [11, 57], and the generalized Weibull distribution [12].

Despite the significant number of new models, few goodness-of-fit (GoF) measures have been proposed for the recent distributions; thus hindering model selection. Considering progressive type-II censored data, Pakyari and Balakrishnan [13] proposed a general GoF test [14] that encompasses the GoF test described in [15]. Such methods are based on distance measures between theoretical and empirical cumulative distribution functions.

Taking a different approach, Linhart and Zucchini [16] proposed information-theoretical measures considering the Akaike and Bayesian information criteria [17] as figures of merit for model selection. An alternative method for GoF assessment was given by the Pearson system diagram for model selection [18], which is based on skewness and kurtosis measures [19, p. 23]. Delignon *et al.* [20] and Vogel and Fennessey [21] applied such diagram for SAR and hydrology data, respectively. Chabert and Tourneret [58] introduced a generalization of the Pearson diagram for bivariate random vectors. Nagahara [59] examined the problem of devising GoF measures for multivariate non-normal distributions by using the Pearson system.

However, in [22], Nicolas noticed that the Pearson diagram tends not to be well-suited for positive random variables. Thus, the LC diagram, which plots the third-kind LC $\widetilde{\kappa}_3$ against the second-kind LC $\widetilde{\kappa}_2$, was introduced as an alternative [22]. The LC diagram offers some advantages over the Pearson diagram. Besides being suitable for positive random variables, its computational implementation is more direct and it also captures the distribution flexibility in the sense of skewness and kurtosis [23]. Such diagram was demonstrated to be relevant for quantitative and qualitative comparison of non-nested distributions in SAR and PolSAR data [27, 28, 60, 61]. A detailed description of the LC diagram is provided in [22, 23]. In [22], Nicolas proposed the application of the

MT [24, p. 50] as an alternative to the usual characteristic function. Li *et al.* [23] also considered the MT-based diagram for the classification of empirical probability density functions (pdf) from SAR imagery data. Nicolas and Maruani [62] compared the MT-based method with the second kind cumulant, moment, lower order moment, and maximum likelihood (ML) methods. Khan and Guida [61] have applied the MT to describe complex vector data having the $\mathcal{G}$ model; whereas Anfinsen and Eltoft [26] have demonstrated that MT can be useful for PolSAR data analysis.

In this chapter, we propose a combination of PWMs and the MT in order to furnish new GoF qualitative and quantitative tools for model selection in classes of generalized distributions [53]. We introduce a general expression for the MT of the beta generalized (beta-G) distributions. Because of analytical tractability and suitability for beta-generalization, we separate the following baseline distributions for investigation: Weibull [63], Fréchet [64, 65], Kumaraswamy [66], and log-logistic [67]. Their corresponding beta-generalizations are: the beta-Weibull (BW) [68], the beta-Fréchet (BF) [69], the beta-Kumaraswamy (BKw) [70], and the beta-log-logistic (BLL) [71] distributions. We introduce closed-form expressions for the Fréchet and Kumaraswamy PWM functions. Moreover, we propose a relationship between the Hotelling's T^2 statistic and the multivariate delta method to obtain asymptotic confidence ellipses for hypothesis testing that involves second kind cumulants. Finally, five actual data sets in the context of survival analysis were submitted to the proposed methodology.

The structure of this work is outlined as follows. Section 2.2 reviews the beta-G class of distributions with four particular cases. In Section 3.3, the MT and its properties are outlined. Moreover, we summarize the PWM theory and derive the PWM for the Fréchet and Kw distributions. Section 2.4 presents new GoF measures for four extended models from the beta-G class. In Section 2.5, numerical results are displayed. Finally, concluding remarks are presented in Section 2.6.

2.2 The Beta-G Distribution Family

The beta-G family of distributions was proposed by Eugene *et al.* [53] and is defined as follows. Let $G(x;\boldsymbol{\tau})$ be a cumulative distribution function (cdf) with parameter vector $\boldsymbol{\tau}$. We refer to such cdf as the baseline distribution. The beta-G approach extends the baseline distribution into another

distribution $F(x)$ according to:

$$F(x) = F(x;a,b,\boldsymbol{\tau}) = I_{G(x;\boldsymbol{\tau})}(a,b) = \frac{1}{B(a,b)} \int_0^{G(x;\boldsymbol{\tau})} \omega^{a-1}(1-\omega)^{b-1}\,\mathrm{d}\omega, \tag{2.1}$$

where $a>0$ and $b>0$ are shape parameters, $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio, $B_y(a,b) = \int_0^y \omega^{a-1}(1-\omega)^{b-1}\,\mathrm{d}\omega$ is the incomplete beta function, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, and $\Gamma(a) = \int_0^\infty \omega^{a-1}\,\mathrm{e}^{-\omega}\,\mathrm{d}\omega$ is the gamma function. The pdf associated with (2.1) is given by:

$$f(x) = f(x;a,b,\boldsymbol{\tau}) = \frac{1}{B(a,b)}\, g(x;\boldsymbol{\tau}) G(x;\boldsymbol{\tau})^{a-1}[1-G(x;\boldsymbol{\tau})]^{b-1}, \tag{2.2}$$

and

$$h(x) = h(x;a,b,\boldsymbol{\tau}) = \frac{g(x;\boldsymbol{\tau})\, G(x;\boldsymbol{\tau})^{a-1}\,[1-G(x;\boldsymbol{\tau})]^{b-1}}{B\,(a,b)\,[1-I_{G(x;\boldsymbol{\tau})}(a,b)]}, \tag{2.3}$$

where $g(x;\boldsymbol{\tau}) = \mathrm{d}\,G(x;\boldsymbol{\tau})/\,\mathrm{d}\,x$ is the baseline pdf. In the next subsections, we separate four particular beta-G distribution, given in Table 2.1, for further assessment and derivation of GoF measures. The quantile function (qf), cdf, and the sample space $\mathcal{X}$ of the baseline distributions in Table 2.1 and models other are addressed in Table 2.2.

Fig. 2.1 presents pdf and hrf curves of BW, BF, BKw, and BLL for several parameters values. Due to the inclusion of shape parameters (a and b), these distributions are more flexible than their baselines and are candidates for modeling positive real data sets [72–74].

2.3 Mellin Transform as a Special PWM: Second Kind Statistics for Beta Generalized Models

The Fourier transform is a central tool in signal analysis [96, 97]. Traditionally a probability distribution can be described by means of its characteristic function (cf) of the first kind, which is the Fourier transform of its pdf. Let X be random variable equipped with cdf $F(x)$. Then, its cf $\Phi_X(t)$ is defined as [98, p. 342]:

$$\Phi_X(t) = \mathrm{E}(\mathrm{e}^{\mathrm{i}tX}) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}F(x), \quad t\in\mathbb{R},$$

[3] For $\alpha \le 0$ we have $\mu \le x < \infty$ and for $\alpha > 0$, $\mu \le x \le \mu + \lambda/\alpha$.

Table 2.1: The pdf and hrf of the addressed models

Models	$f_X(x)$	$h_X(x)$	Reference
BW	$\left(\frac{\alpha}{x}\right)\frac{\upsilon(x)}{B(a,b)}\mathrm{e}^{-b\upsilon(x)}\left[1-\mathrm{e}^{-\upsilon(x)}\right]^{a-1}$	$\left(\frac{\alpha}{x}\right)\frac{\upsilon(x)\mathrm{e}^{-b\upsilon(x)}}{B_{1-\mathrm{e}^{-\upsilon(x)}}(a,b)}\left[1-\mathrm{e}^{-\upsilon(x)}\right]^{a-1}$	[68, 75]
BF	$\left(\frac{\alpha}{x}\right)\frac{\mathrm{e}^{-a\upsilon(x)^{-1}}\left[1-\mathrm{e}^{-\upsilon(x)^{-1}}\right]^{b-1}}{\upsilon(x)B(a,b)}$	$\left(\frac{\alpha}{x}\right)\frac{\mathrm{e}^{-a\upsilon(x)^{-1}}}{\upsilon(x)B_{\mathrm{e}^{-\upsilon(x)^{-1}}}(a,b)}\left[1-\mathrm{e}^{-\upsilon(x)^{-1}}\right]^{b-1}$	[69]
BKW	$\frac{\alpha\lambda x^{\alpha-1}}{B(a,b)}(1-x^{\alpha})^{\lambda b-1}\left[1-(1-x^{\alpha})^{\lambda}\right]^{a-1}$	$\frac{\alpha\lambda x^{\alpha-1}(1-x^{\alpha})^{\lambda b-1}}{B_{1-(1-x^{\alpha})^{\lambda}}(a,b)}\left[1-(1-x^{\alpha})^{\lambda}\right]^{a-1}$	[70]
BLL	$\frac{(\lambda/\alpha)}{B(a,b)}\left(\frac{x}{\alpha}\right)^{\alpha\lambda-1}\left[1+(x/\alpha)^{\lambda}\right]^{-(a+b)}$	$\frac{(\lambda/\alpha)}{B_{x^{\lambda}/(\alpha^{\lambda}+x^{\lambda})}(a,b)}\left(\frac{x}{\alpha}\right)^{\alpha\lambda-1}\left[1+(x/\alpha)^{\lambda}\right]^{-(a+b)}$	[71]

where $\upsilon(x)=\left(\frac{x}{\lambda}\right)^{\alpha}$.

Table 2.2: Quantile and Cumulative Distributions of Adapted PWM

Models	$Q[F(x)]$	$F(x)$	$\mathcal{X}$	References
Weibull	$\lambda[-\log(1-F(x))]^{1/\alpha}$	$1-\exp\left\{-\left(\frac{x}{\lambda}\right)^{\alpha}\right\}$	$\mathbb{R}_+$	[76–78]
Fréchet	$\lambda[-\log(F(x))]^{-1/\alpha}$	$\exp\left\{-\left(\frac{x}{\lambda}\right)^{-\alpha}\right\}$	$\mathbb{R}_+$	––
Kw	$\left[1-(1-F(x))^{1/\lambda}\right]^{1/\alpha}$	$\left[1-(1-x^{\alpha})^{\lambda}\right]$	$[0,\ 1]$	––
Rayleigh	$\{2\sigma^2[-\log(1-F(x))]\}^{1/2}$	$1-\exp\left\{-\frac{x^2}{2\sigma^2}\right\}$	$\mathbb{R}_+$	––
Gumbel	$\mu-\lambda\log[-\log(F(x))]$	$\exp\left\{-\exp\left[-\frac{(x-\mu)}{\lambda}\right]\right\}$	$\mathbb{R}$	[76, 79, 80]
Logistic	$\mu+\lambda[\log(F(x))-\log(1-F(x))]$	$\left[1+\exp\left(-\frac{(x-\mu)}{\lambda}\right)\right]^{-1}$	$\mathbb{R}$	[76]
Log-logistic	$\alpha\left[\frac{F(x)}{1-F(x)}\right]^{1/\lambda}$	$\left[1+\left(\frac{\alpha}{x}\right)^{\lambda}\right]^{-1}$	$\mathbb{R}_+$	[81–84]
GEV ($\alpha\neq 0$)	$\mu+\frac{\lambda}{\alpha}\{[-\log(F(x))]^{-\alpha}-1\}$	$\exp\left[-\left(1+\alpha\frac{(x-\mu)}{\lambda}\right)^{-1/\alpha}\right]$	$\mathbb{R}$	[79, 85–89]
GEV ($\alpha=0$)	$\mu-\lambda\{\log[-\log(F(x))]\}$	$\exp\left[-\exp\left(-\frac{(x-\mu)}{\lambda}\right)\right]$	$\mathbb{R}$	[79, 85–89]
Pareto	$\frac{\lambda}{(1-F(x))^{1/\alpha}}$	$1-\left(\frac{\lambda}{x}\right)^{\alpha}$	$[\lambda,\infty)$	[90–92]
Pareto-G	$\mu+\frac{\lambda}{\alpha}[1-(1-F(x))^{\alpha}]$	$1-\left(1-\alpha\frac{(x-\mu)}{\lambda}\right)^{1/\alpha}$	$(\alpha\neq 0)^3$	[90, 92–94]
Pareto-G	$\mu+\lambda[-\log(1-F(x))]$	$1-\exp\left(-\frac{(x-\mu)}{\lambda}\right)$	$(\alpha=0)^3$	[90, 92–94]
Uniform	$\alpha+F(x)(\beta-\alpha)$	$\frac{x-\alpha}{\beta-\alpha}$	$[\alpha,\beta]$	––
Lambda-G	$\lambda_1+\frac{F(x)^{\lambda_3}-(1-)^{\lambda_4}}{\lambda_4}$	––	$\mathbb{R}$	[76, 95]

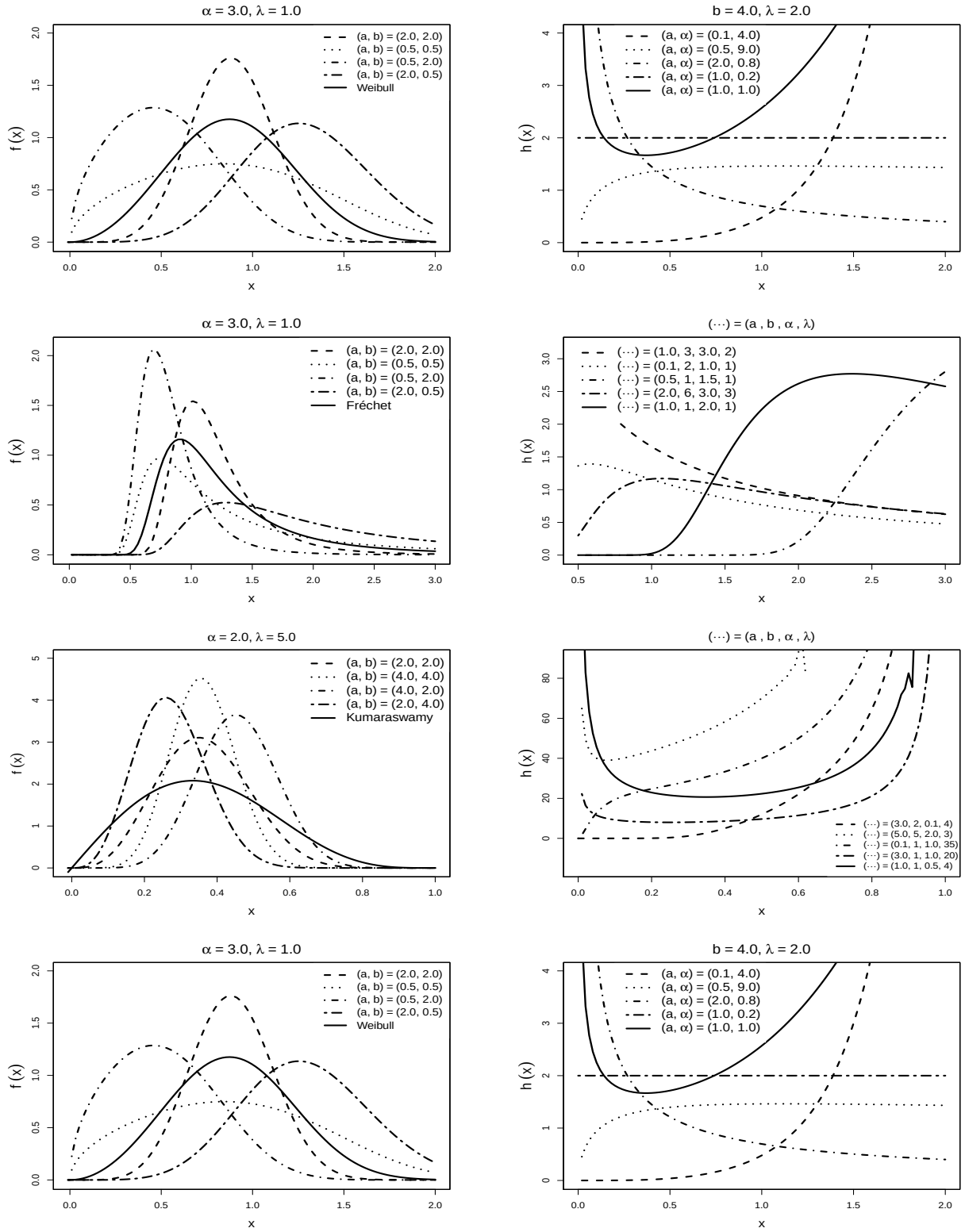


Figure 2.1: The pdf and hrf curves of the BW, BF, BKw, and BLL models for several parameter values.

where $\mathrm{i} = \sqrt{-1}$. However, the cf can be not analytically tractable, as noticed in the BW [68], BF [69], BKw [70], BLL [71], beta-Gumbel [69], and beta log-normal [99] models. To address this issue, Colombo [100] has suggested the MT as an alternative. In [22], Nicolas introduced the second kind statistics based on the MT for analyzing distributions over $\mathbb{R}_+$.

In this section, we show that the MT of beta-G distributions can be directly obtained from the PWM of baseline distributions. For such, in the following, we review PWM for baseline distributions.

2.3.1 *PWM Background*

PWM was introduced by Greenwood *et al.* [76] and consists of a generalized moment expression for probability models. In terms of estimation theory, PWMs can furnish useful closed-form estimators when classical estimators, such as the method of moments and ML, are analytically intractable [76, 79, 101].

The PWM is defined by

$$\mathcal{M}_{l,j,k} = \mathrm{E}\left\{X^l F(X)^j [1-F(X)]^k\right\} = \int_0^1 Q[F(x)]^l F(x)^j (1-F(x))^k \, \mathrm{d}F(x), \tag{2.4}$$

where $l, j, k \in \mathbb{R}$ and $Q(\bullet)$ represents the qf of $F(\bullet)$. Notice that (2.4) generalizes the usual moments, which are obtained by taking $l \in \mathbb{Z}_+$ and $j = k = 0$ ($\mathcal{M}_{l,0,0}$). If $\mathcal{M}_{l,0,0}$ is finite, then PWM $\mathcal{M}_{l,j,k}$ is well-defined for all $j, k \in \mathbb{R}_+$ [76].

Suppose that $j, k \in \mathbb{Z}_+$, from the binomial theorem, we have:

$$\mathcal{M}_{l,0,k} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j \mathcal{M}_{l,j,0} \tag{2.5}$$

and

$$\mathcal{M}_{l,j,0} = \sum_{k=0}^{j} \binom{j}{k} (-1)^k \mathcal{M}_{l,0,k}, \tag{2.6}$$

where if the expression (2.5) exists and X is a continuous variable, thus $\mathcal{M}_{l,0,k}$ exist. The same follows for the expression (2.6). PWM is proportional to $\mathrm{E}(X^l_{j+1:\,j+k+1})$ when l, j, $k \in \mathbb{Z}_+$, since the lth moment about the original of the $(j+1)$th order statistic for a sample size of $(j+k+1)$ is

given by

$$\mathrm{E}\left(X^{l}_{j+1:\,j+k+1}\right) = \frac{\mathcal{M}_{l,\,j,k}}{B(j+1,\,k+1)},$$

where $B(\bullet,\bullet)$ denotes beta function. Other important properties for PWM are [76]:

(i) Let X be a random variable, then lth ordinary moment of the first-order statistic is given by

$$\mathrm{E}(X^{l}_{1:n}) = \begin{cases} (k+1)\mathcal{M}_{l,0,k}, & \text{if} \quad j=0; \\ (j+1)\mathcal{M}_{l,\,j,0}, & \text{if} \quad k=0; \end{cases}$$

(ii) Consider a $n$points random sample drawn from X such that $n = k+1 = j+1$, hence the expected value of the sample range can be determined as

$$\mathrm{E}(X_{n:n} - X_{1:n}) = n\,(\mathcal{M}_{1,n-1,0} - \mathcal{M}_{1,0,n-1}).$$

Table 2.3 presents some PWMs for the distributions listed in Table 2.2.

The PWM can be (i) used to estimate the parameters of a probability model, and (ii) applied as the basis to generate new estimators. In what follows we describe three methods (named M1, M2, and M3) which generators new estimation methods.

Deng and Pandey [102, 103] and Deng *et al.* [104] developed PWM-based estimators by means of the following quantities: (i) partial PWMs (PPWMs), (ii) integer-order PWMs (IPWMs), (iii) fractional PWMs (FPWMs), (iv) fractional partial PWMs (FPPWMs), and (v) integer-order partial PWMs (IPPWMs). The diagram in Fig. 2.2 exhibits the existing estimation PWM-based methods.

Initially, PWM was particularized as follows:

$$\begin{aligned} &\textbf{K1:} \quad \alpha_k = \mathcal{M}_{1,0,k} = \mathrm{E}\{X(1-F)^k\}, \qquad k \in \mathbb{R}_+, \\ &\textbf{K2:} \quad \beta_j = \mathcal{M}_{1,\,j,0} = \mathrm{E}\{X\,F^j\}, \qquad j \in \mathbb{R}_+. \end{aligned} \tag{2.7}$$

[4] Where the constant γ is the Euler-Mascheroni constant.

[5] Summation is 0 when $k = 0$.

Table 2.3: Central Moments and Probability Weighted Moments

Models	$\mathcal{M}_{1,0,0}$	$\mathcal{M}_{l,j,k}$	Restriction
Weibull	$\lambda\Gamma\left(1+\frac{1}{\alpha}\right)$	$\lambda^l\Gamma\left(1+\frac{l}{\alpha}\right)\sum_{r=0}^{\infty}\binom{j}{r}\frac{(-1)^r}{(k+r+1)^{1+l/\alpha}}$	$l,k\in\mathbb{R}$, and $j\in\mathbb{R}_+$
Fréchet	$\lambda\Gamma\left(1-\frac{1}{\alpha}\right)$	$\lambda^l\Gamma\left(1-\frac{l}{\alpha}\right)\sum_{r=0}^{\infty}\binom{k}{r}(-1)^r\frac{1}{(j+r+1)^{1-l/\alpha}}$	$l,j\in\mathbb{R}$ and $k\in\mathbb{R}_+$
Kw	$\lambda B(1+1/\alpha,\,\lambda)$	$\lambda\sum_{r=0}^{\infty}\binom{j}{r}(-1)^r B[1+l/\alpha,\,\lambda(k+r+1)]$	$l,k\in\mathbb{R}$ and $j\in\mathbb{R}_+$
Rayleigh	$\sqrt{2\sigma^2}\Gamma\left(1+\frac{1}{2}\right)$	$\left(2\sigma^2\right)^{l/2}\Gamma\left(1+\frac{l}{2}\right)\sum_{r=0}^{\infty}\binom{j}{r}\frac{(-1)^r}{(k+r+1)^{1+l/2}}$	$l,k\in\mathbb{R}$, and $j\in\mathbb{R}_+$
Gumbel	$\mu+\gamma\lambda$	$\frac{1}{j+1}\{\mu+\lambda[\gamma+\log(j+1)]\}$ [4]	$j\in\mathbb{R}_+, l=1$, and $k=0$
Logistic	μ	$\frac{\mu}{k+1}-\frac{\alpha}{k+1}\sum_{r=0}^{k}\frac{1}{r}$ [5]	$k\in\mathbb{R}_+, l=1$, and $j=0$
Log-logistic	$\frac{\alpha}{\lambda}\left[\frac{\pi}{\sin(\pi/\lambda)}\right]$	$\alpha^l B(j+l/\lambda+1,\,k-l/\lambda+1)$	$j>-(l/\lambda+1)$, and $k>l/\lambda-1$
GEV	$\boldsymbol{\mu}^*$	$\frac{1}{j+1}\left\{\mu-\frac{\lambda}{\alpha}\left[1-(j+1)^{\alpha}\Gamma(1-\alpha)\right]\right\}$	$j\in\mathbb{Z}_+, l=1$, and $k=0$
Pareto	$\frac{\lambda}{1-\alpha},\ \alpha>1$	$\lambda^l B(k+1-\alpha,\,j+1)$	$l\in\mathbb{R}, k>\alpha-1$, and $j>-1$
Pareto-G	$\frac{\lambda}{1+\alpha},\quad \alpha>-1$	$\frac{\lambda}{\alpha}\left[B(j+1,k+1)-B(j+1,k+\alpha+1)\right]$	$j>-1, k>-1$, and $l=1$
Uniform	$\frac{\alpha+\beta}{2}$	$\sum_{r=0}^{l}\binom{l}{r}\alpha^{l-r}\frac{(\beta-\alpha)^r}{(r+j+1)}$	$j\in\mathbb{R}, l\in\mathbb{Z}_+$, and $k=0$
Lambda-G	$\lambda_1+\frac{1}{\lambda_2}\zeta$	$\sum_{r=0}^{l}\binom{l}{r}\lambda_1^{l-r}\lambda_2^{-r}\sum_{h=0}^{r}(-1)^h\binom{r}{h}B(\delta_r,\eta_h)$	$l,j,k\in\mathbb{R}$

where $\boldsymbol{\mu}^* = \begin{cases} \mu+\lambda\left[\frac{\Gamma(1-\alpha)-1}{\alpha}\right], & \text{if} \quad \alpha\neq 0 \quad \text{and} \quad \alpha<1, \\ \mu+\lambda\gamma, & \text{if} \quad \alpha=0, \\ \infty, & \text{if} \quad \alpha\geq 1, \end{cases}$ [4]

$\delta_r=\lambda_3(r-h)+j$, $\eta_h=\lambda_4 h+k+1$, and $\zeta=\left[\frac{1}{\lambda_3+1}-B(1,\lambda_4+1)\right]$.

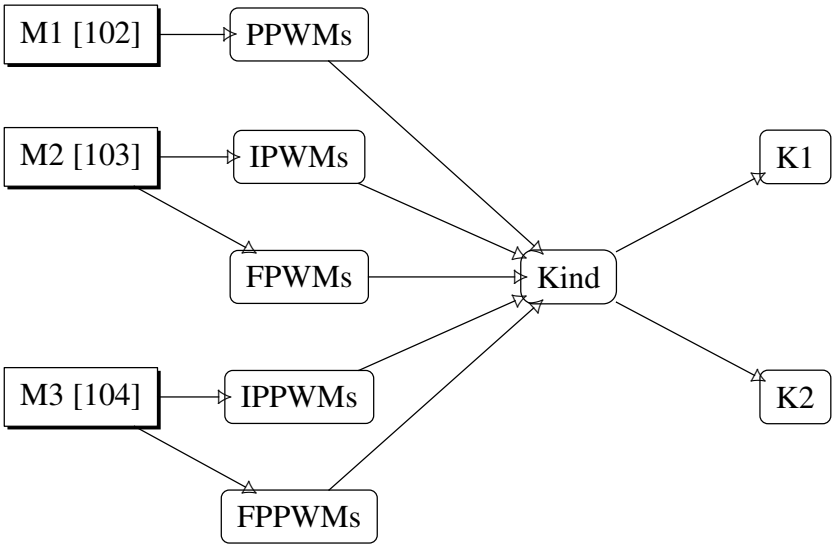


Figure 2.2: Diagram with the PWM-based estimation methods.

Given an ordered sample $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$, the K1 and K2 PWMs can be unbiasedly estimated as

$$a_k = \frac{1}{n}\sum_{r=1}^{n}\left[\binom{n-r}{k}x_r\right] \Big/ \binom{n-1}{k} \qquad \text{and} \qquad b_j = \frac{1}{n}\sum_{r=1}^{n}\left[\binom{n-r}{j}x_r\right] \Big/ \binom{n-1}{j}, \tag{2.8}$$

where $k, j = 0, 1, 2, \ldots, (n-2), (n-1) \in \mathbb{Z}_+$, and the binomial coefficient is given as $\binom{n}{r} = \frac{n!}{r!(n-r)!}$, if $n \geq r \geq 0$, otherwise 0.

The method M1 developed by Deng and Pandey [102], called PPWM, redefines the PWM as

$$\mathcal{M}^p_{l,j,k} = \int_{F(x_0)}^{1} Q[F(x)]^l F(x)^j (1-F(x))^k \, \mathrm{d}F(x), \tag{2.9}$$

where $F(x_0)$ is a lower bound of the censored sample and x_0 is the censoring threshold. Thus,

$$\begin{aligned} \textbf{K1:} \quad & \alpha^p_k = \mathcal{M}^p_{1,0,k} = \int_{F(x_0)}^{1} Q[F(x)](1-F(x))^k \, \mathrm{d}F(x), \\ \textbf{K2:} \quad & \beta^p_j = \mathcal{M}^p_{1,j,0} = \int_{F(x_0)}^{1} Q[F(x)]F(x)^j \, \mathrm{d}F(x). \end{aligned} \tag{2.10}$$

Let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ be an ordered complete sample, the K1 and K2 PPWMs can be unbiasedly estimated by [105]

$$a^p_k = \frac{1}{n}\sum_{r=1}^{n}\left[\binom{n-r}{k}x^p_r\right] \Big/ \binom{n-1}{k} \qquad \text{and} \qquad b^p_j = \frac{1}{n}\sum_{r=1}^{n}\left[\binom{n-i}{j}x^p_r\right] \Big/ \binom{n-1}{j}, \tag{2.11}$$

where $k, j = 0, 1, \ldots, (n-1) \in \mathbb{Z}_+$, and

$$x_r^p = \begin{cases} 0, & \text{if} \quad x_r \leq x_0, \\ x_r, & \text{if} \quad x_r > x_0. \end{cases} \tag{2.12}$$

For a fixed x_0, an empirical frequency estimated of $F(x_0)$ is given by $\widehat{F}(x_0) = n_0/n$, where n_0 is the number of occurrences which do not exceed x_0 in the sample.

We highlight two important observations [102], namely: (i) If $F(x_0) = 0$, α_k^p and β_j^p become α_k and β_j, respectively and (ii) let α_k^p and β_j^p be unbiased estimators, then α_k^p and β_j^p become α_k and β_j, respectively.

On the other hand, Deng and Pandey [103] developed other methods, denoting by IPWMs and FPWMs. The IPWMs method consists in the usage of the expressions (2.7) and (2.8) to generate new estimators. In case of the FPWMs method, we maintain the expressions (2.7)

$$a_k = \frac{1}{n}\sum_{r=1}^{n}[(1-P_r)^k x_r] \quad \text{and} \quad b_j = \frac{1}{n}\sum_{r=1}^{n}[P_r^j x_r],$$

where x_r is the rth order statistics, $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$, and P_r is the probability plotting position of x_r computed by a suitable formula [106], in which

$$P_r = \frac{r-0.35}{n}. \tag{2.13}$$

Finally, we have the method M3 that was developed by [104], denoted FPPWM. This method is the combination of the two previous methods, M1 and M2. The first step is to define the FPPWM as given in (2.9). From there, they designated the K1 and K2 for FPPWMs and IPPWMs as given in (2.10) and (2.11), respectively; and the x_r^p presented in (2.12). Hence, the K1 and K2 FPPWM are estimated from a sample series using plotting position expression

$$a_k^p = \frac{1}{n}\sum_{r=1}^{n}[(1-P_r)^k x_r^p] \qquad \text{and} \qquad b_j^p = \frac{1}{n}\sum_{r=1}^{n}[P_r^j x_r^p],$$

where x_r^p is defined in Eq. (2.12) and P_r in Eq. (2.13).

2.3.2 *PWM of Particular Baseline Distributions*

Greenwood *et al.* [76] and Caiza and Ummenhofer [77] derived the PWM for the following models: the Weibull, the Gumbel [76], the generalized lambda [76], the logistic [76, 77], the

Wakeby [76], and the kappa distribution [76]. Mahdi and Ashkar [83] derived the PWM for the log-logistic model as a means to investigate the generalized probability weighted moments and ML fitting methods. They also showed how to provide an estimation based on PWMs. Mahdi and Ashkar [78, 81] also derived and described PWMs linked to the Weibull and log-logistic models as an alternative to estimation methods such as the generalized PWMs [90], generalized moments [107–109], and ML methods [110].

However, the current literature lacks PWM expressions for the Fréchet and Kw distributions. In the following propositions, we address this literature gap.

Proposition 1 *Let X be a random variable following the Fréchet model with $\lambda > 0$ and $\alpha > 0$ as location and shape parameters, respectively. The PWM of X is given by*

$$\mathcal{M}_{l,j,k} = \lambda^l \Gamma\left(1 - \frac{l}{\alpha}\right) \sum_{r=0}^{\infty} \binom{k}{r} (-1)^r \frac{1}{(j+r+1)^{1-\frac{l}{\alpha}}}.$$

Proposition 2 *Let X be a random variable following the Kw model with shape parameters $\lambda > 0$ and $\alpha > 0$. The PWM of X is given by*

$$\mathcal{M}_{l,j,k} = \lambda \sum_{r=0}^{\infty} \binom{j}{r} (-1)^r B\left[1 + \frac{l}{\alpha}, \lambda(k+r+1)\right].$$

Proofs for the above propositions are given in the Appendix A. A summary of the PWM results is listed in Table 2.3; central moments are also shown as particular cases.

2.3.3 *Mellin Transform*

Let $X \in \mathbb{R}_+$ be a random variable with cdf $F(x)$. Then the first cf of the second kind is defined by means of the MT:

$$\phi_X(s) = \int_0^{\infty} x^{s-1} \, \mathrm{d}F(x) = \mathrm{E}(X^{s-1}), \tag{2.14}$$

where $s \in \mathbb{C}$ is a complex variable [22].

Similarly the cf, the MT has also an inverse. Let $\phi_X(s)$ be well-defined, one can deduce the pdf as [22, 100]

$$f(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} x^{-s} \, \phi_X(s) \, \mathrm{d}s. \tag{2.15}$$

If $f(x)$ is a pdf, the second kind cf satisfies $\phi_X(s)\Big|_{s=1} = 1$.

Next, we have the second cf of the second kind (an alternative to the cgf), which is defined as the natural logarithm of $\phi_X(s)$ [22]:

$$\psi_X(s) = \log[\phi_X(s)]. \tag{2.16}$$

As an illustration, consider the following example [111]: Let $X \in \mathbb{R}_+$ be a random variable following the Gamma distribution with pdf given by

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x}, \quad \text{for} \quad x, \alpha, \beta > 0,$$

where $\Gamma(\alpha)$ is the gamma function [112, p. 254]. From (2.14), one has

$$\phi_X(s) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{(s+\alpha-1)-1} \mathrm{e}^{-\beta x} \mathrm{d}(x).$$

After simple algebraic manipulations,

$$\phi_X(s) = \frac{\Gamma(s+\alpha-1)}{\Gamma(\alpha)} \frac{1}{\beta^{s-1}}.$$

For $\alpha = \beta = 1$ and from the inversion formula (2.15), one has

$$f(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} x^{-s} \Gamma(s)\, \mathrm{d}s = \mathrm{e}^{-x}.$$

The last integral is known as the Cahen-Mellin integral [113]. Assuming $\alpha > 0$ and $\beta > 0$,

$$\begin{aligned} f(x) &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{x^{-s}}{\beta^{s-1}} \frac{\Gamma(s+\alpha-1)}{\Gamma(\alpha)}\, \mathrm{d}s \\ &= \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} (x\beta)^{-(s+\alpha-1)} \Gamma(s+\alpha-1)\, \mathrm{d}s. \end{aligned}$$

Using the Cahen-Mellin integral formula [113],

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x}.$$

Considering the beta-G family, we introduce the following theorem relating the MT to the PWMs.

Theorem 1 *Let X be a random variable having distribution in the beta-G family with cdf and pdf given by* (2.1) *and* (2.2)*, respectively. Then, the MT of X, referred to as $\phi_{BG}(s)$, is given by*

$$\phi_{BG}(s) = \frac{1}{B(a,b)}\,\mathcal{M}_{s-1,a-1,b-1}, \tag{2.17}$$

where $\mathcal{M}_{s-1,a-1,b-1}$ is the PWM of a baseline G.

Proof 1 *Applying* (2.1) *into* (2.14) *with $\overline{G}(x) = 1 - G(x)$ we can show that:*

$$\begin{aligned}
\phi_{BG}(s) &= \int_0^\infty x^{s-1}\,\mathrm{d}F(x) \\
&= \int_0^\infty x^{s-1}\,\frac{g(x)}{B(a,b)}\,G(x)^{a-1}\overline{G}(x)^{b-1}\,\mathrm{d}x \\
&= \frac{1}{B(a,b)}\int_0^\infty x^{s-1}\,G(x)^{a-1}\overline{G}(x)^{b-1}\,\mathrm{d}G(x) \\
&= \frac{1}{B(a,b)}\,\mathrm{E}\left[X^{s-1}\,G(X)^{a-1}\overline{G}(X)^{b-1}\right] \\
&= \frac{1}{B(a,b)}\,\mathcal{M}_{s-1,a-1,b-1}.
\end{aligned}$$

■

Table 2.4 displays the obtained MT for the considered distributions.

Table 2.4: Mellin transform for the BW, BF, BKw, and BLL models

Models	MT	Restriction
BW	$\frac{\lambda^{s-1}}{B(1,b)}\,\Gamma\left(1+\frac{s-1}{\alpha}\right)b^{-(s-1+\alpha)/\alpha}$	for all $a=1, b>0, \alpha>0, \lambda>0$
BF	$\frac{\lambda^{s-1}}{B(a,1)}\,\Gamma\left(1-\frac{s-1}{\alpha}\right)a^{(s-1-\alpha)/\alpha}$	for all $b=1, a>0, \alpha>0, \lambda>0$
BKw	$\frac{\lambda}{B(1,b)}\,B\left(1+\frac{s-1}{\alpha}, \lambda b\right)$	for all $a=1, b>0, \alpha>0, \lambda>0$
BLL	$\frac{\alpha^{s-1}}{B(a,1)}\,B\left(a+\frac{s-1}{\lambda}, 1-\frac{s-1}{\lambda}\right)$	for all $b=1, a>0, \alpha>0, \lambda>0$

2.3.4 *The log-cumulants method*

In what follows, we discuss that log-moments (LMs) arise from the MT likewise ordinary moments come from the cf [22, 114, 115]. Hence, the νth LMs, say $\widetilde{m}_\nu$, is defined by [22]

$$\widetilde{m}_\nu = \left.\frac{\mathrm{d}^\nu \phi_X(s)}{\mathrm{d}s^\nu}\right|_{s=1} = \int_{\mathbb{R}_+} (\log x)^\nu \, \mathrm{d}F(x) = \mathrm{E}[(\log X)^\nu], \quad \forall\, \nu \in \mathbb{N}.$$

Deriving $\psi_X(s)$ and evaluating $s = 1$, we obtain the second kind LCs of order ν [22]:

$$\widetilde{\kappa}_\nu = \left.\frac{\mathrm{d}^\nu \, \psi_X(s)}{\mathrm{d}s^\nu}\right|_{s=1}, \qquad \forall\, \nu \in \mathbb{N}. \tag{2.18}$$

For example, the three first LCs can be written as:

$$\begin{cases} \widetilde{\kappa}_1 = \widetilde{m}_1, \\ \widetilde{\kappa}_2 = \widetilde{m}_2 - \widetilde{m}_1^2, \\ \widetilde{\kappa}_3 = \widetilde{m}_3 - 3\widetilde{m}_1\widetilde{m}_2 + 2\widetilde{m}_1^3. \end{cases}$$

The formulas LMs and LCs are also used for the second-kind moments and cumulants, for instance [116, 117]: Let X be a random variable,

$$\widetilde{\kappa}_1 = \widetilde{m}_1 = \mathrm{E}(\log X) \quad \text{and} \quad \widetilde{\kappa}_2 = \widetilde{m}_2 - \widetilde{m}_1^2 = \mathrm{Var}(\log X).$$

In general, the vth-order LCs can be retrieved from [26, 60]

$$\widetilde{\kappa}_\nu = \widetilde{m}_\nu - \sum_{r=1}^{\nu-1} \binom{\nu-1}{r-1} \widetilde{\kappa}_r \, \widetilde{m}_{\nu-r}. \tag{2.19}$$

In practice, let $\widetilde{\kappa}_\nu$ be a function of the parameter vector $\boldsymbol{\theta}$, then the estimation of $\boldsymbol{\theta}$ is done by replacing $\widetilde{m}_\nu$ by sample LM, say $\widehat{\widetilde{m}}_\nu$, which is defined as [26, 28, 60, 118, 119]

$$\widehat{\widetilde{m}}_\nu = \frac{1}{n}\sum_{r=1}^{n} (\log x_r)^\nu,$$

where n is the sample size and x_r the rth observation of a data set. In [22, 119, 120] the sample LM was defined as

$$\widehat{\widetilde{m}}_{\nu_1} = \frac{1}{n}\sum_{r=1}^{n} \log x_r^{\nu_1}.$$

In addition to these estimation processes, the literature also uses mixed moments and mixed LM given by [119]

$$\widehat{\widetilde{m}}_{\nu_2} = \frac{1}{n}\sum_{r=1}^{n} x_r^{\nu_2+1}\log(x_r),$$

and

$$\widehat{\widetilde{m}}_{\nu_3} = \frac{1}{n}\sum_{r=1}^{n} x_r\log(x_r^{\nu_3}),$$

respectively. Another interesting case is by considering the identities [116, 117, 121, 122]:

$$\begin{aligned}
\widehat{\widetilde{\kappa}}_1 &= \frac{1}{n}\sum_{r=1}^{n}[\log(x_r)],\\
\widehat{\widetilde{\kappa}}_2 &= \frac{1}{n}\sum_{r=1}^{n}[\log(x_r)-\widehat{\widetilde{\kappa}}_1]^2,\\
\widehat{\widetilde{\kappa}}_3 &= \frac{1}{n}\sum_{r=1}^{n}[\log(x_r)-\widehat{\widetilde{\kappa}}_1]^3,\\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
\widehat{\widetilde{\kappa}}_\nu &= \frac{1}{n}\sum_{r=1}^{n}[\log(x_r)-\widehat{\widetilde{\kappa}}_1]^\nu, \quad \text{where} \quad \nu\in\mathbb{N}.
\end{aligned}$$

It follows some examples of LC-based estimators: $\mathcal{G}_I^0$ [120, 123], $\mathcal{G}_A^0$ [120, 123], gamma [62], $\mathcal{K}$ [62], Nakagami [23], Fisher [23], and generalized gamma [23] models. These resulting estimators depend on the polygamma function $\psi^{(\nu)}(z)$ [112, p. 260]. The use of approximations for $\psi^{(\bullet)}(\bullet)$ may be a manner of obtaining approximated closed-form estimators. To that end, the following results are valid [112, p. 259–260] as $z\to\infty$,

$$\begin{aligned}
\psi(z) &= \log(z)-\frac{1}{2z}-\frac{1}{12z^2}+\frac{1}{120z^4}+\dots\\
\psi^{(1)}(z) &\approx \frac{1}{z}+\frac{1}{2z^2}+\frac{1}{6z^3}-\frac{1}{30z^5}+\frac{1}{42z^7}-\frac{1}{30z^9}+\dots\\
\psi^{(2)}(z) &\approx -\frac{1}{z^2}-\frac{1}{z^3}-\frac{1}{2z^4}+\frac{1}{6z^6}-\frac{1}{6z^8}+\frac{3}{10z^{10}}-\frac{5}{6z^{12}}+\dots\\
\psi^{(3)}(z) &\approx \frac{2}{z^3}+\frac{3}{z^4}+\frac{2}{z^5}-\frac{1}{z^7}+\frac{4}{3z^9}-\frac{3}{z^{11}}+\frac{10}{z^{13}}-\dots
\end{aligned} \tag{2.20}$$

For the generalized gamma [23] model with parameters $\mu\neq 0$, $\delta\in\mathbb{R}_+$, and $\sigma\in\mathbb{R}_+$, the LCs

are given by [23]

$$\widetilde{\kappa}_{\nu} = \begin{cases} \log(\sigma) + \dfrac{\psi(\delta) - \log(\delta)}{\mu}, & \text{for} \quad \nu = 1, \\ \dfrac{\psi^{(\nu-1)}(\delta)}{\mu^{\nu}}, & \text{for} \quad \nu \geq 2. \end{cases} \tag{2.21}$$

Thus, using (2.21) and (2.20), estimators for μ and σ may be given in closed-form [23]:

$$\begin{cases} \widehat{\mu} = \sqrt{\frac{1}{\widehat{\widetilde{\kappa}}_2}\left(\frac{1}{\widehat{\delta}} + \frac{1}{2\widehat{\delta}^2}\right)}, \\ \widehat{\sigma} = \exp\left\{\widehat{\widetilde{\kappa}}_1 - \frac{1}{\widehat{\mu}}\left[\left(\frac{1}{\widehat{\delta}} + \frac{1}{2\widehat{\delta}^2}\right) - \log(\widehat{\delta})\right]\right\}. \end{cases}$$

Table 2.5 lists LC-based estimators for some known distributions. Next, in the Table 2.6, presents LCs for the considered beta-G models.

2.3.5 *The Log-cumulants Diagram*

As discussed by Delignon *et al.* [20], the Pearson diagram is a tool for model selection and assessment of fitting. Such diagram is based on skewness and kurtosis measures. Nicolas [22] presented evidence that the Pearson diagram can be analytically intractable and introduced the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram, which is similar to the Pearson diagram, but employs the second kind statistics $\widetilde{\kappa}_3$ and $\widetilde{\kappa}_2$ instead of skewness and kurtosis measures. In [22], it is also shown that the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram is a suitable alternative for classifying SAR images whose associate Pearson diagram is often intractable.

Anfinsen and Eltoft [60] introduced the matrix LC (MLC) diagram as means to visually inspect the multidimensional space where each dimension is represented by one particular MLC with order ν. Thus, such visualization tool facilitates the use of MT and provides intuition to the LC method. The diagram in [60] is an extension of the LC diagram considered by Nicolas [22, 119] for the univariate MT. In [23, 26, 28, 61], the MT-based LC diagram was employed for pdf classification from SAR imagery data.

In this chapter, we employ the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram as a tool for assessing fits under beta-G models. The LCs of such models were derived using PWMs and the MT and are displayed in Table 2.6,

Table 2.5: PDFs, Mellin transform, and Log-cumulants. Here $K_\alpha(\bullet)$ denotes the modified Bessel function of the third kind

Models	Distribution functions	Mellin transform	Log-cumulants	References
Weibull	$\frac{\alpha}{\lambda^\alpha}x^{\alpha-1}\mathrm{e}^{-\left(\frac{x}{\lambda}\right)^\alpha}$	$\lambda^{s-1}\Gamma\left(1+\frac{s-1}{\alpha}\right)$	$\begin{cases}\widetilde{\kappa}_1=\log(\lambda)+\frac{1}{\alpha}\psi(1)\\ \widetilde{\kappa}_\nu=\frac{1}{\alpha^\nu}\psi^{(\nu-1)}(1), \quad \nu\geq 2\end{cases}$	[23, 27, 119, 124]
Log-normal	$\frac{1}{x\sigma\sqrt{2\pi}}\exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right]$	$\mathrm{e}^{(s-1)\mu+\frac{(s-1)^2\sigma^2}{2}}$	$\widetilde{\kappa}_1=\mu, \quad \widetilde{\kappa}_2=\sigma^2$	[27, 124, 125]
Beta	$\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{B(\alpha+s-1,\beta)}{B(\alpha,\beta)}$	$\begin{cases}\widetilde{\kappa}_1=\psi(\alpha)-\psi(\alpha+\beta)\\ \widetilde{\kappa}_\nu=\psi^{(\nu-1)}(\alpha)-\psi^{(\nu-1)}(\alpha+\beta), \quad \nu\geq 2\end{cases}$	[119]
Inverse beta	$\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1+x)^{-\alpha-\beta}$	$\frac{B(\alpha+s-1,\beta-(s-1))}{B(\alpha,\beta)}$	$\begin{cases}\widetilde{\kappa}_1=\psi(\alpha)-\psi(\beta)\\ \widetilde{\kappa}_\nu=\psi^{(\nu-1)}(\alpha)+(-1)^\nu\psi^{(\nu-1)}(\beta), \quad \nu\geq 2\end{cases}$	[119]
Generalized gamma[6]	$\frac{\lvert\eta\rvert}{\mu}\frac{L^{\frac{1}{\eta}}}{\Gamma(L)}\left(\frac{L^{\frac{1}{\eta}}x}{\mu}\right)^{\eta L-1}\mathrm{e}^{-\left(\frac{L^{\frac{1}{\eta}}x}{\mu}\right)^\eta}$	$\mu^{s-1}\frac{\Gamma(L+\frac{s-1}{\eta})}{L^{\frac{s-1}{\eta}}\Gamma(L)}$	$\begin{cases}\widetilde{\kappa}_1=\log(\mu)+\frac{\psi(L)-\log(L)}{\eta}\\ \widetilde{\kappa}_\nu=\frac{\psi^{(\nu-1)}(L)}{\eta^\nu}, \quad \nu\geq 2\end{cases}$	[119, 124]
$\mathcal{K}_A$	$\frac{4\lambda Lx(\lambda Lx^2)^{\frac{(\alpha+L)}{2}-1}}{\Gamma(\alpha)\Gamma(L)}K_{\alpha-L}(2x\sqrt{\lambda L})$	$(L\lambda)^{-\frac{(s-1)}{2}}\frac{\Gamma\left(L+\frac{s-1}{2}\right)\Gamma\left(\alpha+\frac{s-1}{2}\right)}{\Gamma(L)\Gamma(\alpha)}$	$\begin{cases}2\widetilde{\kappa}_1=-\log(L\lambda)+\psi(L)+\psi(\alpha)\\ 2^\nu\widetilde{\kappa}_\nu=\psi^{(\nu-1)}(L)+\psi^{(\nu-1)}(\alpha), \quad \nu\geq 2\end{cases}$	[122–124]
$\mathcal{G}_A^0$	$\frac{2L^L\Gamma(L-\alpha)\gamma^{-\alpha}x^{2L-1}}{\Gamma(L)\Gamma(-\alpha)(\gamma+Lx^2)^{L-\alpha}}$	$\left(\frac{\gamma}{L}\right)^{\frac{(s-1)}{2}}\frac{\Gamma\left(L+\frac{s-1}{2}\right)\Gamma\left(-\alpha-\frac{s-1}{2}\right)}{\Gamma(L)\Gamma(-\alpha)}$	$\begin{cases}2\widetilde{\kappa}_1=\log(\gamma/L)+\psi(L)-\psi(-\alpha)\\ 2^\nu\widetilde{\kappa}_\nu=\psi^{(\nu-1)}(L)+(-1)^{\nu-1}\psi^{(\nu-1)}(-\alpha), \quad \nu\geq 2\end{cases}$	[120, 122–124]
$\mathcal{G}_I^0$	$\frac{L^L\Gamma(L-\alpha)\gamma^{-\alpha}x^{L-1}}{\Gamma(L)\Gamma(-\alpha)(\gamma+Lx)^{L-\alpha}}$	$\left(\frac{\gamma}{L}\right)^{s-1}\frac{\Gamma(-\alpha-s+1)\Gamma(L+s-1)}{\Gamma(L)\Gamma(-\alpha)}$	$\begin{cases}\widetilde{\kappa}_1=\log(\gamma/L)+\psi(L)-\psi(-\alpha)\\ \widetilde{\kappa}_\nu=\psi^{(\nu-1)}(L)+(-1)^{\nu-1}\psi^{(\nu-1)}(-\alpha), \quad \nu\geq 2\end{cases}$	[120, 122, 123, 126]

[f6] We can check that:

1. for $\eta=1$ we find the gamma distribution,
2. for $\eta=2$ we find the Rayleigh-Nakagami distribution,
3. for $\eta=-1$ we find the inverse gamma distribution,
4. for $\eta=-2$ we find the inverse Rayleigh-Nakagami distribution,
5. for $L\eta=1$ we find the generalized normal distribution.

Table 2.6: Log-cumulants of considered models

Model	$\widetilde{\kappa}_1$	$\widetilde{\kappa}_2$	$\widetilde{\kappa}_3$	$\cdots$	$\widetilde{\kappa}_\nu \quad \forall\, \nu > 1$
BW	$\log(\lambda) + \frac{\psi(1)-\log(b)}{\alpha}$	$\frac{1}{\alpha^2}\psi^{(1)}(1)$	$\frac{1}{\alpha^3}\psi^{(2)}(1)$	$\cdots$	$\frac{1}{\alpha^\nu}\psi^{(\nu-1)}(1)$
BF	$\log(\lambda) - \frac{\psi(1)-\log(a)}{\alpha}$	$\frac{1}{\alpha^2}\psi^{(1)}(1)$	$-\frac{1}{\alpha^3}\psi^{(2)}(1)$	$\cdots$	$(-1)^\nu \frac{1}{\alpha^\nu}\psi^{(\nu-1)}(1)$
BKw	$\frac{\psi(1)-\psi(\lambda b+1)}{\alpha}$	$\frac{\psi^{(1)}(1)-\psi^{(1)}(\lambda b+1)}{\alpha^2}$	$\frac{\psi^{(2)}(1)-\psi^{(2)}(\lambda b+1)}{\alpha^3}$	$\cdots$	$\frac{\psi^{(\nu-1)}(1)-\psi^{(\nu-1)}(\lambda b+1)}{\alpha^\nu}$
BLL	$\log(\alpha) + \frac{\psi(a)-\psi(1)}{\lambda}$	$\frac{\psi^{(1)}(a)+\psi^{(1)}(1)}{\lambda^2}$	$\frac{\psi^{(2)}(a)-\psi^{(2)}(1)}{\lambda^3}$	$\cdots$	$\frac{\psi^{(\nu-1)}(a)+(-1)^\nu\,\psi^{(\nu-1)}(1)}{\lambda^\nu}$

where the LCs are given in terms of the digamma and polygamma functions given by $\psi(z) = \frac{\mathrm{d}}{\mathrm{d}z}\log\Gamma(z)$ and $\psi^{(n)}(z) = \frac{\mathrm{d}^{n+1}}{\mathrm{d}z^{n+1}}\log\Gamma(z)$ [112, p. 258–260], respectively.

Fig. 2.3 exhibits the regions in the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram linked to the BW, BF, BKw, and BLL models. These regions can be understood as manifolds [60]. Each distribution is represented by a subspace, whose dimensions depend on the parameter number of the associated distribution [26, 60]. However, the resulting region can degenerate into a curve [26, 60]. For instance, the log-logistic distribution has no shape parameters and its manifold is represented by a line (vertical dashed line), which can be viewed as a zero-dimensional manifold.

The regions linked to the BW and BF distributions are parameterized by one parameter. These regions are represented, respectively, by a solid and dotted curve in Fig. 2.3, being one-dimensional manifolds. On the other hand, the BKw and BLL distributions result in two-dimensional manifold because they are parametrized by two and three parameters, respectively.

2.4 New GoF Tools for Beta-G Models

In recent years, several models have been proposed to describe survival data, such as the beta power exponential [127], McDonald exponentiated gamma [3], gamma extended Weibull [2], and the models considered in this chapter. These models are however in need of accurate GoF tools. In this section, we propose four GoF tools for beta-G models based on the Hotelling's T^2 statistic [128, p. 170].

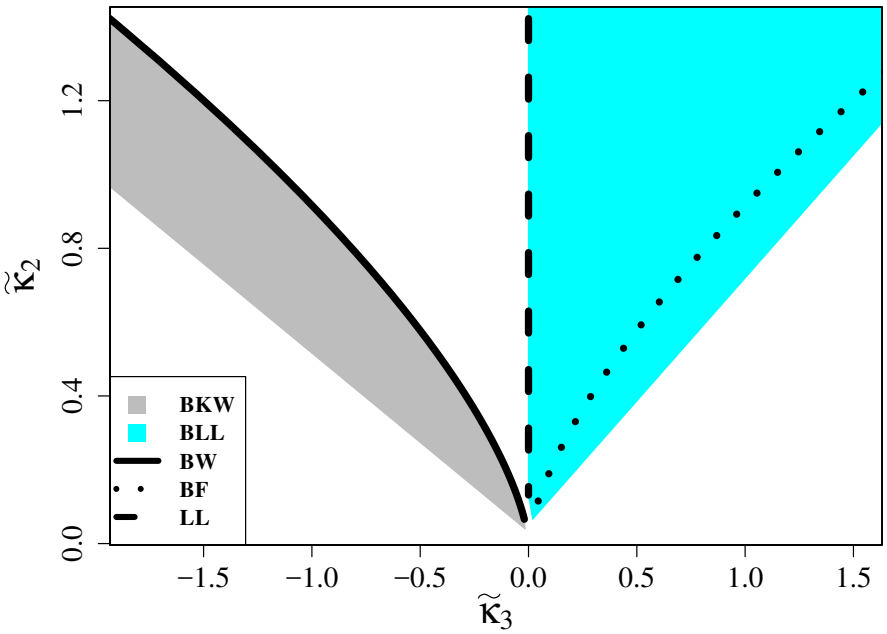


Figure 2.3: Diagram of the LCs $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ for BW, BF, BKw, and BLL models.

2.4.1 *Hotelling's T^2 Statistic*

The Hotelling's T^2 statistic is a generalization of the Student's t statistics [128, p. 170] given by:

$$T^2 = n\,(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{S}^{-1}\,(\overline{\boldsymbol{x}} - \boldsymbol{\mu})\,, \tag{2.22}$$

where $\overline{\boldsymbol{x}} = \frac{1}{n}\sum_{r=1}^{n} \boldsymbol{x}_r$ is the sample mean vector based on a random sample $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ from the ν-variate normal random vector $\boldsymbol{x} \sim \mathcal{N}_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix, respectively; and $\boldsymbol{S} = \frac{1}{n}\sum_{r=1}^{n} \boldsymbol{x}_r \boldsymbol{x}_r^\top - n\overline{\boldsymbol{x}}\,\overline{\boldsymbol{x}}^\top$ is the sample covariance matrix. Such statistics follows the F-Snedecor distribution with ν and $n-\nu$ degrees of freedom denoted by $F_{\nu,n-\nu}$ [128, p. 177].

Considering a significance level η, the likelihood ratio test for the hypothesis $\mathrm{E}(\overline{X}) = \boldsymbol{\mu}$ can be rejected if $T^2 \geq Q_F(1-\eta;\nu,n-\nu)$, where $Q_F(\bullet;\nu,n-\nu)$ is the qf for $F_{\nu,n-\nu}$ [128]. Additionally one may consider ν-dimensional confidence ellipsoids given by [128]:

$$n\,(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{S}^{-1}\,(\overline{\boldsymbol{x}} - \boldsymbol{\mu}) \leq Q_F(1-\eta;\nu,n-\nu).$$

For large samples, the T^2 distribution can be approximated by its limiting distribution, which is the chi-squared distribution with ν degrees of freedom [26]. This result is relevant for the case where $\boldsymbol{x}_r$ is not normal and the exact distribution for (2.22) is not known.

2.4.2 *Hotelling's T^2 statistic and Log-cumulants*

We aim at applying the Hotelling's T^2 statistic as a means for proposing GoF tests based on the LCs. Our goal is to estimate the LCs and then classify the underlying distribution according to the location of the estimated LCs $\left[\widehat{\widetilde{\kappa}}_2 \quad \widehat{\widetilde{\kappa}}_3\right]^\top$ over the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram.

Therefore, we need a test statistics for the null hypothesis $\mathcal{H}_0 : \mathrm{E}\left(\left[\widehat{\widetilde{\kappa}}_2 \quad \widehat{\widetilde{\kappa}}_3\right]\right) = \left[\widetilde{\kappa}_2 \quad \widetilde{\kappa}_3\right]$. Such test would pave the way for accepting or rejecting the pertinence of estimated LCs to particular regions over the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram.

Because the LCs tend to be analytically well-defined quantities, they can be given closed-form expressions, as we showed in Table 2.6, for several beta-G distributions. Such relationship between parameters and LCs can be used to derive estimators for the LCs. In other words, we have that

$$\widehat{\widetilde{\kappa}}_2 = g_2(\widehat{\boldsymbol{\theta}}) \quad \text{and} \quad \widehat{\widetilde{\kappa}}_3 = g_3(\widehat{\boldsymbol{\theta}}),$$

where $\widehat{\boldsymbol{\theta}}$ is the estimated parameter vector; and $g_2(\bullet)$ and $g_3(\bullet)$ are composite functions that return the LCs in terms of the baseline distribution parameters by means of evaluating: (2.18), (2.16), (2.14), and (2.17).

Further, we notice that for large samples, considering the generalized delta method [115], the estimator vector $\left[\widehat{\widetilde{\kappa}}_2 \quad \widehat{\widetilde{\kappa}}_3\right]^\top$ follows the bivariate normal distribution with mean $\left[\widetilde{\kappa}_2 \quad \widetilde{\kappa}_3\right]^\top$ and an asymptotic covariance matrix $\boldsymbol{K}$, as previously shown in [26, 28, 61]. The estimated asymptotic covariance matrix $\widehat{\boldsymbol{K}}$ can be obtained from the asymptotic covariance matrix of the parameter estimators $\boldsymbol{\Sigma}$ [26]. As shown in [26, p. 2769], we can write

$$\widehat{\boldsymbol{K}} = \widehat{\boldsymbol{J}}^\top \cdot \widehat{\boldsymbol{\Sigma}} \cdot \widehat{\boldsymbol{J}},$$

where

$$\widehat{\boldsymbol{J}} = \begin{bmatrix} \frac{\mathrm{d}\, g_2(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_1} & \frac{\mathrm{d}\, g_3(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_1} \\ \frac{\mathrm{d}\, g_2(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_2} & \frac{\mathrm{d}\, g_3(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_2} \\ \vdots & \vdots \\ \frac{\mathrm{d}\, g_2(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_r} & \frac{\mathrm{d}\, g_3(\widehat{\boldsymbol{\theta}})}{\mathrm{d}\,\widehat{\theta}_r} \end{bmatrix}, \tag{2.23}$$

θ_r is the rth parameter of the model and $\widehat{\boldsymbol{\Sigma}}$ is the estimated asymptotic covariance matrix of the estimator vector $\widehat{\boldsymbol{\theta}}$.

The ML estimators for the parameters of beta-G distributions often have no closed-form expressions and its covariance matrix is unknown. This is illustrated by the beta-Pareto [72], beta-Laplace [129], BW [68], and BF [130] distributions.

As presented in [110, p. 181–182] and [25,131], the inverse of Fisher information matrix (FIM) can be employed as an approximation for the variance-covariance matrix, since it is the asymptotic covariance matrix of the ML estimators [110, p. 181–182]. Thus, we have the following approximation for $\boldsymbol{\Sigma}$ [132, Th. 7.3.11]:

$$\boldsymbol{\Sigma} \approx -\left[\mathrm{E}\left(\frac{\mathrm{d}^2\, \ell(\boldsymbol{\theta})}{\mathrm{d}\,\boldsymbol{\theta}^\top\, \mathrm{d}\,\boldsymbol{\theta}}\right)\right]^{-1}, \tag{2.24}$$

under regularity conditions [110].

For the majority of beta-G models, the calculation of their FIMs is analytically intractable. A common solution for this problem is the use of the observed information matrix instead of the FIM, as supported by [72, 129, 133]. Thus, the observed information matrix is an estimator for the FIM [134]. Therefore, in this work, we use the inverse observed information matrix as replacement for the asymptotic covariance matrix of the ML estimators. The observed information matrix has the advantage of being definite positive matrix; thus measuring the observed curvature on the log-likelihood surface. In other words, it provides an indication of how much a multidimensional likelihood surface is rotated with respect to the parameter axes [135]. For the model parameter estimation, we employed the ML estimation because it results in invariant, consistent, and asymptotically efficient estimators [136, p. 3].

Thus, comparing with (2.22) and (2.23), for second and third order LCs $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$, we have that $\nu = 2$, $\boldsymbol{\mu} = \begin{bmatrix} \widetilde{\kappa}_2 & \widetilde{\kappa}_3 \end{bmatrix}^\top$, $\bar{\boldsymbol{x}} = \begin{bmatrix} \widehat{\widetilde{\kappa}}_2 & \widehat{\widetilde{\kappa}}_3 \end{bmatrix}^\top$, where $\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$ are sample estimators for $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$, respectively. The matrix $\mathbf{S}$ can be substituted by the estimated asymptotic covariance matrix $\widehat{\boldsymbol{K}}$ of $[\widehat{\widetilde{\kappa}}_2, \widehat{\widetilde{\kappa}}_3]^\top$. Therefore, we obtain the following statistic:

$$T^2 = n \left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right)^\top \widehat{\boldsymbol{K}}^{-1} \left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right), \tag{2.25}$$

where the inverse of $\widehat{\boldsymbol{K}}$ is obtained via usual matrix inversion [137, 138] if the matrix $\widehat{\boldsymbol{K}}$ is nonsingular [139, p. 508]; otherwise the generalized Moore-Penrose inverse [139, 140] is applied.

For such, we submit the estimated LCs $\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$ from the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram [22] to the Hotelling's T^2 statistic formalism. Considering large samples, the limiting distribution of the random variable T^2 is the χ^2 distribution [26]. Thus, in (2.25), we can adopt the approximation $Q_F(\bullet; \nu, n-\nu) \approx Q_{\chi^2}(\bullet; \nu)$, where $Q_{\chi^2}(\bullet; \nu)$ is the qf for the χ^2 distribution with ν degrees of freedom. Therefore, we can derive a confidence ellipse at significance level η according to:

$$\left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right)^\top \widehat{\boldsymbol{K}}^{-1} \left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right) \leq \frac{1}{n} Q_{\chi^2}(\eta; 2),$$

where $Q_{\chi^2}(\eta; 2)$ is the qf for χ^2_2. The above ellipse is centered at $(\widetilde{\kappa}_2, \widetilde{\kappa}_3)$ and its axes are directed according the eigenvectors of $\widehat{\boldsymbol{K}}$ [141].

2.4.3 *Hotelling's T^2 statistic for Selected Beta-G Distributions*

Based on the last discussion, four GoF measures are proposed for the BW, BF, BKw, and BLL distributions.

Proposition 3 *Let X be a random variable following the BW distribution with parameters $a = 1$, $b > 0$, $\alpha > 0$ and $\lambda > 0$, then the Hotelling's T^2 statistic, here referred to as T^2_{BW}, based on the LCs is given by*

$$T^2_{BW} = \frac{n\widehat{\alpha}^6}{4} \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right)^2 \left(\frac{|\widehat{\boldsymbol{H}}_{BW}|}{U_{\widehat{\alpha}\widehat{\alpha}} U_{\widehat{b}\widehat{b}} - U^2_{\widehat{\lambda}\widehat{b}}} \right),$$

where $|\bullet|$ *is the determinant of a matrix* $\widehat{\boldsymbol{H}}_{\cdot}$*;* $|\widehat{\boldsymbol{H}}_{BW}|$ *is the estimator of* $|\boldsymbol{H}_{BW}|$ *given by*

$$|\boldsymbol{H}_{BW}| = U_{\alpha\alpha}(U_{\lambda\lambda}U_{bb} - U_{\lambda b}^2) + U_{\alpha\lambda}(U_{\alpha b}U_{\lambda b} - U_{\alpha\lambda}U_{bb}) + U_{\alpha b}(U_{\alpha\lambda}U_{\lambda b} - U_{\alpha b}U_{\lambda\lambda}),$$

and $\widehat{\alpha}$*,* $\widehat{b}$*,* $\widehat{\lambda}$ *are the estimators for* α*,* b*, and* λ*, respectively. The quantities* $U_{\cdot}$ *are the entries of the associated information matrix and are given in the Appendix B.*

Proposition 4 *Let* X *be a random variable following the BF distribution with parameters* $a > 0$*,* $b = 1$*,* $\lambda > 0$ *and* $\alpha > 0$*, then the Hotelling's* T^2 *statistic, referred to as* T_{BF}^2*, based on the LCs is given by*

$$T_{BF}^2 = \frac{n\widehat{\alpha}^6}{4}\left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2}\right)^2 \left(\frac{|\widehat{\boldsymbol{H}}_{BF}|}{U_{\widehat{\alpha}\widehat{\alpha}}U_{\widehat{a}\widehat{a}} - U_{\widehat{\lambda}\widehat{a}}^2}\right),$$

where $|\widehat{\boldsymbol{H}}_{BF}|$ *is the estimator of* $|\boldsymbol{H}_{BF}|$ *given by*

$$|\boldsymbol{H}_{BF}| = U_{\alpha\alpha}(U_{\lambda\lambda}U_{aa} - U_{\lambda a}^2) + U_{\alpha\lambda}(U_{\alpha a}U_{\lambda a} - U_{\alpha\lambda}U_{aa}) + U_{\alpha a}(U_{\alpha\lambda}U_{\lambda a} - U_{\alpha a}U_{\lambda\lambda}),$$

and $\widehat{\alpha}$*,* $\widehat{a}$*,* $\widehat{\lambda}$ *are the estimates of the* α*,* a*, and* λ*, respectively. In the Appendix B, the quantities* $U_{\cdot}$ *are fully detailed.*

Proposition 5 *Let* X *be a random variable following the BKw with parameters* $a = 1$*,* $b > 0$*,* $\lambda > 0$ *and* $\alpha > 0$*, then its Hotelling's* T^2 *statistic, here called* T_{BKw}^2*, is given by*

$$T_{BKw}^2 = \frac{n|\widehat{\boldsymbol{H}}_{BKw}|}{\widehat{\delta}_{22}\widehat{\delta}_{33} - \widehat{\delta}_{23}^2} \cdot \left[\widehat{\delta}_{33}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)^2 + \widehat{\delta}_{22}\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)^2 - 2\widehat{\delta}_{23}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)\right],$$

where $|\widehat{\boldsymbol{H}}_{BKw}|$ *is an estimator of* $|\boldsymbol{H}_{BKw}|$ *given by*

$$|\boldsymbol{H}_{BKw}| = U_{\alpha\alpha}(U_{\lambda\lambda}U_{bb} - U_{\lambda b}^2) + U_{\alpha\lambda}(U_{\alpha b}U_{\lambda b} - U_{\alpha\lambda}U_{bb}) + U_{\alpha b}(U_{\alpha\lambda}U_{\lambda b} - U_{\alpha b}U_{\lambda\lambda});$$

$\widehat{\widetilde{\kappa}}_2$ *and* $\widehat{\widetilde{\kappa}}_3$ *are the estimates of the LCs* $\widetilde{\kappa}_2$ *and* $\widetilde{\kappa}_3$*, respectively;* $\widehat{\delta}_{22}$*,* $\widehat{\delta}_{23}$*, and* $\widehat{\delta}_{33}$ *are the estimates for* δ_{22}*,* δ_{23}*, and* δ_{33} *given in compact form by*

$$\begin{aligned}
\delta_{22} &= \begin{bmatrix} J_{12} & J_{22} & J_{32} \end{bmatrix} \cdot \boldsymbol{\Sigma}_{BKw} \cdot \begin{bmatrix} J_{12} & J_{22} & J_{32} \end{bmatrix}^\top, \\
\delta_{23} &= \begin{bmatrix} J_{12} & J_{22} & J_{32} \end{bmatrix} \cdot \boldsymbol{\Sigma}_{BKw} \cdot \begin{bmatrix} J_{13} & J_{23} & J_{33} \end{bmatrix}^\top, \\
\delta_{33} &= \begin{bmatrix} J_{13} & J_{23} & J_{33} \end{bmatrix} \cdot \boldsymbol{\Sigma}_{BKw} \cdot \begin{bmatrix} J_{13} & J_{23} & J_{33} \end{bmatrix}^\top;
\end{aligned}$$

and

$$
\begin{aligned}
J_{12} &= \frac{2}{\alpha^3}\left\{\psi^{(1)}(\lambda b+1)-\psi^{(1)}(1)\right\},\\
J_{13} &= \frac{3}{\alpha^4}\left\{\psi^{(2)}(\lambda b+1)-\psi^{(2)}(1)\right\},\\
J_{22} &= \frac{b}{\alpha^2}\psi^{(2)}(\lambda b+1), \quad J_{23} = \frac{b}{\alpha^3}\psi^{(3)}(\lambda b+1),\\
J_{32} &= \frac{\lambda}{\alpha^2}\psi^{(2)}(\lambda b+1), \quad J_{33} = \frac{\lambda}{\alpha^3}\psi^{(3)}(\lambda b+1).
\end{aligned}
$$

For the sake of brevity, the matrix $\boldsymbol{\Sigma}_{BKw}$ is shown in Appendix B where the above expressions are also given fully expanded forms.

Proposition 6 *Let X be a random variable following the BLL distribution with parameters $a>0$, $b=1$, $\lambda>0$ and $\alpha>0$, then its Hotelling's T^2 statistic, here named T^2_{BLL}, based on the LCs is given by*

$$
T^2_{BLL} = \frac{n|\boldsymbol{H}_{BLL}|}{\widehat{\delta}_{22}\widehat{\delta}_{33}-\widehat{\delta}^2_{23}}\cdot\left[\widehat{\delta}_{33}\left(\widehat{\widetilde{\kappa}}_2-\widetilde{\kappa}_2\right)^2+\widehat{\delta}_{22}\left(\widehat{\widetilde{\kappa}}_3-\widetilde{\kappa}_3\right)^2-2\widehat{\delta}_{23}\left(\widehat{\widetilde{\kappa}}_2-\widetilde{\kappa}_2\right)\left(\widehat{\widetilde{\kappa}}_3-\widetilde{\kappa}_3\right)\right],
$$

where $|\widehat{\boldsymbol{H}}_{BLL}|$ is an estimator of $|\boldsymbol{H}_{BLL}|$ given by

$$
|\boldsymbol{H}_{BLL}| = U_{\alpha\alpha}(U_{\lambda\lambda}U_{aa}-U^2_{\lambda a})+U_{\alpha\lambda}(U_{\alpha a}U_{\lambda a}-U_{\alpha\lambda}U_{aa})+U_{\alpha a}(U_{\alpha\lambda}U_{\lambda a}-U_{\alpha a}U_{\lambda\lambda});
$$

$\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$ are the estimates of the LCs $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$, respectively; $\widehat{\delta}_{22}$, $\widehat{\delta}_{23}$, and $\widehat{\delta}_{33}$ are the estimates for δ_{22}, δ_{23}, and δ_{33} given in compact form by

$$
\begin{aligned}
\delta_{22} &= \begin{bmatrix}0 & J_{22} & J_{32}\end{bmatrix}\cdot\boldsymbol{\Sigma}_{BLL}\cdot\begin{bmatrix}0 & J_{22} & J_{32}\end{bmatrix}^\top,\\
\delta_{23} &= \begin{bmatrix}0 & J_{22} & J_{32}\end{bmatrix}\cdot\boldsymbol{\Sigma}_{BLL}\cdot\begin{bmatrix}0 & J_{23} & J_{33}\end{bmatrix}^\top,\\
\delta_{33} &= \begin{bmatrix}0 & J_{23} & J_{33}\end{bmatrix}\cdot\boldsymbol{\Sigma}_{BLL}\cdot\begin{bmatrix}0 & J_{23} & J_{33}\end{bmatrix}^\top;
\end{aligned}
$$

and

$$
\begin{aligned}
J_{22} &= -\frac{2}{\lambda^3}\left\{\psi^{(1)}(a)+\psi^{(1)}(1)\right\},\\
J_{23} &= -\frac{3}{\lambda^4}\left\{\psi^{(2)}(a)-\psi^{(2)}(1)\right\},\\
J_{32} &= \frac{1}{\lambda^2}\psi^{(2)}(a), \quad J_{33} = \frac{1}{\lambda^3}\psi^{(3)}(a).
\end{aligned}
$$

The matrix $\boldsymbol{\Sigma}_{BLL}$ is shown in Appendix B with the above expressions fully expanded.

2.5 Application to Actual Data

2.5.1 *Selected Data Sets and Descriptive Statistics*

We separated five real data sets to be submitted to our proposed methodology, determining according to the introduced GoF criteria a suitable candidate among the BW, BF, BKw, and BLL models. In the following, we describe briefly the selected data sets:

(i) Breaking data [142]: 100 observations on breaking stress of carbon fibres (in Gba);

(ii) Guinea.pig data [143]: 72 survival times of guinea pigs injected with different doses of tubercle bacilli;

(iii) Stress-rupture data [5]: the stress-rupture life of kevlar 49/epoxy strands subjected to constant sustained pressure at the 90% stress level until failure;

(iv) Airborne data [144, 145]: repair times (in hours) for an airborne communication transceiver;

(v) River flow data [146]: lower discharge of at least seven consecutive days and return period (time) of ten years ($Q_{7,10}$) of the Cuiabá River, Mato Grosso, Brazil.

Table 4.1 gives the descriptive summary for each data set. The first and second data sets are homogeneous with sample variation coefficient (VC) of 38.68% and 37.26%, respectively. The remaining data sets are heterogeneous. The river flow data set has negative skeweness and platykurtic distribution (kurtosis is less than 3). The remaining data sets have positive skeweness with leptokurtic distribution (kurtosis is greater than 3).

2.5.2 $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ *Diagram and Log-cumulants Estimation*

Fig. 2.4 exhibits the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram and regions linked to particular distributions are emphasized. For each data set, we computed the sample LCs according to bootstrap sampling with 1,000 replicates and 90% sample sizes. Each data set is represented by a different dot pattern.

Table 2.7: Descriptive statistics and LCs ($\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$) for selected data sets

Data	Min.	Max.	Mean	Median	SD	Skewness	Kurtosis	VC (%)	$\widehat{\widetilde{\kappa}}_2$	$\widehat{\widetilde{\kappa}}_3$
Breaking	0.39	5.56	2.62	2.70	1.01	0.36	3.10	38.68	0.19	−0.09
Guinea.pig	12.00	376.00	99.82	70.00	81.12	1.80	5.61	81.26	0.50	0.04
Stress-rupture	0.01	7.89	1.02	0.80	1.12	3.00	16.71	109.22	2.02	−2.71
Airborne	0.20	24.50	3.64	1.75	5.07	2.91	11.67	139.34	1.18	0.35
River flow	34.14	186.40	107.60	114.10	40.90	−0.15	1.93	37.26	0.19	−0.06

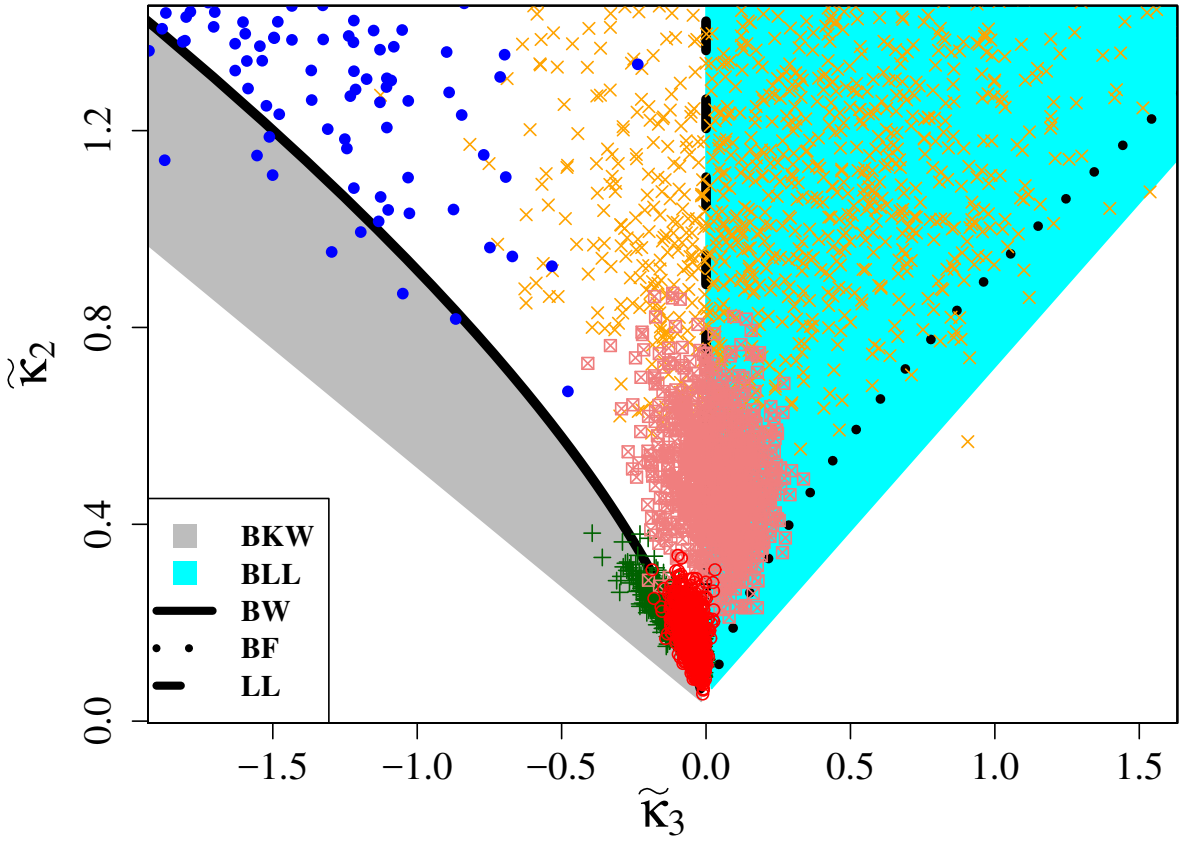


Figure 2.4: Diagram of the LCs $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ showing the manifolds of theoretical LCs for the BKw, BLL, BW, BF, and log-logistic models, as well as a collection of sample LCs representing breaking (+), guinea.pig (⊠), stress-rupture (●), airborne (×), and river flow (○) data sets.

For each data set, we computed the 95% confidence interval in each replications (sample sizes $n \in \{10, 50, 100, 300\}$) and repeated this process 1,000 times to compute these intervals without the Bonferroni correction (WOBC) and with the Bonferroni correction (WBC). Table 2.8 displays centers and ranges of the confidence interval for bootstrap percentile WOBC and WBC. We note that for the five data sets the larger the sample size the smaller confidence interval. Both for the WOBC and WBC methods. Also, we highlight that for all data sets and in all scenarios ($n \in \{10, 50, 100, 300\}$ and the second and third order sample LCs), the method WOBC had a lower confidence interval than WBC.

Qualitatively we have the following analysis. For the breaking stress and stress-rupture data sets, most of the points are located over the regions linked to the BKw and BW distributions. The points derived from the Guinea.pig data are located in the BLL distribution region. The airborne data set has its associated points over the central region, which includes the log-logistic distribution region, and over the BLL and BF distribution regions. Finally, we have the river flow data set, where the BKw distribution region captures most of its points, while some of them spread over the BW and log-logistic distribution regions.

2.5.3 *Hotelling's T^2 statistic Analysis*

The Hotelling's T^2 statistic with p values were computed and are displayed in Table 2.9. The obtained statistics can be interpreted as a measure of the distance between the data and each particular beta-G model. Lower values of Hotelling's T^2 statistic suggest a better agreement between data and model; indicating therefore a better data fitting.

By separating the models linked to the smallest values of Hotelling's T^2 statistic, we have that the BLL distribution is a good model for the guinea.pig and airborne data sets. The stress-rupture and breaking data could be better modeled by the BW distribution. Similarly, the river flow data could be fitted under the BKw distribution. These quantitative results confirm the qualitative analysis provided in Fig. 2.4.

Table 2.8: Centers and ranges of the confidence interval for bootstrap percentile WOBC and WBC

Data	n	Center				Range			
		$\widehat{\widetilde{\kappa}}_2$		$\widehat{\widetilde{\kappa}}_3$		$\widehat{\widetilde{\kappa}}_2$		$\widehat{\widetilde{\kappa}}_3$	
		WOBC	WBC	WOBC	WBC	WOBC	WBC	WOBC	WBC
Breaking	10	0.2540	0.2706	−0.1912	−0.2249	0.3958	0.4545	0.4173	0.5089
	50	0.2058	0.2153	−0.1116	−0.1303	0.1751	0.2153	0.2056	0.2514
	100	0.2005	0.2068	−0.1012	−0.1075	0.1297	0.1531	0.1624	0.1827
	300	0.1980	0.1990	−0.0896	−0.0950	0.0730	0.0881	0.0948	0.1159
Guinea.pig	10	0.4871	0.5261	0.0011	−0.0273	0.6746	0.8103	0.7069	0.8421
	50	0.5058	0.5049	0.0299	0.0275	0.3350	0.3981	0.3068	0.3798
	100	0.5007	0.5036	0.0359	0.0331	0.2326	0.2917	0.2259	0.2637
	300	0.5027	0.5036	0.0411	0.0412	0.1349	0.1660	0.1312	0.1533
Stress-rupture	10	1.9282	2.0325	−2.8507	−3.1065	2.8789	3.3712	6.0492	7.1599
	50	2.0169	2.0158	−2.5566	−2.6439	1.3950	1.6934	2.9275	3.4401
	100	2.0015	2.0187	−2.6613	−2.6314	0.9527	1.1515	2.1224	2.5153
	300	2.0201	2.0181	−2.6782	−2.6759	0.5475	0.6489	1.1837	1.4768
Airborne	10	1.1443	1.1834	0.3735	0.3599	1.4776	1.6600	2.0760	2.6289
	50	1.1758	1.1720	0.3724	0.3786	0.6979	0.7739	1.0802	1.2804
	100	1.1703	1.1741	0.3652	0.3670	0.4798	0.5672	0.7770	0.9387
	300	1.1687	1.1738	0.3551	0.3585	0.2717	0.3209	0.4501	0.5216
River flow	10	0.1797	0.1796	−0.0579	−0.0642	0.2329	0.2681	0.1406	0.1699
	50	0.1868	0.1865	−0.0589	−0.0587	0.1112	0.1260	0.0703	0.0805
	100	0.1855	0.1847	−0.0589	−0.0607	0.0743	0.0876	0.0479	0.0601
	300	0.1854	0.1859	−0.0590	−0.0598	0.0451	0.0538	0.0271	0.0333

Table 2.9: Hotelling's T^2 statistic and p-value (in parentheses) with respect to the data sets

Models	Breaking	Guinea.pig	Stress-rupture	Airborne	River flow
BW	0.1786 (0.8367)	6.7272 (0.0021)	0.0554 (0.9462)	2.9496 (0.0638)	2.6815 (0.0821)
BF	> 10 (≈ 0.00)	5.4663 (0.0062)	> 10 (≈ 0.00)	2.3937 (0.1042)	2.2896 (0.1159)
BKw	6.7720 (0.0018)	> 10 (≈ 0.00)	0.1335 (0.8752)	> 10 (≈ 0.00)	0.2258 (0.7990)
BLL	5.5000 (0.0054)	0.8623 (0.4266)	> 10 (≈ 0.00)	0.3187 (0.7289)	0.9358 (0.4016)

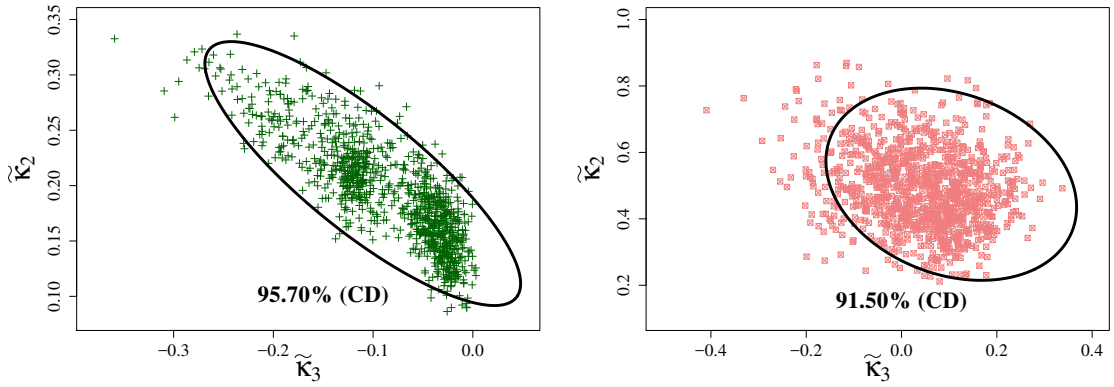
2.5.4 *Confidence Ellipses*

To complement the previous analysis of visual application, we have plotted confidence ellipses for each data set. The construction of the ellipses was based on (2.25). We employed the smallest value of Hotelling's T^2 statistic from Table 2.9 and the associated beta-G distribution using the estimated LCs $\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$; and the sample variance-covariance matrix $\widehat{\Sigma}$.

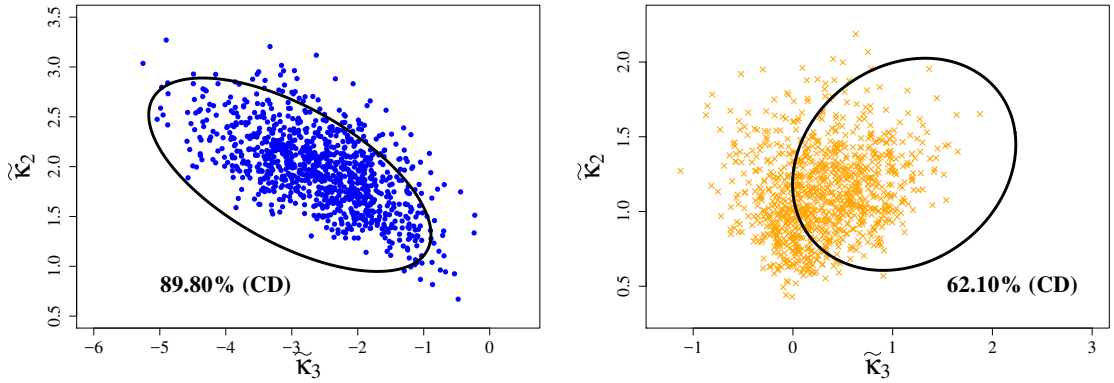
To obtain the ellipse for breaking and stress-rupture data set we used the BW model with the T^2_{BW} statistics given by Proposition 3. For the guinea.pig and airborne data set we apply the BLL model with the T^2_{BLL} statistics given by Proposition 6. In the case of the river flow data, we apply the BKw model with the T^2_{BKw} statistics given by Proposition 5. Fig. 2.5 depicts the obtained ellipses.

2.6 Conclusion

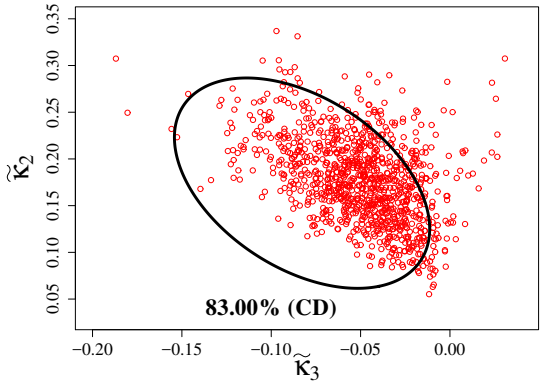
In this chapter, several GoF measures have been proposed for determining good fits at the beta-G class in the survival analysis context. We provided qualitative and quantitative analyses for the introduced GoF tools including numerical and visual inspection approaches. We derived closed-form expressions for the second kind characteristic function, LCs, Hotelling's T^2 statistic, and ellipse of confidence for the LCs of the BW, BF, BKw, and BLL distributions. Proposed measures have been applied to five real data sets in order to demonstrate their applicability.



(a) Confidence ellipses-BW (b) Confidence ellipses-BLL



(c) Confidence ellipses-BW (d) Confidence ellipses-BLL



(e) Confidence ellipses-BKw

Figure 2.5: Confidence ellipses for each data set.

3 New Mellin-based Estimation and GoF Criteria for the Beta-Weibull Model

Abstract

In recent years various probability models have been proposed for describing lifetime data. Such proposals are often made from a class of distributions and aim to gain flexibility to yield asymmetric and heavy tail distributions. The beta-G family proposed by Eugene *et al.* [Beta-normal distribution and its applications. *Communication in Statistics-Theory and Methods*, 31, 497-512] is one of those classes. Although this class as well as other competing proposals can provide models which are able even of characterizing multimodal data, efficient estimation processes for their parameters are mandatory. Works about new distributions often present maximum likelihood estimators (MLEs). In general, although has well-defined asymptotic properties they do not provide closed-form and requires the use of interactive optimization sources. In this chapter, we propose a new estimation method based on log-cumulant (LC) expressions of the beta-Weibull (BW) distribution, an important model in the beta-G generator. This process is called LC estimators. Further, we furnish a new BW goodness-of-fit measure as well as its asymptotic behavior. This latter tool combines the Hotelling's statistics with the generalized delta method. Unlike MLEs, our proposal has both expression and asymptotic standard errors in closed-form. In order to quantify the performance of proposed LC estimators and to compare it with that due to MLEs, we realize Monte Carlo experiments. Finally, we apply the proposed method based on log-cumulant (LC) expressions with real data. Results suggest new LC estimators may outperform meaningfully the MLEs.

Keywords: Beta-Weibull model. Mellin transform. Probability weighted moments.
Goodness-of-fit measure. Monte Carlo simulation.

3.1 Introduction

Survival analysis tools have been required in several reliability issues; as, for example, in lifetimes [35] and industrial [147] areas, biomedical sciences [5], and reliability engineering [148]. In particular, survival data impose some difficulties (e.g., censured nature and resulting empirical distributions having heavy tail and asymmetric behavior) and, therefore, require the proposal of tailored modeling and inference methods. Several probability distributions have been proposed as answers to those issues, but few works have addressed the proposal of alternative inference methods for these extended models. In this chapter, we advance in the last sense.

In recent years, several works have addressed to extend classical models through generators of distributions. Some examples are: the Marshall and Olking (MO)-G class proposed by Marshall and Olkin [7], the generalized exponential (Exp)-G class by Gupta and Kundu [54], the Beta-G class by Eugene *et al.* [53], the Gamma-G class by Zografos and Balakrishnan [55] and Ristic and Balakrishnan [56], the Kw-G class by Cordeiro and Castro [8], the McDonald (Mc)-G class by Cordeiro *et al.* [9], the TX-G class by Alzaatreh *et al.* [149], and by Aljarra *et al.* [11], and the Weibull-G class by Cordeiro *et al.* [12]. The diagram 5.4 displays a relationship among these classes.

We illustrate some issues which may arrive to apply likelihood-based inference in models obtained from those classes.

Consider the extended Weibull model proposed by Marshall and Olkin [7]. Hirose [150] applied this model in breakdown voltage estimation. He employed the maximum likelihood estimates (MLEs) to this end. MLEs do not often present closed-form and, as a consequence, require the use of interactive optimization sources. Dias [29] discussed about some numerical issues which come from using MLEs in extended distributions. It is known MLEs have well-defined asymptotic properties, but they may present an expressive bias for small and moderate sample sizes. In particular, Dias [29] presented evidence that this bias may be more pronounced for the exponential-Poisson distribution.

The gamma, Weibull and lognormal models remains largely an open field for parameter estimation research [27, 116]. It can be by the fact that both classical moments method and maximum

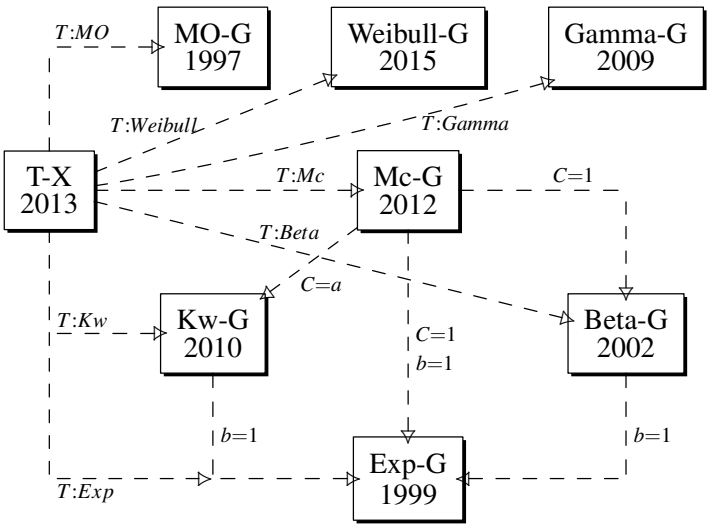


Figure 3.1: Diagram relating the current models.

likelihood parameter estimation techniques bring to systems of non-linear ill behaved systems of equations [151–153] that do not allow the use of classical numerical estimation approaches, such as Newton-Raphson approach which reported persistent divergence in a panel of cases [153, 154].

The estimation procedure based on the MT has achieved a prominent position in several areas, such as radar image and signal processing. The estimation by means of the MT was worked by Nicolas and Maruani [62]. They showed evidence that such estimators present variance which can be comparable with that of MLEs. The methodology of estimation by second type LCs (or by the MT) consists in to equal the sample and theoretical LC estimators (LCEs). This estimation has shown good performance for flexible distributions arising from the stochastic product or sum model.

Nicolas [22] introduced a diagram alternative to the Pearson system (developed by Rhind [155] and, subsequently, by Pearson [156]) like a qualitative measure of GoF to assess fits of matrix distributions in radar image processing. Nicolas and Maruani [62] provided a simulation study on which the performance of estimation based on LCs was compared with the ML and moment estimations. Recently, Khan and Guida [61] proposed an estimation procedure in terms of log-moments for the matrix $\mathcal{G}$ distribution. Anfinsen and Eltoft [60] have demonstrated matrix LCEs may be great value in polarimetry SAR image processing, since it is possible to derive from them estimators having lower bias and variance. Moreover, Li *et al.* [23] proposed the generalized

gamma model for the empirical modeling of SAR images. With this, they developed the second-kind cumulants, its properties, and applied to the SAR images data set.

In this chapter, we derive a new estimation procedure and a quantitative and qualitative GoF measure for the BW distribution [68]. Both proposals are formulated in terms of theoretic results linked to the BW MT, which is also proposed in this chapter. In particular, we derive that estimators based on LCs for the BW parameters present closed-form and their asymptotic covariances are analytically tractable. We also propose a type Hotelling statistic for the BW LCEs and MLEs which may be used as a GoF measure. We perform a Monte Carlo study on which the proposed estimation method is compared with MLEs adopting bias and mean squared error (MSE) as figures of merit. Besides, we perform an application with actual lifetime data in order to quantify the impact of applying MLEs and LCs on the proposed GoF measure. Results point out that our proposal may outperform that based on the likelihood function (widely used in the survival analysis).

The chapter is outlined as follows. Section 3.2 presents the BW distribution. In Section 3.3, we discuss the MT as a special PWM for the BW model. In Section 3.4, we present a new estimation criterion and a GoF measure. Numerical results are displayed in Section 3.5. Finally, the conclusions are exhibited in Section 3.6.

3.2 The BW model

The Weibull distribution is commonly used to model lifetimes of systems. However, its hrf does not accommodate non-monotone hazard rates and this fact motivates the proposal of extended Weibull models which outperform this gap. Famoye *et al.* [68] introduced the BW model which has been widely employed. Other properties of the BW model were derived and studied by Lee *et al.* [6], and Cordeiro *et al.* [157, 158]. From (2.2) and (2.3), the BW pdf and hrf are respectively:

$$f(x) = \frac{\alpha x^{\alpha-1}}{B(a,b)}\frac{1}{\lambda^{\alpha}}\mathrm{e}^{-b\left(\frac{x}{\lambda}\right)^{\alpha}}\left[1-\mathrm{e}^{-\left(\frac{x}{\lambda}\right)^{\alpha}}\right]^{a-1},$$

and

$$h(x) = \frac{\alpha x^{\alpha-1}}{\lambda^{\alpha}} \frac{e^{-b\left(\frac{x}{\lambda}\right)^{\alpha}}\left[1-e^{-\left(\frac{x}{\lambda}\right)^{\alpha}}\right]^{a-1}}{B_{1-e^{-\left(\frac{x}{\lambda}\right)^{\alpha}}}(a,b)}.$$

This distribution is denoted by $X \sim BW(a,b,\lambda,\alpha)$. Fig. 2.1 displays some BW pdf and hrf curves. One can note that its hrf may assume constant, increase, decrease, bathtub and upside-down bathtub forms. These behaviors are sought because they are found in practical situations.

3.2.1 *MLE of BW model*

Let $x_1, x_2, \ldots, x_n$ be an observed sample obtained from $X \sim \mathrm{BW}(a=1,b,\lambda,\alpha)$, the associated log-likelihood function is given by

$$\ell(\boldsymbol{\theta}) = n\log(\alpha) - n\alpha\log(\lambda) - n\log B(1,b) + (\alpha-1)\sum_{r=1}^{n}\log(x_r) - b\sum_{r=1}^{n}\left(\frac{x_r}{\lambda}\right)^{\alpha},$$

and the corresponding components of the score vector,

$$\mathrm{U}(\boldsymbol{\theta}) = (\mathrm{U}_{\alpha}, \mathrm{U}_{\lambda}, \mathrm{U}_{b}) = \left(\frac{\mathrm{d}\,\ell(\boldsymbol{\theta})}{\mathrm{d}\,\alpha}, \frac{\mathrm{d}\,\ell(\boldsymbol{\theta})}{\mathrm{d}\,\lambda}, \frac{\mathrm{d}\,\ell(\boldsymbol{\theta})}{\mathrm{d}\,b}\right)^{\top},$$

are give by:

$$\begin{aligned}
\mathrm{U}_{\alpha} &= \frac{n}{\alpha} + \sum_{r=1}^{n}\log\left(\frac{x_r}{\lambda}\right) - b\sum_{r=1}^{n}\left(\frac{x_r}{\lambda}\right)^{\alpha}\log\left(\frac{x_r}{\lambda}\right), \\
\mathrm{U}_{\lambda} &= -n\frac{\alpha}{\lambda} + b\frac{\alpha}{\lambda^{\alpha+1}}\sum_{r=1}^{n}x_r^{\alpha}, \\
\mathrm{U}_{b} &= -n\{\Psi(0,b) - \Psi(0,1+b)\} - \frac{1}{\lambda^{\alpha}}\sum_{r=1}^{n}x_r^{\alpha}.
\end{aligned} \tag{3.1}$$

where we employed the polygamma function given by: $\Psi(k,x) = \frac{\mathrm{d}^{k+1}}{\mathrm{d}x^{k+1}}\log\Gamma(x)$ [22].

The MLE for $\boldsymbol{\theta}$ is defined as

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^4}{\arg\max}[\ell(\boldsymbol{\theta})],$$

or, equivalently, as a solution of system obtained from making Equations (3.1) equal to null vector. One can note that MLEs do not present closed-form expressions and interactive techniques are

required, such as BFGS (Broyden-Fletcher-Goldfarb-Shanno) [159]. Examples these models are the beta-Pareto [72], beta-Laplace [129], BW [68], and BF [160] distributions.

For covariance matrix estimation, and T^2 Statistic, we can employ as an approximation the inverse of FIM [110, p. 181–182], $\boldsymbol{\Sigma}_{\text{ML}}$ say, given in (2.24), since it is the asymptotic covariance matrix (ACM) of the ML estimators [110, p. 181–182].

Not always the calculation of FIMs for the beta-G model is analytically tractable, thereupon a common solution for this problem is the use of the observed information matrix instead of the FIM, as supported by [72, 129, 133]. Therefore, the observed information matrix is an estimator for the FIM [134], this is,

$$\boldsymbol{\Sigma}_{\text{ML}} \approx \left[-\frac{\mathrm{d}^2\,\ell(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}^{\top}\,\mathrm{d}\boldsymbol{\theta}}\right]^{-1} = \begin{bmatrix} \mathrm{U}_{\alpha\alpha} & \mathrm{U}_{\alpha\lambda} & \mathrm{U}_{\alpha b} \\ \mathrm{U}_{\lambda\alpha} & \mathrm{U}_{\lambda\lambda} & \mathrm{U}_{\lambda b} \\ \mathrm{U}_{b\alpha} & \mathrm{U}_{b\lambda} & \mathrm{U}_{bb} \end{bmatrix}^{-1}, \tag{3.2}$$

where $\mathrm{U}_{\alpha\alpha} = \frac{n}{\alpha^2} + b\xi_2^{\text{BW}}$, $\mathrm{U}_{\alpha\lambda} = \mathrm{U}_{\lambda\alpha} = \frac{1}{\lambda}(n - b\xi_3^{\text{BW}})$, $\mathrm{U}_{\alpha b} = \mathrm{U}_{b\alpha} = \xi_1^{\text{BW}}$, $\mathrm{U}_{\lambda\lambda} = {}^{[b\alpha(\alpha+1)\xi_0^{\text{BW}} - n\alpha]}\!/\!_{\lambda^2}$, $\mathrm{U}_{\lambda b} = \mathrm{U}_{b\lambda} = -\frac{\alpha}{\lambda}\xi_0^{\text{BW}}$, $\mathrm{U}_{bb} = n[\Psi(1,b) - \Psi(1,1+b)]$, $\xi_s^{\text{BW}} = \sum_{r=1}^{n}\left(\frac{x_r}{\widehat{\lambda}}\right)^{\widehat{\alpha}}\log^s\left(\frac{x_r}{\widehat{\lambda}}\right)$, for $s = 0,1,2$; and $\xi_3^{\text{BW}} = \sum_{r=1}^{n}\left(\frac{x_r}{\lambda}\right)^{\alpha}\left[\log\left(\frac{x_r}{\lambda}\right)^{\alpha} + 1\right]$.

It is known that MLEs present a bias with order $O(n^{-1})$ and it suggests the proposal of alternative estimation strategies, where $O(\bullet)$ represents magnitude order. Dias [29] presented evidence that plane likelihood events can occur of using MLEs in extend distributions. Thus, new estimation procedures for distribution classes (such as the beta-G model) are sought. In what follows, we present some theoretical results which aim to propose new GoF measure and estimation procedure for the BW model.

3.3 Second Kind Statistics for the BW Model

In this section, is explored the theoretical properties for the BW model. Because, the PWM and MT have already been reported in the previous chapter. This discussion aims to provide contexts for our theoretical contributions.

3.3.1 *Theoretical properties for the BW model*

The BW MT is given by the following Theorem.

Theorem 2 *Let $X \sim BW(1,b,\lambda,\alpha)$, its MT is given by*

$$\phi_{BW}(s) = \frac{\lambda^{s-1}}{B(1,b)} b^{-\frac{(s-1+\alpha)}{\alpha}} \Gamma\left(1+\frac{s-1}{\alpha}\right).$$

■

The proof is detailed in [161]. Assuming $s = \nu + 1$ in Theorem 2, one has that the log-moments of the BW model, m_ν, for $\nu \in \mathbb{N}$, are given by

$$m_\nu = \frac{\lambda^\nu}{B(1,b)} b^{-\frac{(\nu+\alpha)}{\alpha}} \Gamma\left(1+\frac{\nu}{\alpha}\right). \tag{3.3}$$

Several authors [6,68,157] have proposed expansions for BW log-moments, Equation (3.3) presents a closed-form expression. Hence, using Theorem 2 in (2.18), we have the BW LCs:

Proposition 7 *If $X \sim BW(1,b,\lambda,\alpha)$, its second kind cumulants are given by*

$$\widetilde{\kappa}_\nu = \begin{cases} \log(\lambda) + \frac{\Psi(1)-\log(b)}{\alpha}, & \text{for } \nu = 1; \\ \\ \frac{1}{\alpha^\nu}\Psi^{(\nu-1)}(1), & \text{for } \nu > 1, \end{cases}$$

One can note that BW LCs with order higher than two depend only on α. We remark that the third order LC is negative.

3.4 Estimation and GoF for the BW model

This section aims to present both estimation procedures by LC and maximum likelihood and a GoF measure for the BW distribution. First we present the LCEs for BW parameters based on Proposition 7. Finally, we propose a GoF measure for the BW distribution, combining the asymptotic behaviors of LCEs with the Hotelling T^2 statistics.

3.4.1 *LCEs for the BW model*

Now, we are in position to derive the LCEs for b, λ and α parameters, say $\widehat{b}_{LC}$, $\widehat{\lambda}_{LC}$ and $\widehat{\alpha}_{LC}$, respectively. Combining the Proposition 7 with the relations in (2.19) replacing $\widetilde{m}_{\nu}$ by the corresponding sample log-moments, one has closed-form estimators for λ and α are given by, respectively: Assuming that $\widehat{b}_{LC}$ known,

$$\widehat{\lambda}_{LC} = \exp\left\{\widehat{\widetilde{m}}_1 - \frac{\psi(1) - \log(\widehat{b}_{LC})}{\widehat{\alpha}_{LC}}\right\},$$

and

$$\widehat{\alpha}_{LC} = \sqrt{\frac{\psi^{(1)}(1)}{\widehat{\widetilde{m}}_2 - \widehat{\widetilde{m}}_1^2}}.$$

Further, LCE for b is defined as a solution of the following non-linear equation:

$$\widehat{b}_{LC} = \exp\left\{-\sqrt[3]{\Lambda\,\Upsilon^{3/2}}\left[\psi(1) - \log(\widehat{b}_{LC})\right] + \psi(1)\right\},$$

where

$$\Lambda = \frac{\psi^{(2)}(1)}{\widehat{\widetilde{m}}_3 - 3\widehat{\widetilde{m}}_1\widehat{\widetilde{m}}_2 + 2\widehat{\widetilde{m}}_1^3} \quad \text{and} \quad \Upsilon = \frac{\widehat{\widetilde{m}}_2 - \widehat{\widetilde{m}}_1^2}{\psi^{(1)}(1)}.$$

3.4.2 *Covariance matrix of LCEs*

First we derive the covariance matrix of BW LCEs. Subsequently, we furnish a new GoF measure for the BW model. Consider the following quantities;

$$\widetilde{\boldsymbol{m}} = (\widetilde{m}_1, \ldots, \widetilde{m}_p)^\top \quad \text{and} \quad \widehat{\widetilde{\boldsymbol{m}}} = (\widehat{\widetilde{m}}_1, \ldots, \widehat{\widetilde{m}}_p)^\top.$$

The central limit theorem and the Cramér-Wold theorem [162] proved that

$$\sqrt{n}(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_{\nu}(\boldsymbol{0},\, \boldsymbol{M}_{\nu}),$$

where $\boldsymbol{x} \sim \mathcal{N}_{\nu}(\boldsymbol{\mu},\, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and [22, 26]

$$\boldsymbol{M}_{\nu} = n\,\mathrm{E}\left\{(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}})(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}})^\top\right\} = \left\{\widetilde{m}_{r+j} - \widetilde{m}_r\widetilde{m}_j\right\}\Big|_{r,j=1}^{\nu}. \tag{3.4}$$

Now, let $g_\nu : \mathbb{R}^\nu \to \mathbb{R}$ be moment-to-cumulant transformation function (which are continuously differentiable),

$$\widetilde{\kappa}_\nu = g_\nu(\widetilde{m}_1, \ldots, \widetilde{m}_\nu) \quad \text{and} \quad \widehat{\widetilde{\kappa}}_\nu = g_\nu(\widehat{\widetilde{m}}_1, \ldots, \widehat{\widetilde{m}}_\nu).$$

Thence, together with the Cramér-Wold theorem, we have [163]

$$\sqrt{n}(\widehat{\widetilde{\boldsymbol{\kappa}}} - \widetilde{\boldsymbol{\kappa}}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_\nu(\mathbf{0},\, \boldsymbol{K}_\nu),$$

where

$$\boldsymbol{K}_\nu = \boldsymbol{J}_\nu^\top \cdot \boldsymbol{M}_\nu \cdot \boldsymbol{J}_\nu, \tag{3.5}$$

$$\boldsymbol{\nabla} g_\nu = \left[\frac{\mathrm{d}}{\mathrm{d}\widetilde{m}_1} g_\nu(\widetilde{\boldsymbol{m}}), \ldots, \frac{\mathrm{d}}{\mathrm{d}\widetilde{m}_\nu} g_\nu(\widetilde{\boldsymbol{m}}) \right]^\top, \text{ for } \nu = 1, \ldots, r,$$

and

$$\boldsymbol{J}_\nu = \left[\boldsymbol{\nabla} g_1 \middle| \cdots \middle| \boldsymbol{\nabla} g_\nu \right]^\top.$$

In this chapter, we use both asymptotic results of LCEs (given in (3.5)) and of MLEs. These last have behavior similar to (3.5), replacing $\boldsymbol{M}_\nu$ in (3.5) by the observed information matrix as estimator for its expected counterpart which is analytically intractable. The following proposition presents these matrices.

Proposition 8 *Let $X \sim BW(1, b, \lambda, \alpha)$ $\nu = 3$ in (3.5), its ACM of LCE is given by*

$$\boldsymbol{K}_{LC} = \frac{1}{\alpha^6} \begin{bmatrix} \widetilde{\kappa}_{22} & \widetilde{\kappa}_{23} \\ \widetilde{\kappa}_{32} & \widetilde{\kappa}_{33} \end{bmatrix},$$

where

$$\begin{aligned}
\widetilde{\kappa}_{22} &= \alpha^2[\Psi(3,1) + 2\Psi^2(1,1)], \\
\widetilde{\kappa}_{23} &= \alpha[\Psi(4,1) + 6\Psi(1,1)\Psi(2,1)], \\
\widetilde{\kappa}_{33} &= \Psi(5,1) + 9\Psi(1,1)\Psi(3,1) + 9\Psi^2(2,1) + 6\Psi^3(1,1).
\end{aligned}$$

Further the ACM of MLEs is given [161]

$$\boldsymbol{K}_{ML} = \left(\frac{U_{\alpha\alpha}U_{bb} - U_{\lambda b}^2}{\alpha^8 |\boldsymbol{H}|} \right) \cdot \begin{bmatrix} 4\alpha^2\Psi^2(1,1) & 6\alpha\Psi(1,1)\Psi(2,1) \\ 6\alpha\Psi(1,1)\Psi(2,1) & 9\Psi^2(2,1) \end{bmatrix},$$

where

$$|\boldsymbol{H}| = U_{\alpha\alpha}(U_{\lambda\lambda}U_{bb} - U_{\lambda b}^2) + U_{\alpha\lambda}(U_{\alpha b}U_{\lambda b} - U_{\alpha\lambda}U_{bb}) + U_{\alpha b}(U_{\alpha\lambda}U_{\lambda b} - U_{\alpha b}U_{\lambda\lambda}),$$

and the proof of this proposition is provided in Appendix C.

3.4.3 *Hotelling's statistic for BW model*

In order to compare our proposal with MLEs, we propose a new GoF measure for the BW model. To that end, we combine the Hotelling's statistic [164] with the generalized delta method [165] as follows.

The following result is known from the multivariate analysis literature [128]: Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^4$ be a random sample drawn from a 4-points random vector. The T^2 Hotelling statistic is given by

$$T^2 = (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)^\top \boldsymbol{S}^{-1} (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0), \tag{3.6}$$

where $\overline{\boldsymbol{x}} = \frac{1}{n}\sum_{r=1}^n \boldsymbol{x}_r$, $\quad \boldsymbol{S} = \frac{1}{n-1}\sum_{r=1}^n (\boldsymbol{x}_r - \overline{\boldsymbol{x}})(\boldsymbol{x}_r - \overline{\boldsymbol{x}})^\top$ and $\boldsymbol{\mu}_0 \in \mathbb{R}^n$ is a mean vector associated with any random vector.

In practice, (3.6) is often considered to test $\mathcal{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ when the covariance matrix is unknown. As a decision rule, if an outcome of T^2 is large (representing that $\boldsymbol{x}$ is distinct of $\boldsymbol{\mu}_0$), one has sample evidence to reject $\mathcal{H}_0$. This criterion is feasible because the T^2 statistics follows asymptotically the Snedecor $\mathcal{F}$ model given by $(n-1)\nu(n-\nu)^{-1}X$ with $X \sim \mathcal{F}_{\nu,n-\nu}$, where "$\mathcal{F}_{\nu,n-\nu}$" denoted the Snedecor $\mathcal{F}$ distribution with ν and $n-\nu$ degrees of freedom.

Now we are in position to provide a new GoF measure. Note that, from the stochastic majoring $T^2 \leq C^2$, one can obtain an ellipsoid equation centered on LCEs having axes in the directions of the eigenvectors of $\boldsymbol{K}$. Based on the generalized delta method [165], $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$ follows asymptotically

bivariate normal distribution; i.e.,

$$\left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right) \overset{a}{\sim} \mathcal{N}_2(\mathbf{0},\, \boldsymbol{K}),$$

where $\widehat{\widetilde{\kappa}}_2 = g_2(\widehat{\boldsymbol{\theta}})$, $\widehat{\widetilde{\kappa}}_3 = g_3(\widehat{\boldsymbol{\theta}})$, $\widehat{\boldsymbol{\theta}} = (\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_\nu)^\top$, $\boldsymbol{K} = \boldsymbol{\nabla}^\top \boldsymbol{\Sigma} \boldsymbol{\nabla}$,

$$\boldsymbol{\Sigma} = \begin{cases} \boldsymbol{\Sigma}_{\text{ML}}, & \text{for MLEs,} \\ \\ \boldsymbol{\Sigma}_{\text{LC}}, & \text{for LCEs,} \end{cases}$$

where $\boldsymbol{\Sigma}_{\text{ML}}$ is given in (3.2), and $\boldsymbol{\Sigma}_{\text{LC}} = \text{Cov}(\widehat{m}_r, \widehat{m}_j)\Big|_{r,j=1}^{\nu} = \frac{1}{n}(\widetilde{m}_{r+j} - \widetilde{m}_r \widetilde{m}_j)$; and

$$\boldsymbol{\nabla} = \begin{cases} \boldsymbol{\nabla}_{\text{ML}}, & \text{for MLEs ,} \\ \\ \boldsymbol{\nabla}_{\text{LC}}, & \text{for LCEs ,} \end{cases}$$

where

$$\boldsymbol{\nabla}_{\text{ML}}^\top = \begin{bmatrix} \frac{\mathrm{d}\, g_2(\boldsymbol{\theta})}{\mathrm{d}\,\theta_1} & \frac{\mathrm{d}\, g_2(\boldsymbol{\theta})}{\mathrm{d}\,\theta_2} & \cdots & \frac{\mathrm{d}\, g_2(\boldsymbol{\theta})}{\mathrm{d}\,\theta_r} \\ \\ \frac{\mathrm{d}\, g_3(\boldsymbol{\theta})}{\mathrm{d}\,\theta_1} & \frac{\mathrm{d}\, g_3(\boldsymbol{\theta})}{\mathrm{d}\,\theta_2} & \cdots & \frac{\mathrm{d}\, g_3(\boldsymbol{\theta})}{\mathrm{d}\,\theta_r} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\nabla}_{\text{LC}}^\top = \begin{bmatrix} \frac{\mathrm{d}\, g_2(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_1} & \frac{\mathrm{d}\, g_2(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_2} & \cdots & \frac{\mathrm{d}\, g_2(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_r} \\ \\ \frac{\mathrm{d}\, g_3(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_1} & \frac{\mathrm{d}\, g_3(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_2} & \cdots & \frac{\mathrm{d}\, g_3(\widetilde{\boldsymbol{m}})}{\mathrm{d}\,\widetilde{m}_r} \end{bmatrix}.$$

Hence, we have the associated contour equation with $(1-\eta)$ confidence are given by: For η as a specific nominal level,

$$T^2 = n \left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right)^\top \widehat{\boldsymbol{K}}^{-1} \left(\begin{bmatrix} \widehat{\widetilde{\kappa}}_2 \\ \widehat{\widetilde{\kappa}}_3 \end{bmatrix} - \begin{bmatrix} \widetilde{\kappa}_2 \\ \widetilde{\kappa}_3 \end{bmatrix} \right) \leq \frac{\chi^2_{2,\eta}}{n},$$

where $\chi^2_{2,\eta}$ denotes the ηth percentile of the chi-squared distribution with two degrees of freedom, $\boldsymbol{K}^{-1} = \left\{ \widetilde{\kappa}^{r,j} \right\}_{r,j=2,3}$, and the probability of random vector belonging to the contour is

$$P\left(T^2 \leq \chi^2_{2,\eta}\right) = 1 - \eta.$$

Based on the last discussion, we can define a GoF measure for the BW distribution according to the approach of Vasconcelos *et al.* [161].

Proposition 9 *Let X be a random variable following the BW distribution with parameters $a = 1$, $b > 0$, $\alpha > 0$ and $\lambda > 0$, then the T^2 statistic based on LCEs is given by*

$$T^2 = \begin{cases} T^2_{ML}, & \textit{for MLEs} \\ \\ T^2_{LC}, & \textit{for LCEs} \end{cases}, \tag{3.7}$$

where

$$T^2_{ML} = \frac{n\widehat{\alpha}^6}{4}\left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2}\right)^2 \left(\frac{|\widehat{\boldsymbol{H}}|}{U_{\widehat{\alpha}\widehat{\alpha}}U_{\widehat{b}\widehat{b}} - U^2_{\widehat{\lambda}\widehat{b}}}\right),$$

and

$$T^2_{LC} = \frac{n\alpha^6}{\widetilde{\widetilde{\kappa}}_{33}\widetilde{\widetilde{\kappa}}_{22} - \widetilde{\widetilde{\kappa}}^2_{23}}\left\{\widetilde{\widetilde{\kappa}}^{22}\left[\Psi(1,1)\left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2}\right)\right]^2 + \widetilde{\widetilde{\kappa}}^{33}\left[\Psi(2,1)\left(\frac{1}{\widehat{\alpha}^3} - \frac{1}{\alpha^3}\right)\right]^2 - 2\widetilde{\widetilde{\kappa}}^{23}\Psi(1,1)\Psi(2,1)\left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2}\right)\left(\frac{1}{\widehat{\alpha}^3} - \frac{1}{\alpha^3}\right)\right\},$$

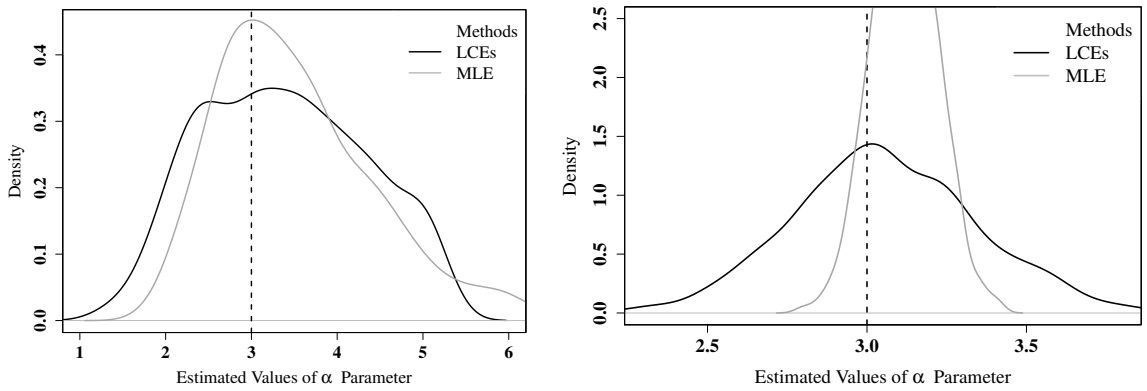
where $|\widehat{\boldsymbol{H}}|$, $\widetilde{\widetilde{\kappa}}^{22}$, $\widetilde{\widetilde{\kappa}}^{33}$, $\widetilde{\widetilde{\kappa}}^{23}$, $\widehat{\alpha}$, $\widehat{\lambda}$, and $\widehat{b}$ are the estimators of $|\boldsymbol{H}|$, $\widetilde{\kappa}^{22}$, $\widetilde{\kappa}^{33}$, $\widetilde{\kappa}^{23}$, α, λ, and b respectively; the proof of this Proposition is presented in Appendix C. We provide (3.7) *as a GoF measure for the BW law.*

3.5 Numerical results

This section discusses results of a simulation study, performed to quantify the performance of discussed methods. Beyond, experiments with real data are made to detect the effect of the use of estimates into the new GoF measure.

3.5.1 *Simulation study*

To assess the performance of LCEs, a Monte Carlo simulation is accomplished with sample sizes $n = 10, 200, 300, 500$, $\lambda = 1$, $\alpha \in \{3,4,5\}$ and $b \in \{4,5,6,7\}$. For each combination (b, λ, α, n) we generate 1.000 Monte Carlo replications on which proposed estimators are assessed



(a) Sample size $n = 10$ (b) Sample size $n = 500$

Figure 3.2: Density of estimates of parameter α by the two methods.

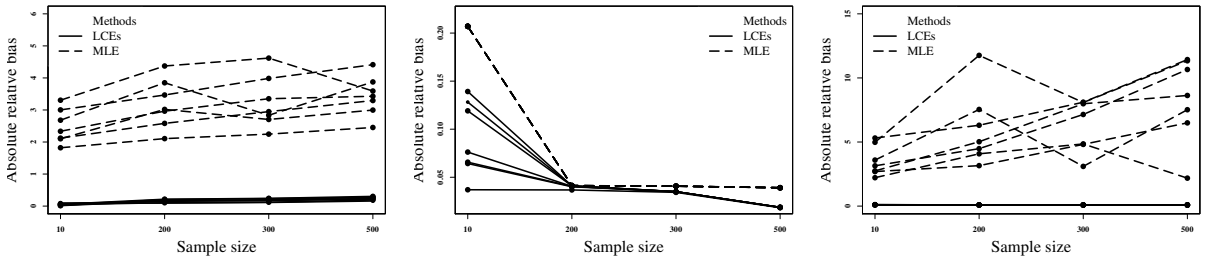


Figure 3.3: The absolute relative bias of the estimates of parameters λ (a), α (b) and b (c) by two methods.

for several sample sizes. This study is carried out using the `R` statistical software with `BB` and `MaxLik` packages and `BBsolve` and `maxBFGS` functions. As figures of merit, we use the following quantities to compare our proposal with MLEs: (a) absolute bias $|\mathrm{E}(\widehat{\theta}_i) - \theta_i|$, (b) relative absolute bias $|\mathrm{E}(\widehat{\theta}_i) - \theta_i|/\theta_i$, and (c) square root of mean square error, $\sqrt{\mathrm{MSE}}$.

Table 3.1 presents the values of adopted criteria. One can note that the estimates based on LCEs present values of $\sqrt{\mathrm{MSE}}$ and bias that decrease when increasing the sample size, while MLEs have a good performance only for α. Fig. 3.3 shows values of the absolute relative bias. In all cases, our method outperform MLEs. Fig. 3.2 displays empirical densities of estimates for both ML and LC criteria. It is noticeable that our method converges faster for the expected asymptotic behavior.

Table 3.1: Absolute biases and root mean squared error between parentheses

n	Parameters	LCE			MLE		
	(λ, α, b)	$\widehat{\lambda}$	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{\lambda}$	$\widehat{\alpha}$	$\widehat{b}$
10	(1, 3, 4)	0.0970 (0.4375)	0.3847 (1.0240)	0.4378 (0.4829)	2.6814 (2.7717)	0.6209 (1.3079)	14.3624 (16.0946)
	(1, 3, 5)	0.0772 (0.4541)	0.4175 (1.0681)	0.5065 (0.5529)	3.0021 (3.1089)	0.6209 (1.3079)	13.4747 (14.4035)
	(1, 3, 6)	0.0199 (0.4761)	0.3574 (1.0116)	0.5762 (0.6176)	3.3048 (3.4247)	0.6209 (1.3079)	13.2928 (14.3557)
	(1, 4, 5)	0.0443 (0.4900)	0.3047 (1.1650)	0.5569 (0.6260)	2.1011 (2.1849)	0.8279 (1.7439)	24.8966 (30.4742)
	(1, 4, 6)	0.0143 (0.4720)	0.2572 (1.1157)	0.6202 (0.6940)	2.1212 (2.1891)	0.8279 (1.7439)	18.8525 (23.6321)
	(1, 4, 7)	0.0570 (0.4921)	0.2628 (1.1406)	0.6881 (0.7547)	2.3351 (2.4381)	0.8279 (1.7439)	19.2111 (23.5978)
	(1, 5, 7)	0.0588 (0.5330)	0.1851 (1.2944)	0.7533 (0.8554)	1.8208 (1.8918)	1.0348 (2.1796)	37.1777 (47.9914)
200	(1, 3, 4)	0.0926 (0.2496)	0.1232 (0.4600)	0.3684 (0.3892)	3.8489 (3.8671)	0.1238 (0.2049)	30.1620 (31.0186)
	(1, 3, 5)	0.1499 (0.2758)	0.1232 (0.4600)	0.4383 (0.4618)	3.4671 (3.4700)	0.1238 (0.2049)	15.7810 (15.8675)
	(1, 3, 6)	0.2042 (0.3101)	0.1232 (0.4600)	0.5220 (0.5446)	4.3724 (4.3806)	0.1238 (0.2049)	24.4781 (24.7023)
	(1, 4, 5)	0.1049 (0.3198)	0.1620 (0.6089)	0.4576 (0.4862)	3.0176 (3.0288)	0.1651 (0.2732)	58.8000 (60.4853)
	(1, 4, 6)	0.1570 (0.3422)	0.1606 (0.6072)	0.5381 (0.5713)	2.5815 (2.5937)	0.1651 (0.2732)	26.8884 (28.0325)
	(1, 4, 7)	0.2149 (0.3665)	0.1620 (0.6089)	0.6114 (0.6448)	2.9631 (2.9700)	0.1651 (0.2732)	35.1489 (35.7311)
	(1, 5, 7)	0.1726 (0.4039)	0.1842 (0.7340)	0.6334 (0.6714)	2.1059 (2.1142)	0.2063 (0.3415)	44.1454 (47.8689)
300	(1, 3, 4)	0.1129 (0.2321)	0.1050 (0.3982)	0.3581 (0.3768)	2.8348 (2.8431)	0.1222 (0.1829)	12.3868 (13.0370)
	(1, 3, 5)	0.1710 (0.2608)	0.1050 (0.3982)	0.4360 (0.4565)	3.9843 (3.9903)	0.1223 (0.1829)	24.0742 (24.1696)
	(1, 3, 6)	0.2325 (0.3065)	0.1050 (0.3982)	0.5152 (0.5372)	4.6184 (4.6247)	0.1223 (0.1829)	29.0995 (29.4373)
	(1, 4, 5)	0.1214 (0.3014)	0.1399 (0.5309)	0.4519 (0.4778)	2.7040 (2.7105)	0.1630 (0.2439)	40.4146 (41.5011)
	(1, 4, 6)	0.1809 (0.3257)	0.1399 (0.5309)	0.5285 (0.5551)	2.9484 (2.9590)	0.1630 (0.2439)	42.8914 (43.9331)
	(1, 4, 7)	0.2413 (0.3599)	0.1399 (0.5309)	0.6067 (0.6330)	3.3507 (3.3622)	0.1630 (0.2439)	55.9035 (58.6638)
	(1, 5, 7)	0.1975 (0.3904)	0.1716 (0.6592)	0.6154 (0.6467)	2.2460 (2.2532)	0.2037 (0.3049)	56.7250 (59.4779)
500	(1, 3, 4)	0.1582 (0.2151)	0.0560 (0.3004)	0.3506 (0.3652)	3.8750 (3.8775)	0.1172 (0.1569)	30.0781 (30.1264)
	(1, 3, 5)	0.2149 (0.2612)	0.0560 (0.3004)	0.4340 (0.4489)	4.4134 (4.4171)	0.1172 (0.1569)	32.4681 (32.5623)
	(1, 3, 6)	0.2719 (0.3087)	0.0560 (0.3004)	0.5223 (0.5380)	3.5910 (3.6114)	0.1172 (0.1568)	13.0823 (14.9157)
	(1, 4, 5)	0.1890 (0.2716)	0.0747 (0.4005)	0.4374 (0.4570)	2.9964 (3.0019)	0.1562 (0.2091)	56.6910 (57.5881)
	(1, 4, 6)	0.2405 (0.3093)	0.0747 (0.4006)	0.5251 (0.5465)	3.2949 (3.3088)	0.1562 (0.2091)	63.9427 (66.7045)
	(1, 4, 7)	0.3029 (0.3621)	0.0747 (0.4005)	0.6002 (0.6193)	3.4282 (3.4387)	0.1562 (0.2091)	60.4176 (62.7743)
	(1, 5, 7)	0.2761 (0.3682)	0.0917 (0.4979)	0.6077 (0.6321)	2.4534 (2.4596)	0.1953 (0.2614)	80.0111 (83.1250)

Table 3.2: Description of data sets

Data	Minimum	Mean	Median	Mode	SD	Skewness	Kurtosis	Maximum	VC
Voltage	2.00	177.03	196.50	300.00	114.99	−0.28	1.45	300.00	64.95
Electronic	0.03	1.94	1.79	NA	1.44	0.60	2.72	5.09	74.20
Glass	0.55	1.51	1.59	1.61	0.32	−0.90	3.92	2.24	21.51

3.5.2 *Experiments with actual data*

In this section, we provide two experiments with actual data to illustrate that the LCEs for BW parameters may outperform meaningfully that due to ML in the lifetime data context. To that end, we consider:

i) Voltage data: This base represents the time of failure and running times for a sample of devices from a field-tracking study of a larger system [157, 166].

ii) Electronic data: This base data represents lifetimes of 20 Electronic Components [167, p. 83–100].

iii) Glass data: This base represents the first set of data on the strengths of 1.5 cm glass fibre, measured at the National Physical Laboratory, England [168].

Table 3.2 displays a descriptive summary of each data set. The first and third has one mode, and second amodal. The electronic and glass data sets presents the smallest standard deviation (SD) (1.44 and 0.32, respectively); while the highest is given in voltage data set. Voltage and glass data sets have the mean small in than median, which results a negative skewness, see in Table 3.2. Also, for voltage and glass data sets, we have the distribution of these data are negative skewness because its symmetry are negative (-0.28 and -0.90). The glass data set has a high kurtosis what implies a difficulty in choosing the possible model for estimating and the others has a small kurtosis. The first data set is homogeneous (small variance) with sample variation coefficient (VC) of 21.51%. The remaining data sets are heterogeneous.

Fig. 3.4 displays the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram, which is a tool alternative to the Pearson's diagram as discussed by Vasconcelos *et al.* [161]. It consists in resulting curves and areas over plane obtained

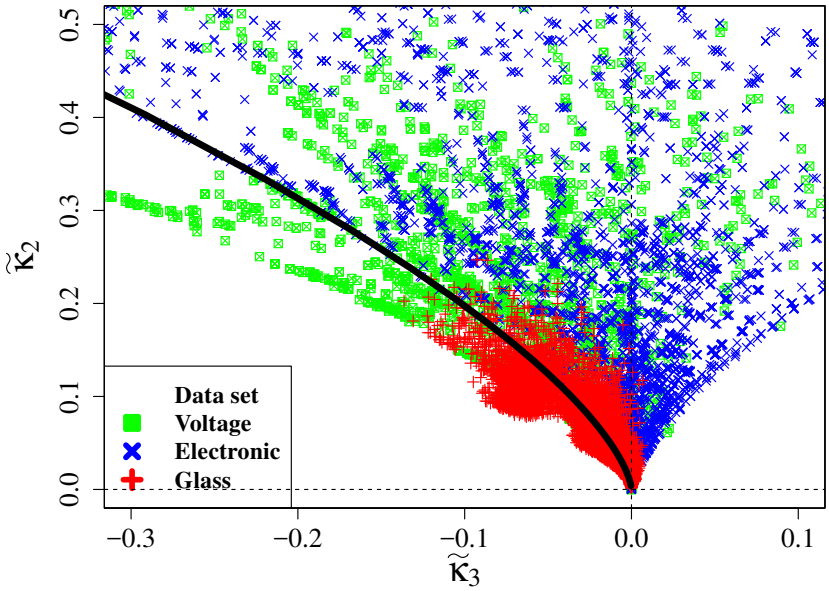
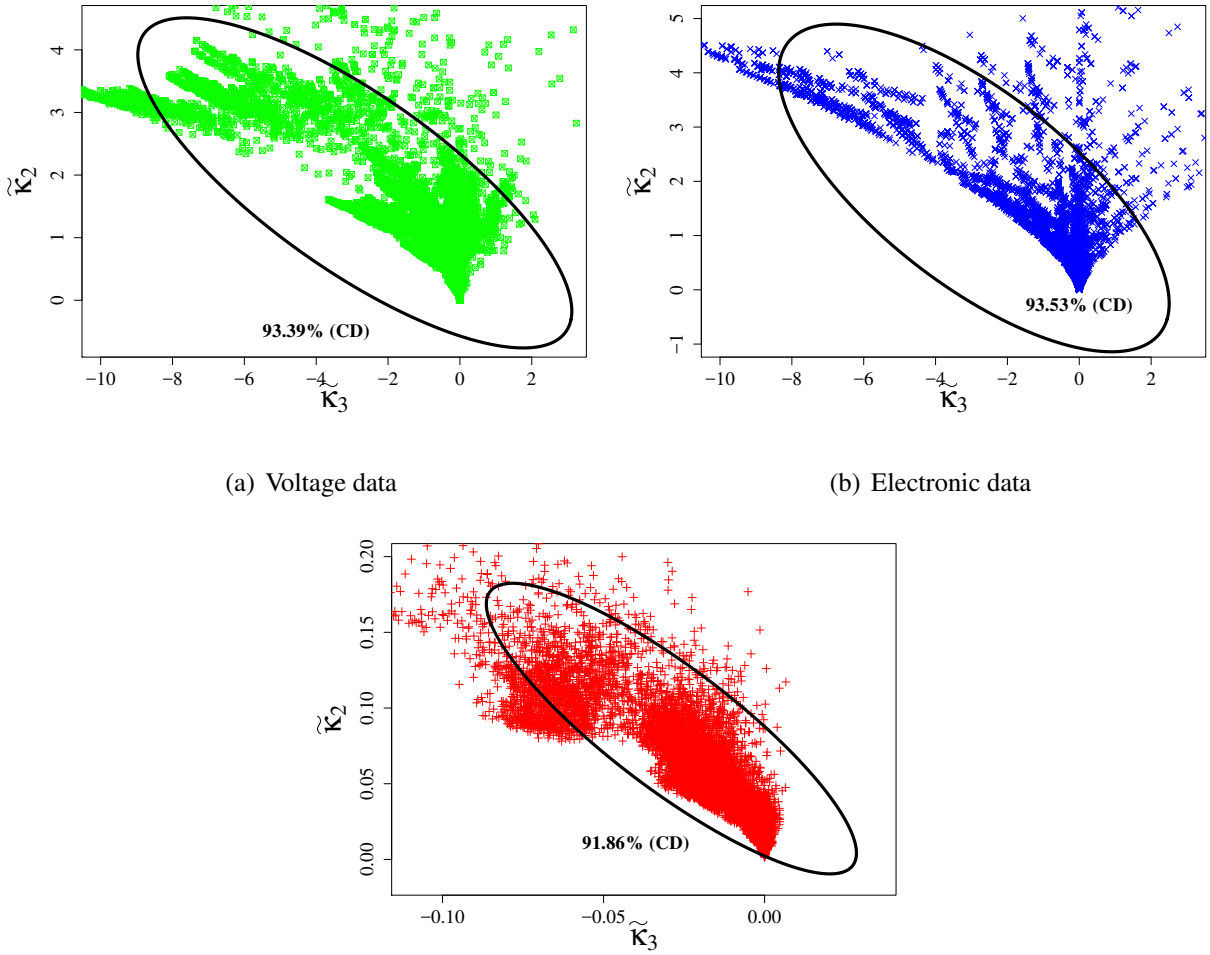


Figure 3.4: Diagram of the LCs $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ showing the manifold of theoretical LCs for the BW, as well as a collection of sample LCs representing voltage (⊠), electronic (×), and glass (+) data sets.

from of writing $\widetilde{\kappa}_2$ in terms of $\widetilde{\kappa}_3$ for all points of the BW parametric space. After sample LCEs were plotted over the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ plane for each data set. It is noticeable considered data sets are over the curve, which indicates the BW model like a good descriptor for considered data sets. This graph can be understood as a qualitative index defined on the Proposition 7 for the BW model.

Now, we wish to assess the impact of both LCEs and ML criteria with respect to the 95% confidence region derived from in Proposition 9. Fig. 3.5 shows these regions adopted for three data sets and for two estimation methods. By a visual inspection, one can initially see that the use of LCEs are more recommended than MLEs with respect to the use of proposed GoF measure.

In order to confirm the previous analysis, we quantify the degree of coverage of confidence curves in Fig. 3.5. Table 3.3 exhibits values of the kind Hotelling statistic using LCEs and MLE methods as well as their p-values. As a conclusion, the T^2 statistic equipped with LCEs presented better performance than that in terms of MLEs.



(a) Voltage data

(b) Electronic data

(c) Glass data

Figure 3.5: Confidence ellipses for each data set.

Table 3.3: Hotelling statistics and p-value with respect to the data sets

Data	LCE		MLE	
	T^2	p-value	T^2	p-value
Voltage	0.2461	0.7835	1.7562	0.1912
Electronic	0.1588	0.8544	1.4353	0.2640
Glass	0.0002	0.9998	2.1485	0.1254

3.6 Conclusion

In this chapter, we have proposed a new estimation procedure based on the MT for the BW model. Moreover, a novel GoF measure for the BW distribution has also been derived. These proposals have aimed to outperform issues with the use of MLE for the BW parameters as well as the absence of a GoF quantity with a qualitative appeal in the lifetime context.

Monte Carlo experiments were conducted in sense to compare our proposal with MLEs. Using biases and MSE as figures of merit, estimators based on LCs for BW parameters may outperform meaningfully MLE. Finally, experiments with actual data were performed, indicating that the proposed GoF measure equipped with LCEs may present better coverage degrees than if it is evaluated at MLEs.

4 The Compound Truncated Poisson Cauchy Model: A Descriptor for Multimodal Data

Abstract

Multimodal data come from several applications in practice such as in the processing of Synthetic Aperture Radar (SAR) images involving multiple clusters. This case has been often modeled by using probability mixtures, but this solution may involve a great number of parameters and the inference stage becomes hard. In order to overcome this gap we propose a probability distribution having only three parameters, which is able to describe multimodal data. Our model is defined by the sum of a random number, following the truncated Poisson law, of independent random variables with the Cauchy model, called *Compound truncated Poisson Cauchy* (CTPC) distribution. We derive some of its properties: characteristic function (cf) and a distance measure between cfs. Further, we provide two estimation procedures for the CTPC parameters: maximum likelihood estimators (MLEs) and quadratic distance estimators (QDEs). Furthermore, we derive a new goodness-of-fit (GoF) measure stemmed from the CTPC law and based on empirical cf. To quantify the performance of both proposed estimators and GoF statistic, we make a Monte Carlo simulation study. According to three figures of merit, results indicate QDEs may present better performance than MLEs. Finally, an experiment with actual SAR data is performed. It aims to describe a segment of SAR intensities (positive real feature of SAR) with at least two types of textures. Our model can outperform six classic distributions and one mixed distribution for modeling SAR intensities: Weibull, gamma, generalized gamma, $\mathcal{K}$, $\mathcal{G}^0$, beta generalized normal, and *Compound truncated Poisson Gamma* models.

Keywords: Stochastic sums. Multimodal distribution. CTPC distribution. Stochastic distance. PolSAR imagery.

4.1 Introduction

Multimodal empirical distributions have been used in several applications in practice; for instance, biomedical sciences [5], image processing [51], industrial [2] areas, breaking strengths of glass fibers [3], and survival times [6]. Probability mixture models are often used to describe data of this nature. However, these models impose a large number of parameters and, as a consequence, making inferences on such parameters becomes a hard task. becomes a hard task.In order to furnish a more parsimonious solution, we propose a probability distribution with only three parameters which is able to describe data with multiple modes. The proposed model is defined by the sum of a random number N of independent random variables following the Cauchy distribution with location and scale parameters, μ and γ, respectively. As subsequently detailed, we assume that the quantity N follows the truncated Poisson distribution. Such distribution is denoted by $X \sim \mathcal{C}(\mu, \gamma)$.

The Cauchy distribution has been employed on several contexts; e.g., to model experiments with circular [31, 32], financial [33], and optical data [34]. In recent years, the Cauchy model has been extended resulting in a number of derived models: (i) the generalized odd half-Cauchy family proposed by Cordeiro *et al.* [12], (ii) the beta-Cauchy distribution proposed by Alshawarbeh *et al.* [169], (iii) the generalized skew-Cauchy distribution proposed by Huang and Chen [37], (iv) the Weibull-power Cauchy distribution proposed by Tahir *et al.* [170], (v) the generalized Cauchy family proposed by Alzaatreh *et al.* [171], (vi) the half-Cauchy distribution proposed by Bosch [40], (vii) the Kumaraswamy-half Cauchy distribution proposed by Ghosh [172], and (viii) the beta-half-Cauchy distribution proposed by Cordeiro and Lemonte [173]. In a general sense, the above models were generated from the T-X (in which X is a random variable, "the transformer", is used to transform another random variable T, "the transformed"). Class proposed by Alzaatreh *et al.* [149]. However, such generator often provide unimodal or bimodal distributions.

Another approach to generate probability distributions is by "compounding", a technique worked by Grushka [43]. Such procedure generate new models by means of employing scale transformations of random vectors. Some examples of such transformations are: (i) minimum or maximum of a random sample; (ii) sum of a random vector; and (iii) minimum on which is one random vec-

tor of its parameters. Golubev [44] proposed the exponentially modified Gaussian model which is suitable for analysis of variabilities featured by a number of biological phenomena often thought to be associated with the lognormal distributions. Teich and Diament [174] showed that, if the parameter of the compound Poisson distribution follows the range distribution, then the negative binomial model is obtained. In Karlis and Xekalaki [30] have addressed an important type of such compositions, the so-called compound N distribution, in which a random variable describes by the following sum:

$$S = X_1 + X_2 + \cdots + X_N, \tag{4.1}$$

where both N and $\{X_l\}$ are independent random variables. Revfeim [46] proposed the compound Poisson exponential distribution for describing as to how little extra information might bere covered from a record of daily rainfall totals compared with a record of monthly rainfall totals. Because (4.1) describes a number of natural phenomena, it is a widely considered model [47]. Revfeim [46] proposed the compound Poisson exponential distribution for describing the total precipitation per day, where the number of days with precipitation is Poisson distributed and the precipitation amount follows the exponential distribution. Second Panger [47], the compound Poisson and compound negative binomial distributions are extensively used in risk theory, in particular to model total claims incurred a fixed period of time. Finally, Thompson [48] applied the compound Poisson distribution for modeling the total monthly rainfall.

Synthetic aperture radar (SAR) imagery have been used as important tool for remote sensing. Each entry of SAR image is associated to a complex-valued element whose squared norm is known as *intensity*. Two questions are important to take into account when working with this type of images: (i) they are contaminated by speckle noise which hampers their processing and (ii) some of their scenarios result in multimodal data. Delignon and Pieczynski [175] assumed that the returning back-scattered field F from an illuminated area is modeled by:

$$F = \sum_{k=1}^{n} F_k, \tag{4.2}$$

where n is the number of scatters and F_k are complex-valued quantities representing the individual scatters.

In this chapter, we adopt the the above-mentioned summation modeling to describe intensities captured by SAR. Further, based on the results in [176–178], we have evidence that the Cauchy distribution could meaningful represent the intensity. Moreover, in [179, 180] we find indications that the truncated Poisson distribution could be a suitable candidate distribution for modeling the number of resolution cells. Therefore, combining these two evidences, we proposed the random sum of the Cauchy distributions with the number of terms following the truncated Poisson law as a descriptor for the squared norm of F. As a result, we introduce a new probabilistic model referred to as the CTPC distribution. The proposed distribution inherits the parameters from its parent distributions, presenting three parameters: two shape and one scale parameter.

Here, we use the random sum of the Cauchy distributions with the number of terms following the truncated Poisson law as a descriptor for the squared norm of F. Two drawbacks arrives from this supposition: (i) the proposed model is defined in $\mathbb{R}$ instead of on $\mathbb{R}_+$ and (ii) it is not resulting from the multiplicative model (as, for instance, $\mathcal{K}$ and $\mathcal{G}^0$ classic distribuitions) based on SAR image physical formation. With respect the first issue, we show that the model proposed under restriction is quasi positive (we assume this notation means “a real variable whose probability of intervals on $\mathbb{R}_-$ is negligible"). About (ii), although the CTPC model is not defined as product of two independent random variables (where one represents the speckle noise and other describes the relief, known multiplicative model), this model is linked to the fact that the signal of a resolution cell is influenced additively by several reflectors whose number of them varies of cell by cell. Note that

$$||F||^2 = \sum_{l=1}^{N} F_l \bar{F}_j + 2\,\mathrm{Re}[\sum_{l<j} F_l \bar{F}_j],$$

where $||\cdot||$ represents the square norm of a complex argument. Here, the second term is assumed to be null and, consequently, $||F||^2 = \sum_{l=1}^{N} F_l \bar{F}_j$. Thus, Eq. (4.2) can be used for describing SAR intensities. On the other hand, our model can assume uni and multimodals behaviors, unlike of models obtained from the multiplicative approach. This gain may be very important before SAR scens having many kind of textures.

The goal of this chapter is three-fold. First, we aim at analytically derive the CTPC distribution. Then, two of its properties are derived and discussed: cf and a distance measure between cfs.

Moreover, we provide two estimation procedures for CTPC parameters: MLE and QDE. Second, we furnish a CTPC GoF measure. A Monte Carlo study is made to quantify the performance of the estimators in terms of some figures of merit. Results indicate QDEs may be more accurate than MLEs. Third, we present evidence that the CTPC model may assume positive and multimodal nature being able to describe intensities distributed over SAR images with distinct clusters. Then an experiment with actual SAR data is performed. From this study, the proposed model is compared against six classic distributions and one mixed distribution for SAR intensities: Weibull, gamma, generalized gamma (GΓ) [181], $\mathcal{K}$ [182, 183], $\mathcal{G}^0$ [123], beta generalized normal (BGN) distribution [51], and Compound truncated Poisson Gamma (CTPG) [184]. Results are favorable to the introduced model.

This chapter is outlined as follows. In Section 4.2, we make a brief review of Compound N family. In Section 4.3 we define the CTPC model, present some of its properties, we provide the CTPC cf and a GoF measure based on empirical cf. Two estimation procedures for CTPC parameters are developed in Section 4.4. In Section 6.4, we make a simulation study to assess the proposed tools. Experiments with real SAR data are provided in Section 4.6. Finally, we furnish concluding remarks in Section 4.7.

4.2 Compound N family

Karlis and Xekalaki [30] have addressed an important distribution generator called compound N family, which is defined by the sum (4.1). Now, we detail this family briefly. Let $N \in \mathbb{Z}_+$ be a discrete random variable and $X_1, X_2, \ldots, X_N$ be a random sample obtained from X. Consider $N, X_1, X_2, \ldots, X_N$ be mutually independent. Thence, the cdf of (4.1) is given by

$$\begin{aligned} F_S(x) = P(S \leq x) = P\left(\sum_{h=1}^{N} X_h \leq x\right) &= \sum_{r=1}^{\infty} P(N=r)\, P\left(\sum_{h=1}^{N} X_h \leq x \,|\, N=r\right) \\ &= \sum_{r=1}^{\infty} P(N=r)\, P\left(\sum_{h=1}^{r} X_h \leq x\right) \\ &= \sum_{r=1}^{\infty} P(N=r)\, F_{S_r}(x), \end{aligned} \tag{4.3}$$

where $F_{S_r}(x)$ is the cdf of the sum of r independent random variables with distribution X, say S_r. Its pdf follows from differentiating (4.3):

$$f_S(x) = \sum_{r=1}^{\infty} P(N = r) f_{S_r}(x),$$

where f_{S_r} is the pdf of the sum of r independent random variables with distribution X.

Additionally, the hazard rate function (hrf) of the CTPC distribution is

$$\begin{aligned} h_S(x) &= \frac{f_S(x)}{1 - F_S(x)} = \frac{\sum_{k=1}^{\infty} P(N = k)\,(1 - F_{S_k}(x))\, h_{S_k}(x)}{\sum_{k=1}^{\infty} P(N = k)\,(1 - F_{S_k}(x))} \\ &= \sum_{k=1}^{\infty} w_k(x)\, h_{S_k}(x), \end{aligned}$$

where

$$w_k(x) = \frac{P(N = k)\,(1 - F_{S_k}(x))}{\sum_{k=1}^{\infty} P(N = k)\,(1 - F_{S_k}(x))},$$

such that $\sum_{k=1}^{\infty} w_k(x) = 1$. Thus, one has that the CTPC hrf is a mixture of Cauchy sum hrfs.

Fig. 4.1 show CTPC pdf plot. It is noticeable that the proposed model can describe multimodal and positive data.

Moreover, hrf may be of three different forms: increasing, decreasing and inverted bathtub.

The Fourier transform [J. B. J. Fourier, 1768–1830] [24, p. 32] is one of the most important tools in signal processing. Traditionally, a probability distribution can be characterized by means of its cf, which is the Fourier transform of its pdf. Let X be random variable following a probability distribution equipped with pdf $f(x)$ and cdf $F(x)$. Then, its cf, $\Phi_X(t)$, is defined as [98, p. 342]:

$$\Phi_X(t) \coloneqq \mathrm{E}(\mathrm{e}^{itX}) = \int_{-\infty}^{\infty} \mathrm{e}^{itx}\, \mathrm{d}F(x), \quad \text{for} \quad t \in \mathbb{R}, \tag{4.4}$$

where $i = \sqrt{-1}$ and $\mathrm{E}(\bullet)$ is the expectation operator. The cf of a compound N family can be determined according to the following theorem [185, Theorem 17.10.1].

Theorem 3 *Let N be a discrete random variable over positive integer having cf $\Phi_N(t)$ and $S = \sum_{r=1}^{N} X_r$ such that X_i cf $\Phi_X(t)$ for all $i = 1,\ldots,N$ and $X_1, X_2, \ldots, X_N, N$ be mutually independent. Then the following result holds*

$$\Phi_S(t) = \Phi_N(-i \log \Phi_X(t)).$$

Based on the cf, the nth moment of the random variable S is given as follows. The nth moment of S is the ratio between the nth derivative of cf, evaluated at the zero point, and the nth power of the imaginary number; i.e.,

$$\mathrm{E}(S^n) = \frac{\Phi_S^{(n)}(0)}{i^n}, \qquad n \in \mathbb{N}.$$

In particular, the two first moments are given by

$$\begin{aligned}
\mathrm{E}(S) &= \frac{\Phi_S(0)}{i} = \frac{1}{i}\left[\frac{\mathrm{d}}{\mathrm{d}t}\Phi_S(t)\right]_{t=0} \\
&= \left[\frac{1}{i}\Phi_N^{(1)}(-i\log\Phi_X(t))\left(\frac{-i\Phi_X^{(1)}(t)}{\Phi_X(t)}\right)\right]_{t=0} \\
&= i\,\Phi_N^{(1)}(0)\,\Phi_X^{(1)}(0) \\
&= \mathrm{E}(N)\,\mathrm{E}(X)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}(S^2) &= \frac{1}{i^2}\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi_S(t)\right]_{t=0} \\
&= \frac{\Phi_N^{(2)}(0)}{i^2}[-i\Phi_X^{(1)}(0)]^2 + (-i)\Phi_N^{(1)}(0)\left\{\frac{\Phi_X^{(2)}(0) - [\Phi_X^{(1)}(0)]^2}{i^2}\right\} \\
&= \mathrm{E}(N^2)[\mathrm{E}(X)]^2 + \mathrm{E}(N)\,\mathrm{Var}(X),
\end{aligned}$$

where $\mathrm{Var}(X)$ is the variance operator.

Another way to characterize a random variable is through the moment generating function (mgf). This function is also known as the Laplace transform [P. S. Laplace, 1749–1827] [24, p. 36], particularly in non-probabilistic contexts [98, p. 278].

Let X be a random variable such that $\mathrm{E}(\mathrm{e}^{tX}) < \infty$ [186, p. 194], then the mgf of X is defined as

$$\mathrm{M}_X(t) \coloneqq \mathrm{E}(\mathrm{e}^{tX}) = \int_{-\infty}^{\infty} \mathrm{e}^{tx}\,\mathrm{d}F(x) \quad \text{for} \quad t \in \mathbb{R}.$$

Thereby, the mgf of the compound N family is given by

$$\begin{aligned}\mathrm{M}_S(t) &= \int_{-\infty}^{\infty} \mathrm{e}^{t\,x} f_S(x)\,\mathrm{d}(x) \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{t\,x} \sum_{r=1}^{\infty} P(N=r) f_{S_r}(x)\,\mathrm{d}(x) \\ &= \sum_{r=1}^{\infty} P(N=r) \int_{-\infty}^{\infty} \mathrm{e}^{t\,x} f_{S_r}(x)\,\mathrm{d}(x).\end{aligned}$$

As mentioned before, note that f_{S_r} is the pdf of the sum of r independent random variables with distribution X. Thence, the mgf can be given as [185]

$$\mathrm{M}_S(t) = \sum_{r=1}^{\infty} P(N=r)[\mathrm{M}_X(t)]^r = \mathrm{M}_N\{\log[\mathrm{M}_X(t)]\},$$

where $\mathrm{M}_N(\bullet)$ and $\mathrm{M}_X(\bullet)$ are the mgfs of the N and X random variables, respectively. As a consequence, the cumulant generating function (cgf) [98, p. 147]

$$\mathrm{K}_S(t) = \log[\mathrm{M}_S(t)] = \log\{\mathrm{M}_N[\log \mathrm{M}_X(t)]\} = \mathrm{K}_N[\mathrm{K}_X(t))].$$

Likewise that we obtain the expected value by means of the cf, one can do through mgf:

$$\mathrm{E}(S^n) = \frac{\mathrm{d}^n \mathrm{M}_S(t)}{\mathrm{d}\,t^n}\bigg|_{t=0}.$$

Supposing mgf exists and is known, then $\Phi_S(t) = \mathrm{M}_S(it)$. Thus, all properties of the mgf can be extended to the cf. The main gain of the cf with respect to mgf is that it always exists for any $t \in \mathbb{R}$ [187, p. 153].

To illustrate the previous discussion of this section, consider (4.1) with $N \sim \mathrm{TP}(\lambda)$ having cf $\Phi_N(t) = \frac{\mathrm{e}^{\lambda \mathrm{e}^{it}} - 1}{\mathrm{e}^{\lambda} - 1}$ [188]. Using $X_i \sim \mathcal{N}(\mu, \sigma^2)$, Belchior *et al.* [189] proposed $S \sim \mathrm{CTPN}(\lambda, \mu, \sigma^2)$; and Weverson *et al.* [189] introduced $S \sim \mathrm{CTPG}(\lambda, \alpha, \beta)$ assuming $X_i \sim \Gamma(\alpha, \gamma)$. Their cdf and pdf are given by, respectively,

$$F_S(x) = \begin{cases} \left(\dfrac{1}{\mathrm{e}^{\lambda} - 1}\right) \displaystyle\sum_{r=1}^{\infty} \frac{\lambda^r}{r!} \Phi\left(\frac{x - r\mu}{\sigma\sqrt{r}}\right), & \mathrm{CTPN}, \\ \left(\dfrac{1}{\mathrm{e}^{\lambda} - 1}\right) \displaystyle\sum_{r=1}^{\infty} \frac{\lambda^r}{r!} \frac{\gamma(r\alpha, x\beta)}{\Gamma(r\alpha)}, & \mathrm{CTPG}, \end{cases}$$

and

$$f_S(x) = \begin{cases} \left(\dfrac{1}{e^{\lambda}-1}\right) \displaystyle\sum_{r=1}^{\infty} \dfrac{\lambda^r}{r!} \dfrac{1}{\sigma\sqrt{2r\pi}} \exp\left[-\dfrac{1}{2}\dfrac{(x-r\mu)^2}{r\sigma^2}\right], & \text{CTPN}, \\ \left(\dfrac{1}{e^{\lambda}-1}\right) \displaystyle\sum_{r=1}^{\infty} \dfrac{(\lambda\beta^{\alpha})^r}{r!} x^{r\alpha-1} e^{-\beta x}, & \text{CTPG}, \end{cases}$$

where $\Phi(\bullet)$ is the cdf of the $\mathcal{N}(0,1)$, $\lambda > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$, $\alpha > 0$, and $\beta > 0$. More details and properties are given in [184, 189].

4.3 The CTPC distribution

Let $X_1, X_2, \ldots, X_k$ a random variable following a Cauchy distribution, called $\mathcal{C}(\mu,\gamma)$, then the cdf and pdf of S_k are, respectively, given by:

$$F_{S_k}(x;\mu,\gamma) = \left[\frac{1}{\pi}\arctan\left(\frac{x-k\mu}{k\gamma}\right) + \frac{1}{2}\right], \tag{4.5}$$

and

$$f_{S_k}(x;\mu,\gamma) = \left\{\pi k\gamma\left[1+\left(\frac{x-k\mu}{k\gamma}\right)^2\right]\right\}^{-1},$$

where $x \in \mathbb{R}$ and $\mu \in \mathbb{R}$ and $\gamma > 0$ are the location and scale parameters, respectively. Let us further assume that N obeys the truncated Poisson distribution, denoted as $N \sim \text{TP}(\lambda)$. Thus we have [190]:

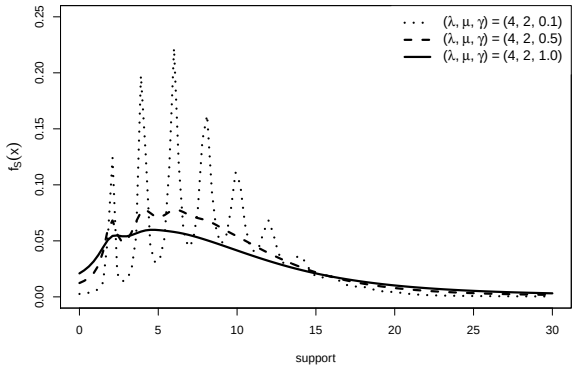
$$P(N=n) = \frac{1}{e^{\lambda}-1}\frac{\lambda^n}{n!}, \quad \text{for} \quad n = 1,2,\ldots. \tag{4.6}$$

By applying (4.5) and (4.6) in (4.3), it follows that:

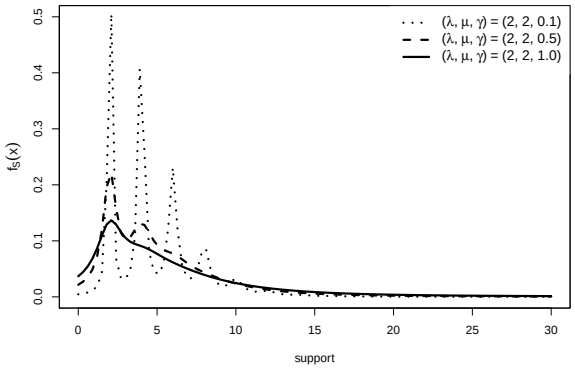
$$F_S(x;\boldsymbol{\theta}) = \frac{1}{e^{\lambda}-1}\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}\left[\frac{1}{\pi}\arctan\left(\frac{x-k\mu}{k\gamma}\right) + \frac{1}{2}\right],$$

where $\boldsymbol{\theta} = (\lambda,\mu,\gamma)^{\top}$ and the associated pdf is given by:

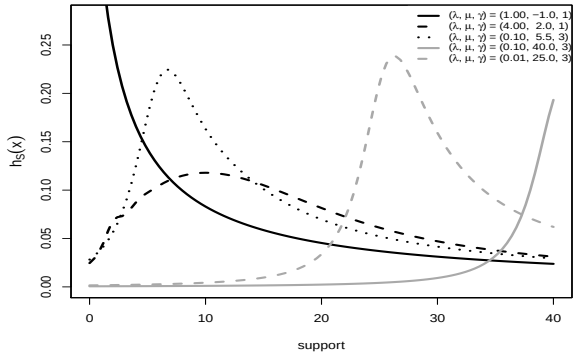
$$f_S(x;\boldsymbol{\theta}) = \frac{1}{e^{\lambda}-1}\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}\left\{\pi k\gamma\left[1+\left(\frac{x-k\mu}{k\gamma}\right)^2\right]\right\}^{-1}. \tag{4.7}$$



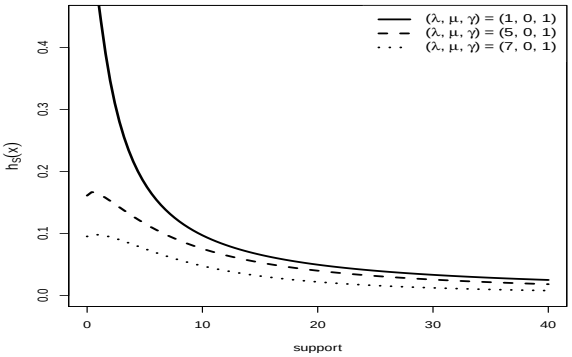
(a) PDF plot



(b) PDF plot



(c) HRF plot



(d) HRF plot

Figure 4.1: Plots of the pdf and hrf of CTPC model for several parameter values.

In this work, the above probability law is referred to as the *Compound Poisson-truncated Cauchy* distribution and it is denoted by $S \sim \text{CTPC}(\lambda, \mu, \gamma)$.

Fig. 4.1 show CTPC pdf and hrf plot. It is noticeable that the proposed model can describe multimodal and positive data. Moreover, hrf may be of three different forms: increasing, decreasing and inverted bathtub. In what follows, we derive some results for the CTPC model.

Initially, we have that the cf of N and X are, respectively, given by [189]:

$$\Phi_N(t) = \frac{e^{\lambda\, e^{it}} - 1}{e^{\lambda} - 1} \quad \text{and} \quad \Phi_X(t) = \exp\left(\mu i t - \gamma|t|\right),$$

Applying the above Theorem 3 we have that the CTPC cf is given by:

$$\Phi_S(t) = \frac{\mathrm{e}^{\lambda \exp\{\mathrm{i}[-\mathrm{i}(\mathrm{i}\mu t - \gamma|t|)]\}} - 1}{\mathrm{e}^{\lambda} - 1} = \frac{\mathrm{e}^{\lambda \exp\{\mathrm{i}\mu t - \gamma|t|\}} - 1}{\mathrm{e}^{\lambda} - 1}. \tag{4.8}$$

Invoking the Euler's identity, we can rewrite (4.8) in the following form:

$$\Phi_S(t) = \frac{\Lambda_t \left[\cos(\delta_t) + \mathrm{i}\sin(\delta_t)\right] - 1}{\mathrm{e}^{\lambda} - 1}, \tag{4.9}$$

where $\Lambda_t = \mathrm{e}^{\lambda \cos(\mu t)\, \mathrm{e}^{-\gamma|t|}}$ and $\delta_t = \lambda \sin(\mu t)\, \mathrm{e}^{-\gamma|t|}$. Thus,

$$\mathrm{Re}\left[\Phi_S(t)\right] = \frac{1}{\mathrm{e}^{\lambda} - 1}\left(\Lambda_t \cos(\delta_t) - 1\right) \quad \text{and} \quad \mathrm{Im}\left[\Phi_S(t)\right] = \frac{1}{\mathrm{e}^{\lambda} - 1}\Lambda_t \sin(\delta_t).$$

Additionally, the second type cf or cumulant generate function of S_N is given by

$$\Psi_S(t) = \log\left(\Phi_S(t)\right) = \log\left\{\frac{\mathrm{e}^{\lambda \exp[\mathrm{i}\mu t - \gamma|t|]} - 1}{\mathrm{e}^{\lambda} - 1}\right\}.$$

Expression (4.9) can be used for various goals. In this chapter, we propose two estimation criteria based on it. In what follows, we discuss the derivation of these methods as well as the maximum likelihood procedure for the CTPC model.

4.3.1 *Empirical cf method*

Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from X with cdf $F(x;\boldsymbol{\theta})$. The associated empirical cf (ecf) is defined as

$$\Phi_{X_n}(t) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}tX_j} = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\, dF_n(x;\boldsymbol{\theta}), \tag{4.10}$$

where $F_n(x;\boldsymbol{\theta})$ is the empirical cdf associated with $F(x;\boldsymbol{\theta})$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^{\top} \in \mathbb{R}^p$. Based on (4.10), several inferential methods have been proposed. In particular, Groparu and Doray [191] proposed a new hypothesis test and a point estimation procedure in terms of (4.10).

Assuming $t \in \mathbb{R}$ in (4.10), Press *et al.* [192] and Paulson *et al.* [193] introduced an estimator which minimizes the following distance:

$$D_{\mathrm{v}}(\boldsymbol{\theta};\boldsymbol{x}) = \int_{\mathbb{R}} \left|\Phi_{X_n}(t) - \Phi_X(t)\right|^{\mathrm{v}}\, dG(t,\kappa), \tag{4.11}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top$ and $G(t, \kappa)$ is a continuous weighting function.

As an example of (4.11) with $\nu = 2$, Knight and Yu [194] introduced an estimator based on the distance

$$D_2(\boldsymbol{\theta}; \boldsymbol{x}) = \int_{\mathbb{R}} |\Phi_{X_n}(t) - \Phi_X(t)|^2 \, dG(t, \kappa). \tag{4.12}$$

Further, Besbeas and Morgan [195] and Matsui and Takemura [196] have used successfully the weight function $G(t, \kappa) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa|t|} \, \mathrm{dt}$ for $\kappa > 0$, which renders the integral in (4.12) finite. In general, these estimators are called *ecf* estimators (ECFE) and are defined as

$$\widehat{\boldsymbol{\theta}}_{\text{ECFE}} = \underset{\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p}{\arg\min} \, D_2(\boldsymbol{\theta}; \boldsymbol{x}),$$

which obeys the asymptotic behavior [162]

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}^{-1} \boldsymbol{\Omega} \boldsymbol{\Sigma}^{-1}),$$

where "$\xrightarrow{\mathcal{D}}$" indicates *convergence in distribution*,

$$\boldsymbol{\Sigma} = \mathrm{E}\left(\frac{\mathrm{d}^2 D_2(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}^\top \, \mathrm{d}\boldsymbol{\theta}} \right) \quad \text{and} \quad \Omega = \mathrm{E}\left(\frac{\mathrm{d} D_2(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}^\top} \frac{\mathrm{d} D_2(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}} \right).$$

In this chapter, we aim to derive a distance-based GoF criterion for the CTPC model. First, consider the following result proposed by Yu [197].

Lemma 1 *Let $X \sim \mathcal{C}(\mu, \gamma)$ with cf given by*

$$\Phi_X(t) = \mathrm{e}^{it\mu - \gamma|t|}, \quad \textit{for} \quad t \in \mathbb{R}.$$

Assuming $G(t, \kappa) = \mathrm{e}^{-\kappa|t|}$ and $\kappa > 0$, the distance (4.12) *is given by*

$$D_C(\boldsymbol{\theta}; \boldsymbol{x}) = \frac{2}{n^2} \sum_{j,l=1}^{n} \frac{\kappa}{\kappa^2 + \left(x_j - x_l\right)^2} - \frac{4}{n} \sum_{j=1}^{n} \frac{\kappa + \gamma}{(\kappa + \gamma)^2 + \left(x_j - \mu\right)^2} + \frac{2}{2\gamma + \kappa},$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top$ is an outcome of $(X_1, X_2, \ldots, X_n)^\top$.

The proof this lemma is in [195, 196].

Taking into account the previous lemma, we propose the subsequent proposition.

Proposition 10 *Let $X \sim \text{CTPC}(\lambda,\mu,\gamma)$ with cf given in* (4.8) *and $G(t,\kappa) = \mathrm{e}^{-\kappa|t|}$ for $\kappa > 0$. In this case, the distance* (4.12) *is given by*

$$\begin{aligned} D_{CTPC}(\boldsymbol{\theta};\boldsymbol{x}) = \frac{2}{n^2}\sum_{j,l=1}^{n}\frac{\kappa}{\kappa^2+(x_j-x_l)^2} + \frac{4}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n}\frac{\kappa}{\kappa^2+x_j^2} \\ - \frac{2}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n} g_3(\lambda,\gamma,\mu,\kappa,x_j) + \frac{2}{\kappa(\mathrm{e}^{\lambda}-1)^2} \\ + \frac{1}{(\mathrm{e}^{\lambda}-1)^2} g_1(\lambda,\gamma,\mu,\kappa) - \frac{2}{(\mathrm{e}^{\lambda}-1)^2} g_2(\lambda,\gamma,\mu,\kappa), \end{aligned}$$

where

$$g_1(\boldsymbol{\theta};\kappa) = \int_{\mathbb{R}} \mathrm{e}^{2\lambda\cos(\mu t)\,\mathrm{e}^{-\gamma|t|}}\,\mathrm{e}^{-\kappa|t|}\,dt, \quad g_2(\boldsymbol{\theta};\kappa) = \int_{\mathbb{R}} \Lambda_t\cos(\delta_t)\,\mathrm{e}^{-\kappa|t|}\,dt,$$

and

$$g_3(\boldsymbol{\theta};\kappa,x_j) = \int_{\mathbb{R}} \cos(\delta_t - t x_j)\,\Lambda_t\,\mathrm{e}^{-\kappa|t|}\,dt.$$

■

The proof of this proposition is given in Appendix D. From this proposition, the following corollary holds.

Corollary 1 *Let $X \sim \text{CTPC}(\lambda,0,1)$. The GoF measure* (4.12) *is given by*

$$\begin{aligned} D_{CTPC}(\boldsymbol{\theta};\boldsymbol{x}) = \frac{2}{n^2}\sum_{j,l=1}^{n}\frac{\kappa}{\kappa^2+(x_j-x_l)^2} + \frac{4}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n}\frac{\kappa}{\kappa^2+x_j^2} \\ - \frac{2}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n}\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}\frac{(k+r)}{x_j^2+(k+r)^2} + \frac{2}{\kappa(\mathrm{e}^{\lambda}-1)^2} \\ + \frac{1}{(\mathrm{e}^{\lambda}-1)^2}\frac{2}{(-2\lambda)^{\kappa}}\left[\Gamma(\kappa)-\Gamma(\kappa,-2\lambda)\right] - \frac{4}{(\mathrm{e}^{\lambda}-1)^2(-\lambda)^{\kappa}}\left[\Gamma(\kappa)-\Gamma(\kappa,-\lambda)\right]. \end{aligned}$$

■

The above particular case is important to define an alternative measure for multimodal data to be compared with other obtained from one parameter standard models (such as exponential, Rayleigh and Nakagami distributions).

To check the behavior of $D_{\text{CTPC}}(\boldsymbol{\theta};\boldsymbol{x})$ at $\boldsymbol{\theta} = (\lambda,\mu,\gamma)$, where $\lambda \in \{0.1, 1.0, 5.0, 10.0\}$ and $\mu \in \{-10,-9,\ldots,10\}$, we follow the steps:

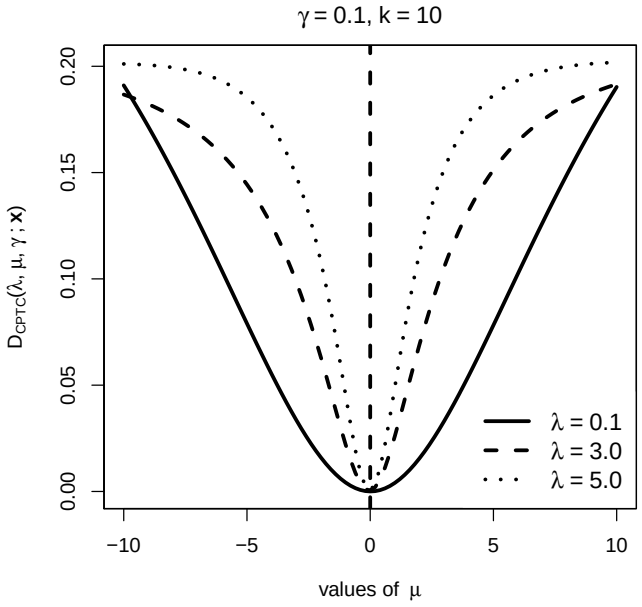


Figure 4.2: Illustration of GoF measure of CTPC model.

1. Given $\lambda = \lambda_0$, generate one CTPC$(\lambda_0, 0, 0.1)$ sample of size $n = 100$, say $\boldsymbol{x} = (x_1, \ldots, x_n)$;

2. At each Monte Carlo replication, quantify $D_{\text{CTPC}}((\lambda_0, \mu, 0.1); \boldsymbol{x})$ assuming $\kappa = 10$ for each $\mu \in \{-10, -9, \ldots, 10\}$, yielding the distance vector $d_{1,i}, \ldots, d_{100,i}$ for ith replication;

3. Return, for each λ, the vector $(d_1, d_2, \ldots, d_{100})$ such that $d_k = 100^{-1} \sum_{i=1}^{100} d_{i,k}$ and $\kappa = 1, \ldots, 10$.

It is expected that $D_{\text{CTPC}}(\lambda_0, 0, 1) \leq D_{\text{CTPC}}(\lambda_0, \mu, 0.1)$, where the lower bound must be as close to zero as possible. Fig. 4.2 illustrates the values of the distances obtained. One can observe the expected behaviour tends to be achieved for higher values of λ; i.e., this distance is sensitive to the value of the parameter λ. Although it is not the focus of this chapter, the measure proposed in Proposition 10 can be combined to non-parametric procedures to provide a GoF tools for the CTPC law.

In what follows, we furnish estimation criteria based on CTPC distances.

4.4 New Estimation and GoF criteria

This section aims to provide two estimation methods. One of them is the maximum likelihood procedure and the other two are based on the minimization of a distance measure between empirical and theoretical cfs. Moreover, the distance proposed in the previous section is understood as a new GoF criterion.

4.4.1 *Maximum Likelihood Method*

Let $S_1, S_2, \ldots, S_n$ be a random sample obtained from $S \sim \text{CTPC}(\lambda, \mu, \gamma)$ having pdf given in (4.7). The log-likelihood function at the parameter vector $\boldsymbol{\theta} = (\lambda, \mu, \gamma)^\top$ based on one outcome of that sample, say $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top$, is given by

$$\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}; \boldsymbol{x}) = \log \prod_{l=1}^{n} f(x_l; \lambda, \mu, \gamma) = n \log[(e^{\lambda} - 1)^{-1}] + \sum_{l=1}^{n} \log \left\{ \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} f_{S_k}(x_l) \right\}, \tag{4.13}$$

where $f_{S_k}(x_l) \equiv f_{S_k}(x_l; \mu, \gamma)$.

Based on (4.13), the associated score function,

$$\mathrm{U}(\boldsymbol{\theta}) = (\mathrm{U}_\lambda, \mathrm{U}_\mu, \mathrm{U}_\gamma) = \left(\frac{\mathrm{d}\, \ell(\boldsymbol{\theta})}{\mathrm{d}\, \lambda}, \frac{\mathrm{d}\, \ell(\boldsymbol{\theta})}{\mathrm{d}\, \mu} \frac{\mathrm{d}\, \ell(\boldsymbol{\theta})}{\mathrm{d}\, \gamma} \right)^\top,$$

is given by:

$$\mathrm{U}_\lambda(\boldsymbol{\theta}) = \left[\sum_{l=1}^{n} \frac{1}{f_S(x_l)(e^{\lambda} - 1)} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} f_{S_{k+1}}(x_l) \right] - n \frac{e^{\lambda}}{(e^{\lambda} - 1)},$$

$$\begin{aligned} \mathrm{U}_\mu(\boldsymbol{\theta}) &= \frac{1}{(e^{\lambda} - 1)} \sum_{l=1}^{n} \frac{1}{f_S(x_l)} \sum_{k=1}^{\infty} \left\{ \frac{\lambda^k}{k!} f_{S_k}(x_l) \frac{\mathrm{d} \log[f_{S_k}(x_l)]}{\mathrm{d}\, \mu} \right\} \\ &= \frac{2}{e^{\lambda} - 1} \sum_{l=1}^{n} \frac{1}{f_S(x_l)} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \left[\frac{k(x_l - k\mu)}{k^2\gamma^2 + (x_l - k\mu)^2} \right] f_{S_k}(x_l). \end{aligned}$$

$$\begin{aligned} \mathrm{U}_\gamma(\boldsymbol{\theta}) &= \frac{1}{(e^{\lambda} - 1)} \sum_{l=1}^{n} \frac{1}{f_S(x_l)} \sum_{k=1}^{\infty} \left\{ \frac{\lambda^k}{k!} f_{S_k}(x_l) \frac{\mathrm{d} \log[f_{S_k}(x_l)]}{\mathrm{d}\, \gamma} \right\} \\ &= \frac{2}{(e^{\lambda} - 1)\gamma} \sum_{l=1}^{n} \frac{1}{f_S(x_l)} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \left[\frac{y_{kl} - 1}{y_{kl} + 1} \right] f_{S_k}(x_l), \end{aligned}$$

where $y_{kl} = \left(\frac{x_l - k\mu}{k\gamma}\right)^2$.

The MLEs can be obtained from $\mathrm{U}_\lambda = \mathrm{U}_\mu = \mathrm{U}_\gamma = 0$. These equations do not yield closed-form estimates for λ, μ, and γ. Hence, we need to use iterative techniques; such as a BFGS (Broyden-Fletcher-Goldfarb-Shannon) method [159, 198].

4.4.2 *QDE based on cf*

Let $X_1, X_2, \ldots, X_n$ be a random sample obtained from $X \sim \mathrm{CTPC}(\lambda, \mu, \gamma)$ having cf $\Phi_S(t)$ given in (4.8). Thus, the associated empirical cf is defined by

$$\Phi_n(t_j) = \frac{1}{n}\sum_{l=1}^{n} \mathrm{e}^{\mathrm{i}t_j X_l},$$

for $j = 1, 2, \ldots, k$, where k is fixed and represents the number of points used to discretization of the support of the cf. Also, for the same set of points $t_1, t_2, \ldots, t_k$, let $\boldsymbol{v}(\boldsymbol{\theta})$ be the vector containing the real and imaginary parts of the cf, as follows:

$$\boldsymbol{v}(\boldsymbol{\theta}) = \left(\mathrm{Re}\left[\Phi(t_1)\right], \ldots, \mathrm{Re}\left[\Phi(t_k)\right], \mathrm{Im}\left[\Phi(t_1)\right], \ldots, \mathrm{Im}\left[\Phi(t_k)\right]\right)^\top,$$

and let $\boldsymbol{v}_n$ be the vector containing the real and imaginary parts of the empirical cf, as follows:

$$\boldsymbol{v}_n = \left(a_{t_1}, a_{t_2}, \ldots, a_{t_k}, b_{t_1}, b_{t_2}, \ldots, b_{t_k}\right)^\top,$$

where

$$a_{t_j} = \mathrm{Re}\left[\Phi_n(t_j)\right] = \frac{1}{n}\sum_{l=1}^{n} \cos(t_j X_l) \quad \text{and} \quad b_{t_j} = \mathrm{Im}\left[\Phi_n(t_j)\right] = \frac{1}{n}\sum_{l=1}^{n} \sin(t_j X_l), \quad j = 1, 2, \ldots, k.$$

We consider two QDEs which are defined as the argument in $\Theta \in R^p$ which minimizes the distances D_{CTPC_1} and D_{CTPC_2} defined by:

$$D_{\mathrm{CTPC}_1}(\boldsymbol{\theta}) = [v_n - v(\theta)]^\top I_{2k} [v_n - v(\theta)],$$

and

$$D_{\mathrm{CTPC}_2}(\boldsymbol{\theta}) = [v_n - v(\theta)]^\top \Sigma^{-1}(\theta) [v_n - v(\theta)],$$

where I_{2k} is the identity matrix with order $2k$,

$$\Sigma(\theta) = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{R}} & \mathbf{\Sigma}_{\mathrm{RI}} \\ \mathbf{\Sigma}_{\mathrm{IR}} & \mathbf{\Sigma}_{\mathrm{I}} \end{bmatrix},$$

$$\Sigma_R = n\,\mathrm{E}\left[\left(a_{t_l} - \mathrm{Re}\left[\Phi(t_l)\right]\right)\left(a_{t_j} - \mathrm{Re}\left[\Phi(t_j)\right]\right)\right]\Big|_{l,j=1}^{k},$$

$$\Sigma_I = n\,\mathrm{E}\left[\left(b_{t_l} - \mathrm{Im}\left[\Phi(t_l)\right]\right)\left(b_{t_j} - \mathrm{Im}\left[\Phi(t_j)\right]\right)\right]\Big|_{l,j=1}^{k},$$

and

$$\Sigma_{IR} = \Sigma_{RI} = n\,\mathrm{E}\left[\left(a_{t_l} - \mathrm{Re}\left[\Phi(t_l)\right]\right)\left(b_{t_j} - \mathrm{Im}[\Phi(t_j)]\right)\right]\Big|_{l,j=1}^{k}.$$

Before proposing our estimation criterion, it is important to determine guidelines to choose the values of t_l. To that end, Thaler [199] and Groparu [200] derived the following proposition.

Proposition 11 *Let $X_1,\ldots,X_n$ be a random sample drawn from X with cf $\Phi(t)$. Let $\Phi_n(t)$ be the ecf associated with $\Phi(t)$, then the following identities hold true:*

$$\begin{aligned}
\mathrm{E}\,(a_t, b_t) &= \{Re[\Phi(t)], Im[\Phi(t)]\}, \\
Cov\,(a_t, a_s) &= \frac{1}{2n}\Big\{Re[\Phi(t+s)] + Re[\Phi(t-s)] - 2Re[\Phi(t)]Re[\Phi(s)]\Big\}, \\
Cov\,(b_t, b_s) &= \frac{1}{2n}\Big\{Re[\Phi(t-s)] - Re[\Phi(t+s)] - 2Im[\Phi(t)]Im[\Phi(s)]\Big\}, \\
Cov\,(a_t, b_s) &= \frac{1}{2n}\Big\{Im[\Phi(t+s)] - Im[\Phi(t-s)] - 2Re[\Phi(t)]Im[\Phi(s)]\Big\}.
\end{aligned}$$

The proof of this proposition can be found in Thaler [199], Groparu [200], Groparu and Doray [191], and Luong and Doray [201, 202].

From Proposition 11, for $t = s$, we have $\mathrm{Re}[\Phi(0)] = 1$ and $\mathrm{Im}[\Phi(0)] = 0$; therefore, we have that:

$$\begin{aligned}
\mathrm{Var}\,(a_t) &= \frac{1}{2n}\{\mathrm{Re}[\Phi(2t)] + 1 - 2\,(\mathrm{Re}[\Phi(t)])^2\}, \\
\mathrm{Var}\,(b_t) &= \frac{1}{2n}\{1 - \mathrm{Re}[\Phi(2t)] - 2\,(\mathrm{Im}[\Phi(t)])^2\}, \\
\mathrm{Cov}\,(a_t, b_t) &= \frac{1}{2n}\{\mathrm{Im}[\Phi(2t)] - 2\mathrm{Re}[\Phi(t)]\mathrm{Im}[\Phi(t)]\}.
\end{aligned}$$

Assuming $t \to 0$, one can show that $\mathrm{Re}[\Phi_n(t)]$ and $\mathrm{Im}[\Phi_n(t)]$ are uncorrelated and converge for $\mathrm{Re}[\Phi(t)]$ and $\mathrm{Im}[\Phi(t)]$ for larger samples, respectively. Therefore, it is recommended to define $t_1, t_2, \ldots, t_k$ as close to zero as possible.

Theorem 4 *Let $X \sim \mathrm{CTPC}(\lambda, \mu, \gamma)$ be a random variable. The $D_{CTPC_1}(\boldsymbol{\theta})$ is given by*

$$D_{CTPC_1}(\boldsymbol{\theta}) = \sum_{j=1}^{k} \left[\left(\frac{1}{n} \sum_{l=1}^{n} \cos(t_j x_l) - \frac{C_j \cos(S_j) - 1}{\mathrm{e}^{\lambda} - 1} \right)^2 + \left(\frac{1}{n} \sum_{l=1}^{n} \sin(t_j x_l) - \frac{C_j \sin(S_j)}{\mathrm{e}^{\lambda} - 1} \right)^2 \right],$$

where

$$C_j = \exp\left\{ \lambda \cos(\mu t_j)\, \mathrm{e}^{-\gamma |t_j|} \right\}, \quad \textit{and} \quad S_j = \lambda \sin(\mu t_j)\, \mathrm{e}^{-\gamma |t_j|}.$$

■

From Theorem 4, one has the quadratic distance estimator (QDE) is given by

$$\widehat{\boldsymbol{\theta}}_{\mathrm{CTPC}_1} = \underset{\boldsymbol{\theta} \in \Theta}{\arg\min} \left[D_{\mathrm{CTPC}_1}(\boldsymbol{\theta};\, \boldsymbol{x}) \right].$$

Theorem 5 *Let $X \sim \mathrm{CTPC}(\lambda, \mu, \gamma)$ be a random variable. The $D_{CTPC_2}(\boldsymbol{\theta})$ is given by*

$$D_{CTPC_2}(\boldsymbol{\theta}) = \left[\boldsymbol{v}_n - \boldsymbol{v}(\boldsymbol{\theta}) \right]^{\top} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \left[\boldsymbol{v}_n - \boldsymbol{v}(\boldsymbol{\theta}) \right],$$

where

$$\boldsymbol{v}_n = \left[a_{t_1}, a_{t_2}, \ldots, a_{t_k}, b_{t_1}, b_{t_2}, \ldots, b_{t_k} \right],$$

$$\boldsymbol{v}(\boldsymbol{\theta}) = \frac{1}{\mathrm{e}^{\lambda} - 1} \left[C_1 \cos(S_1) - 1, \ldots, C_k \cos(S_k) - 1, C_1 \sin(S_1), \ldots, C_k \sin(S_k) \right],$$

and the $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ matrix is given in Appendix E in the form of matrix (6)*.*

■

Thus, a second estimator based on D_{CTPC_2}, denoted by VQDE (The QDE using the variance matrix given in (6)), is defined by

$$\widehat{\boldsymbol{\theta}}_{\mathrm{CTPC}_2} = \underset{\boldsymbol{\theta} \in \Theta}{\arg\min} \left[D_{\mathrm{CTPC}_2}(\boldsymbol{\theta};\, \boldsymbol{x}) \right].$$

4.4.3 *GoF Measure*

The Distance measure proposed in Proposition 10 may be understood as a discrepancy measure between a supposed probabilistic model and an empirical distribution drawn from data by means of cf. We want to use such distance measure to propose adherence test in terms of the CTPC cf, $\mathcal{H}_0 : X \sim F_{\text{CTPC}}$. To this end, it is necessary to define a pivotal statistics.

Baringhaus and Henze [203] proposed to solve this problem through the statistic

$$T_{n,G}(\widehat{\boldsymbol{\theta}}) = n \int |\Phi_{X_n}(t) - \Phi_X(t, \widehat{\boldsymbol{\theta}})|^2 \mathrm{d}G(t, \kappa), \tag{4.14}$$

where $\Phi_X(t, \widehat{\boldsymbol{\theta}})$ is the CTPC cf, $\Phi_{X_n}(t)$ is the empirical cf (both given in (4.8) and (4.10)), $G(t, \kappa)$ is the weight function given in Eq. (4.4), and $\widehat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta}$. Jiménez-Gamero *et al.* [204] studied some properties of (4.14). Two of them are (i) $T_{n,\widehat{G}}(\widehat{\boldsymbol{\theta}})$ and $T_{n,G}(\widehat{\boldsymbol{\theta}})$, where $\widehat{G}$ is an estimator of G, have similar expansions and (ii) if $\widehat{\boldsymbol{\theta}}$ is strongly consistent for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $T_{n,G}(\widehat{\boldsymbol{\theta}}) \longrightarrow \infty$ almost surely when $\mathcal{H}_0$ does not hold. Therefore, as general rule, the test equipped by (4.14) that rejects $\mathcal{H}_0$ for large values of $T_{n,G}(\widehat{\boldsymbol{\theta}})$ is strongly consistent.

The asymptotic null distribution of $T_{n,G}(\widehat{\boldsymbol{\theta}})$, supposing that $\mathcal{H}_0$ is true, is given by

$$T_{n,G}(\widehat{\boldsymbol{\theta}}) \xrightarrow[n\to\infty]{\mathcal{D}} \sum_{j=1}^{\infty} \zeta_j \chi^2_{1j},$$

where $\chi^2_{11}, \chi^2_{12}, \ldots,$ are independent variables having chi-square distributions with one degree of freedom, the set $\{\zeta_j\}$ represent the eigenvalues of an operator in terms of $\boldsymbol{\theta}$ according to Jiménez-Gamero *et al.* [204].

In practice, it is hard to determine the asymptotic distribution of $T_{n,G}(\widehat{\boldsymbol{\theta}})$. A way to work with this statistic is to use the bootstrap procedure. In what follows, we outline how to use parametric bootstrap to approximate the null distribution of $T_{n,G}(\widehat{\boldsymbol{\theta}})$.

Let us first define the bootstrap distribution under the null hypothesis of $T_{n,G}(\widehat{\boldsymbol{\theta}})$. Let T^* be the parametric bootstrap version of $T_{n,G}(\widehat{\boldsymbol{\theta}})$. Thus, for $\tau^*_{n,\alpha} = \inf\{\tau : P(T^* \geq \tau) \leq \alpha\}$ and $p^* = P(T^* \geq T_{obs})$, the test is defined by (for a nominal level α)

$$\phi^* = \begin{cases} 1, & \text{if } T^* \geq \tau^*_{n,\alpha} \text{ or } p^* \leq \alpha, \\ 0, & \text{otherwise.} \end{cases}$$

In practice, this bootstrap test is executed through the following algorithm: Consider the problem to check if a data array $\{x_i\}_{1\leq i\leq n}$ can be understood as an outcome of a random sample from the CTPC model.

1. Estimate the CTPC parameters, say $\widehat{\boldsymbol{\theta}}$, by a consistent method (as, for instance, by maximum likelihood).

2. Based on the estimates of the previous item, calculate $T_{n,G}(\widehat{\boldsymbol{\theta}})$, using Equation of the distance given in Proposition 10, say T_{obs}.

3. Generate B bootstrap samples $\{x_i^{(k)}\}$, for $1 \leq i \leq n$, $1 \leq k \leq B$, from $F_{\text{CTPC}}(x; \widehat{\boldsymbol{\theta}})$.

4. Calculate $T_{n,G}(\widehat{\boldsymbol{\theta}})$ for each bootstrap sample and denote the resultant value by T_{k}^*, for $k = 1,2,\ldots,B$.

5. Approximate the p-value by means of the expression
$$\widehat{p}_{\text{boot}} = \frac{\sum_{k=1}^{B} \mathbb{I}(T_{\text{k}}^* \geq T_{\text{obs}})}{B},$$
in which $\mathbb{I}(\omega)$ is the indicator variable that is equal to 1 if the event ω is true or to zero otherwise. The boostrap test rejection rule is
$$\text{Reject } \mathcal{H}_0 \text{ if } \widehat{p}_{\text{boot}} \leq \alpha,$$
with α denoting the required significance level. An equivalent way to define such decision rule is by the bootstrap critical value $T_{\text{a:B}}^*$, where $a = [(1-\alpha)B] + 1$, $[x]$ is the greatest integer less than x, and $T_{\text{1:B}}^* \leq T_{\text{2:B}}^* \leq \cdots \leq T_{\text{B:B}}^*$ are the values T_{k}^*, in increasing order. In this case, the rejection rule is
$$\text{Reject } \mathcal{H}_0 \text{ if } T_{\text{obs}} > T_{\text{a:B}}.$$

4.5 Synthetic studies

4.5.1 *Performance of the estimators*

We carry out a simulation study in order to compare the performance of the proposed estimation methods. To that end, we consider several combinations of parameters and $\lambda \in \{0.5, 1.0\}$, $\mu \in$

$\{-2,-1,0,0.5,1.5\}$, $\gamma = 8$ and sample sizes $n \in \{10,200\}$. For each combination (λ, μ, γ, n), we generate 1,000 Monte Carlo replications on which proposed estimators are assessed for several sample sizes.

For the QDE and VQDE methods, we use two kinds of $\boldsymbol{t} = (t_1, t_2, \ldots, t_k)$:

(a) for $\boldsymbol{t} = (0.1, 0.2, \ldots, 1.0)$ with associated estimators denoted by QDE_1 and VQDE_1 and

(b) for $\boldsymbol{t} = (0.1, 0.2, \ldots, 5.0)$, denoted as QDE_2 and VQDE_2.

As one merit figure, we use the sample root mean squared error by parameter (RMSEp) $\text{RMSE}(\widehat{\theta}_i) = \sqrt{1/n \sum_{k=1}^{1000} (\widehat{\theta}_{ik} - \theta_i)^2}$ for $i = 1,2,3$, where $\widehat{\theta} \in (\widehat{\lambda}, \widehat{\mu}, \widehat{\gamma})^\top$ and $\theta_i \in (\lambda, \mu, \gamma)^\top$. We also use the sample root mean squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{1000} \sum_{k=1}^{1000} E_k},$$

where $E_k = (\widehat{\lambda}_k - \lambda_k)^2 + (\widehat{\mu}_k - \mu_k)^2 + (\widehat{\gamma}_k - \gamma_k)^2$. Better results are associated to smaller values of RMSEp and RMSE.

Numerical results are shown in Fig.s 4.3 and 4.4, where the numbers $\{1,2,\ldots,8\}$ corresponds the parameter values equivalent the combination

$$\begin{aligned}(\lambda, \mu, \gamma) = &\{(0.5,-1,8),(0.5,0,8),(0.5,0.5,8),(0.5,1.5,8),(1.0,-2,8),(1.0,0,8),(1.0,0.5,8),\\ &(1.0,1.5,8)\}.\end{aligned}$$

We observe that, as expected, for all methods, the larger sample sizes are associated to smaller RMSEp values. In addition, we emphasize that for both small samples the VQDE_1 method outperforms the others and large samples the QDE_1. Further, with respect to the effect number of points in the vector $\boldsymbol{t}$ on QDE_i and VQDE_i, one can notice that its increase implies VQDE_2 and QDE_1 outperform VQDE_1 and QDE_2, respectively.

Finally, we analyze the impact of the QDE and VQDE methods of the distance between the points in the vector t. For that, we considered $t_\varepsilon = (0.1\varepsilon, 0.2\varepsilon, \ldots, 5\varepsilon)$, where $\varepsilon \in \{0.1, 0.5, 1, 2, 10\}$, $\lambda = 0.5$, $\gamma = 8$ and $\mu \in \{-1.0, 0.5, 1.5\}$. Fig. 4.5 displays the results. We note that VQDE outperforms QDE and that such difference vanishes as ε increases.

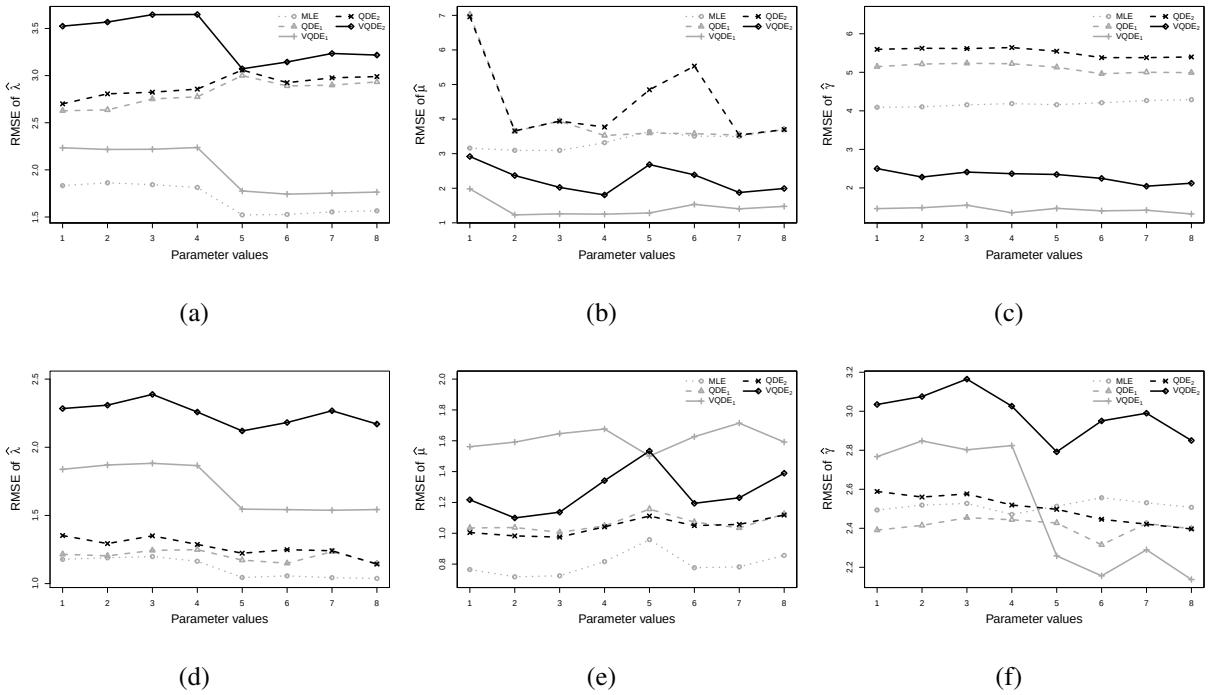


Figure 4.3: RMSEp for several parameter values. (a), (b), and (c) $n = 10$; and (d), (e), and (f) $n = 200$.

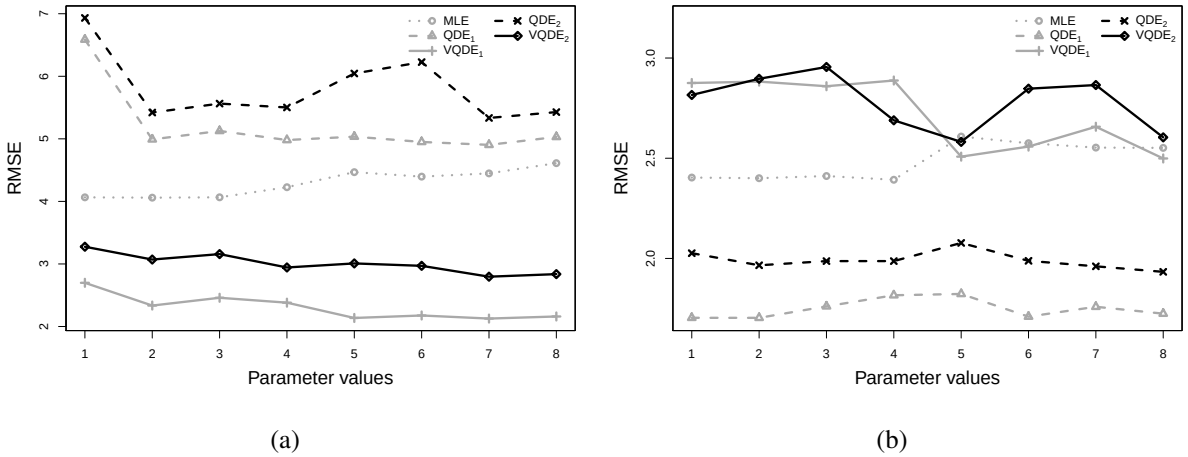


Figure 4.4: RMSE for several parameter values: (a) $n = 10$ and (b) $n = 200$.

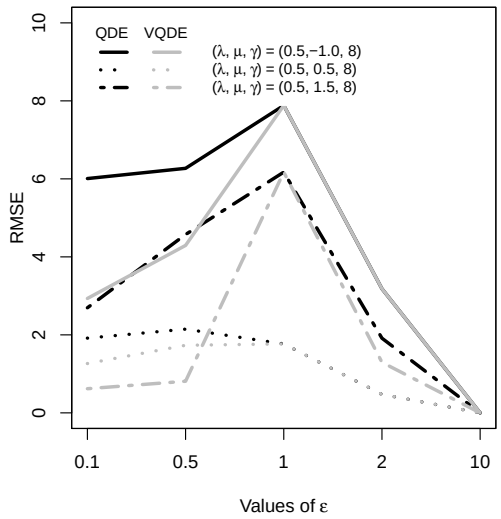


Figure 4.5: RMSE of the QDE and VQDE methods with respect to ε.

4.5.2 *Performance of the new GoF measure*

Now, the performance of $T_{n,G}(\widehat{\boldsymbol{\theta}})$ in (4.14) is assessed on synthetic data. The simulation study is driven by the algorithm below:

Algorithm 1: Estimated type I error for $T_{n,G}(\widehat{\boldsymbol{\theta}})$ at the CTPC model.

Step 1. Set $\boldsymbol{\theta} = (\lambda, \mu, \gamma) \in \{(1.3, -0.1, 1); (1.3, -0.1, 0.5); (\log(2), 0, 1);$ $(1, 0, 1); (1.3, 0, 1); (1.3, 0, 0.5); (\log(2), 0.1, 1); (1, 0.1, 1); (1.3, 0.1, 1); (1.3, 0.1, 0.5)\}$;

Step 2. Generate a sample $x_1, x_2, \ldots, x_n$ from the CTPC distribution at $\boldsymbol{\theta}$;

Step 3. Estimate $\boldsymbol{\theta}$ by MLE;

Step 4. Calculate $T_{n,G}(\widehat{\boldsymbol{\theta}})$;

Step 5. Design $1,000$ Monte Carlo replications and, on each one, 250 n-points bootstrap samples, $(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \ldots, \boldsymbol{x}_{250}^*)$, at sample sizes $n \in \{20, 50, 100\}$; **Step 6.** Calculate the Monte Carlo estimates for $\alpha \in \{0.01, 0.05, 0.10\}$.

Fig. 4.6 displays the estimated test results. In general, estimates are close to the adopted nominal levels and, in particular, better results are found in higher sample sizes. Significance levels

smaller than 10% are overestimated; while, they are underestimated for $\alpha = 10\%$.

4.6 Application

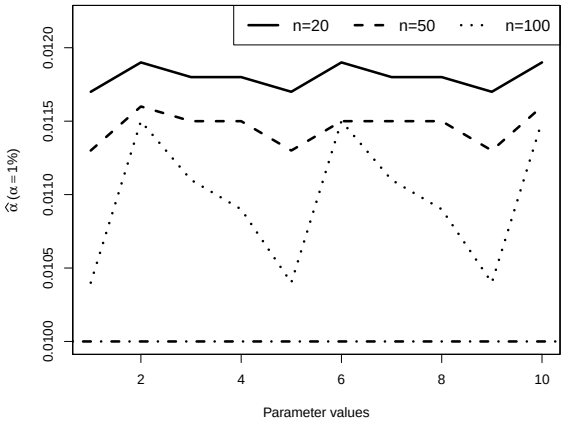
Intensities extracted from SAR images are important features for understanding SAR scenarios. These components can also be understood as the result of compounding sums as (4.1) [175]. In this chapter, we propose the CTPC model as a possible compound distribution, which is able to describe multimodal (among other behaviors, such as negative support, symmetric form, etc). In what follows, we aim to describe SAR intensity obtained by EMISAR (see [205]) sensor of scenes of Foulum (Denmark) regions.

SAR images are processed obeying the following dynamic: Polarized pulses oriented at horizontal (H) and vertical (V) directions are emitted towards to an under-study target. Subsequently, such pulses are also recorded at H and V directions. Thus, each SAR entry is equipped with informations due to three polarization channels, say S_{HH}, S_{HV}, $S_{\text{VV}} \in \mathbb{C}$, representing the Horizontal-Horizontal (HH), Horizontal-Vertical (HV) and Vertical-Vertical (VV) states, respectively. In practice, HV and VH states are similar and, as a consequence, one assumes that $S_{\text{HV}} = S_{\text{VH}}$.

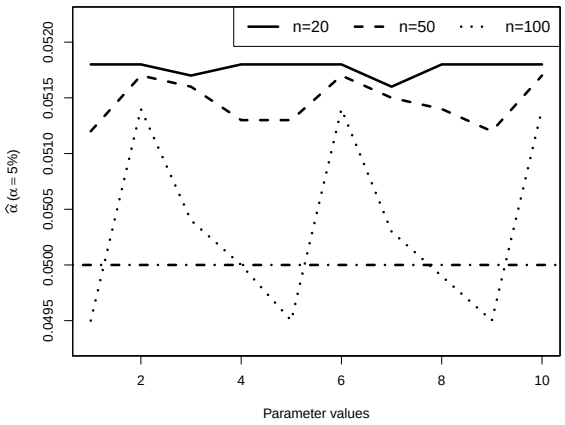
Fig. 4.7 shows an intensity map of Foulum SAR regions due to the HH and VV channel. First we provide a brief descriptive discussion in the Table 4.1. One can observe three behaviors for both channels: (a) "median $<$ mean" which indicates that the empirical density is right asymmetric (confirmed by the positive asymmetry coefficient), (b) the sample variation coefficients of 73.03% and 73.33%, respectively, indicate high variability data and (c) values of associated kurtosis are greater than three what suggest a leptokurtic empirical distributions; i.e., they are more concentrated than the Gaussian model.

Several distributions have been proposed to model SAR intensity data: the classical Γ and Weibull models, the GΓ law proposed by Stacy [181], the $\mathcal{K}$ and $\mathcal{G}^0$ models proposed by Jakeman and Pusey [182], Lee *et al.* [183] and Frery *et al.* [123] and the BGN distribution derived recently by Cintra *et al.* [51]. We compare the proposed model fit for SAR intensity data using the CTPC and those six models.

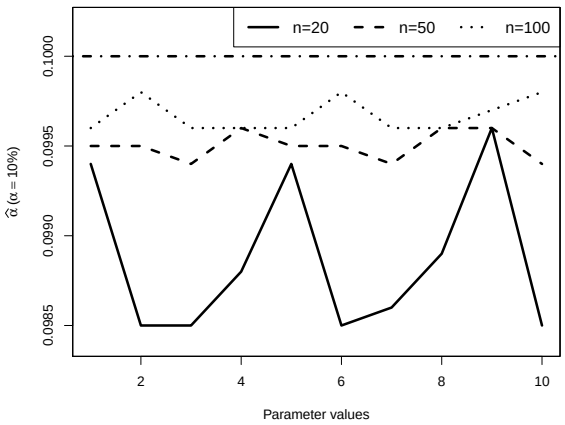
In order to check which distribution better fits the data from highlighted SAR region, we plot



(a)

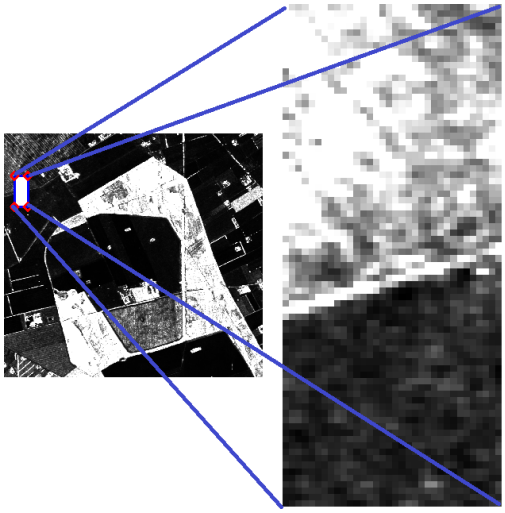


(b)



(c)

Figure 4.6: Estimated type I error probability for several parameters with $\alpha = 1\%$, $\alpha = 5\%$, and $\alpha = 10\%$, respectively.



(a)

Figure 4.7: PolSAR image with selected region Denmark.

Table 4.1: Statistical descriptive measures

Measures	Channel HH	Channel VV
Minimum	0.0025	0.0055
Mean	0.0165	0.0267
Median	0.0146	0.0208
SD	0.0121	0.0195
Skewness	1.1511	2.5267
Kurtosis	4.4181	12.8274
Maximum	0.0864	0.1923

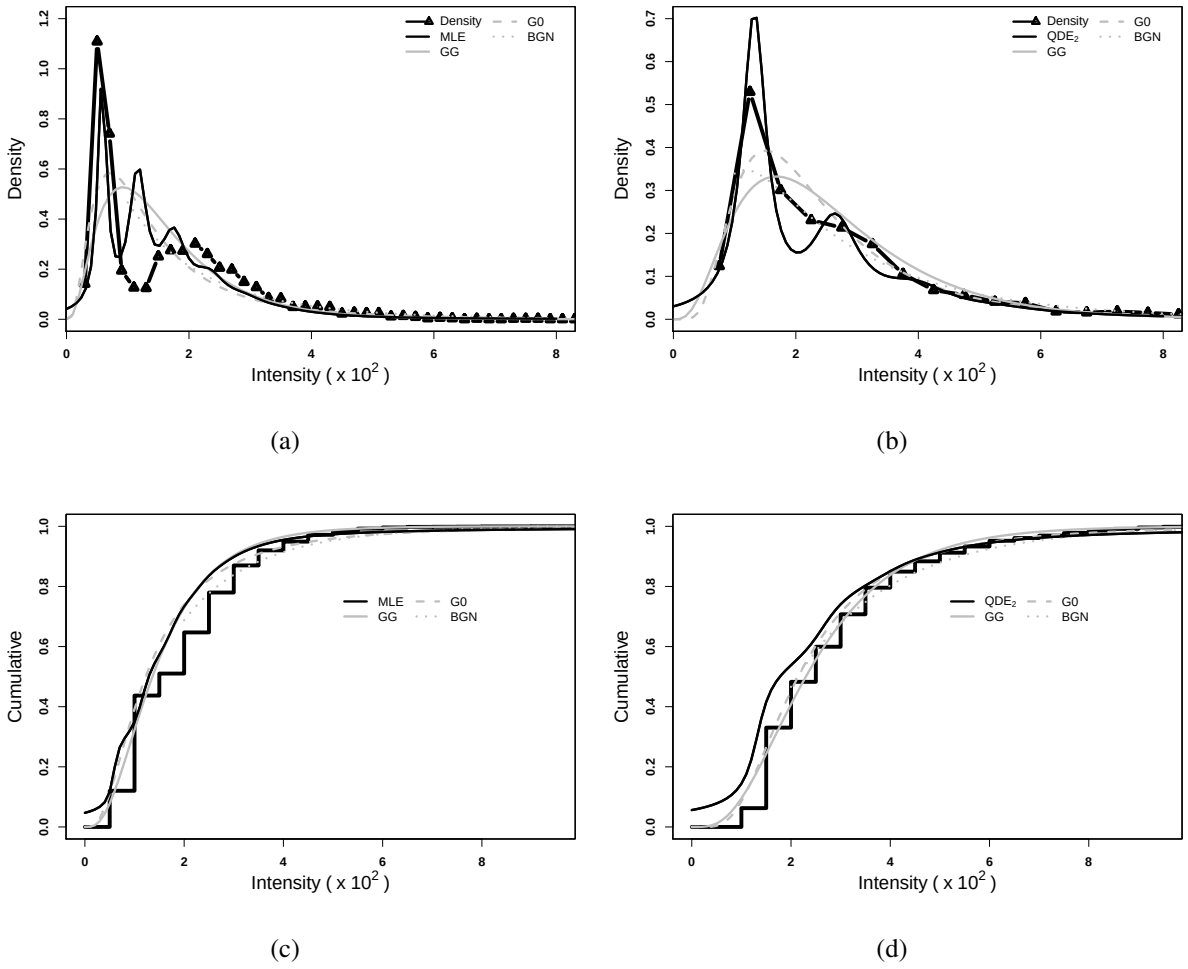


Figure 4.8: Densities adjusted by estimation method (MLE and $VQDE_2$)of the sub-region for two channels (HH-Fig. 4.8(a) and Fig. 4.8(c)- and VV-Fig. 4.8(b) and Fig. 4.8(d)- respectively).

pdf and cdf curves in Fig. 4.8. One can observe that CTPC pdf and cdf fits are closer to the corresponding empirical functions of the channels. To confirm it, we apply GoF test based on the Kolmogorov-Smirnov (KS) statistics. In general, smaller KS values indicate better fit. Results are presented in Table 4.2. The best performance is obtained by the CTPC distribution with MLE and QDE_2 methods. Moreover, regarding the proposed GoF tool, p-values associated with statistics designed in Algorithm 1 for MLE and QDE2 are 0.492 and 0.484, respectively; i.e., both databases can be understood as outcomes of the CTPC model.

Table 4.2: Estimates of the model parameters and the Statistic of KS test of the channels

	Distributions	Methods	Estimates $(\lambda, \mu, \gamma, \alpha, \beta)$					KS
Channel HH	CTPC	**MLE**	**2.3463**	**0.5819**	**0.0864**	−−−	−−−	**0.1074**
		QDE_1	1.9790	0.7045	0.1049	−−−	−−−	0.1524
		QDE_2	1.5885	0.5742	0.1945	−−−	−−−	0.1973
		$VQDE_1$	1.9014	0.7460	0.1062	−−−	−−−	0.1828
		$VQDE_2$	2.2609	0.6037	0.0622	−−−	−−−	0.1153
	Gamma	MLE	1.9636	−−−	1.1907	−−−	−−−	0.2450
	Weibull	MLE	1.4406	−−−	0.5469	−−−	−−−	0.1378
	GG	MLE	9.8567	0.0144	0.4957	−−−	−−−	0.1857
	$\mathcal{K}$	MLE	6.0458	−−−	3.6724	2.8570	−−−	0.1405
	$\mathcal{G}^0$	MLE	−−−	−3.2020	3.7833	4.7838	−−−	0.1272
	BGN	MLE	0.1329	0.2405	0.8651	4.5401	0.1506	0.1413
	CTPG	MLE	2.0001	−−−	−−−	0.6094	0.1021	0.1190
Channel VV	CTPC	MLE	1.3275	1.3039	0.2165	−−−	−−−	0.0919
		QDE_1	1.0291	1.5102	0.1439	−−−	−−−	0.1024
		QDE_2	**1.2196**	**1.3202**	**0.2354**	−−−	−−−	**0.0597**
		$VQDE_1$	0.9574	1.6379	0.1589	−−−	−−−	0.1533
		$VQDE_2$	1.3972	1.2946	0.2188	−−−	−−−	0.0929
	Gamma	MLE	2.7334	−−−	1.0248	−−−	−−−	0.1177
	Weibull	MLE	1.5488	−−−	0.3336	−−−	−−−	0.1190
	GG	MLE	9.9020	0.0366	0.5455	−−−	−−−	0.0980
	$\mathcal{K}$	MLE	12.0986	−−−	4.5895	3.8570	−−−	0.0921
	$\mathcal{G}^0$	MLE	−−−	−4.3179	8.8194	10.9331	−−−	0.0659
	BGN	MLE	0.3917	0.5408	0.9218	3.1273	0.2498	0.0883
	CTPG	MLE	1.3498	−−−	−−−	1.3011	0.2125	0.1091

4.7 Conclusion

We have proposed a new model with three parameters to describe multimodal data, called CTPC. In particular, the CTPC model formation may be related to radar data dynamics. Our model has been used to describe features obtained from SAR images, named intensity.

Some mathematical properties of the CTPC model have been derived and studied; such as, cf and a distance measure (as a possible GoF measure). Further, we proposed two estimation methods: maximum likelihood and quadratic QD estimators. A Monte Carlo study has indicated QD estimators can present better results. We also made an application to real SAR data. Our model may furnish the best fit when compared with six SAR models: Γ, Weibull, GΓ, $\mathcal{K}$, $\mathcal{G}^{\circ}$, beta generalized normal distributions, and CTPG.

5 An Extension for the Family of $\mathcal{G}$ Distributions

Abstract

Synthetic aperture radar (SAR) is a central tool for mapping scenarios on the Earth's surface. Like its processing considers the coherent illumination, images obtained by such system are strongly corrupted by an interference pattern, called "speckle noise". This phenomenon can significantly degrade the perceived SAR image quality, making it difficult to analyze and interpret. Thus, the proposal of statistical models which are able for describing speckled data is sought as an important preprocessing activity. In this sense, the multiplicative approach has received considerable attention. An important particular case from this class is the $\mathcal{G}$ family proposed by Frery *et al.* [IEEE Transactions on Geoscience and Remote Sensing, vol. 35, no. 11, pp. 648–659, 1994], which is a wide class of models for speckle data. In this chapter, we extend such class, proposing models more extensive than classical distributions from the $\mathcal{G}$ family: such as $\mathcal{G}^0$ and $\mathcal{K}$ models. Additionally, the statistical properties such as moment, and Mellin-based log-cumulants (LCs), and maximum likelihood methods concerning to the new distributions are derived. Finally, applications to actual data provide evidence that the new distributions outperform usual $\mathcal{G}$ models.

Keywords: SAR imagery. Speckle noise. Multiplicative model. $\mathcal{G}$ family.

5.1 Introduction

Synthetic aperture radar has been widely used as an important tool for monitoring informations in remote sensing applications. This fact can be explained by means of its capability of operating in all-weather and all-day as well as of providing images with high spacial resolution. However, like its resulting images stem from a coherent illumination process, they are affected

by a signal-dependent granular noise called *speckle* [206]. A primary source of noise is the coherent interference of the signal backscatter by the elements of the target surface. Due to this contamination, rendered imagery presents a granular appearance which precludes its immediate interpretation. In this sense, the proposal of appropriate statistical models has been sought as a necessary preprocessing step.

Many works have recently aimed at finding adequate models to fit speckled data [20, 123, 175, 207, 208]. For example, Delignon *et al.* [20] proposed the KUBW Pearson system for the statistical modeling of ocean SAR images. This approach was extended for the non-Rayleigh speckle distribution [175]. In recent years, results on empirical distributions have favored the search for statistical models [208]. Importantly, the multiplicative model is one of the most successful approaches for describing speckled data [207]. Unlike others methods, this modeling has a phenomenological nature which is closely tied to the physics of the image formation. In this sense, the $\mathcal{G}$ distribution proposed by Frery *et al.* [123] represents one of the most successful models due to its flexibility. In particular, the $\mathcal{G}$ distribution has two practical submodels, namely, the $\mathcal{K}$ and $\mathcal{G}^0$ distributions.

The $\mathcal{K}$ law was originally proposed by Jakeman and Pulsey [209] in order to extend the Rayleigh distribution which is often used for describing homogeneous surfaces. Several works have addressed its parameter estimation by means of the maximum likelihood (ML) [210, 211] and the moment method (MM) [212]. On the other hand, the $\mathcal{G}^0$ distribution was derived by Frery *et al.* [123]. This model is similar to the Fisher distribution, which is a reparametrization of the $\mathcal{G}^0$ distribution itself [213]. The estimation of its parameters is often based on ML [214], MM [215], robust procedures [215–218], and bias correction [219–221]. Additionally, parameter hypothesis tests based on classical inference [222], information theory [223], and non-parametric method [213, 224, 225] have been suggested. In [208, 226] the $\mathcal{K}$ and $\mathcal{G}^0$ models have been examined and found not to be accurate descriptions for all polarization channel intensities in PolSAR images.

The main goal of this chapter is the proposition of a new probability distribution capable of encompassing the $\mathcal{G}$ distribution family. In particular, the $\mathcal{K}$ and $\mathcal{G}^0$ submodels are sought to be extended with the inclusion of a single additional parameter. Additionally, new evidence corroborating with conclusions in [208, 226] is sought. We also aim at deriving statistical properties of

proposed distributions and apply them to actual SAR data. Further, we developed Mellin-based LCs as well as their covariance matrix. The proposed models are assessed according to detailed computational experiments with SAR data, considering goodness-of-fit (GoF) measures as figures of merit.

The rest of this chapter is structured as follows. First, the multiplicative model theory is revisited in Section 5.2. Section 5.3 details the mathematical derivation of the proposed distribution. Some of its statistical properties are examined and the relationships with other models are schematized. An application to real data is performed in Section 5.4. GoF measures are employed as figures of merit to assess the proposed models. Section 5.5 concludes this chapter.

5.2 Multiplicative model and the $\mathcal{G}$ distribution

One of first models for speckled data obtained from a SAR mechanism was considered by Arsenault [227]. Thereafter, statistical multiplicative modeling for SAR images was suggested by Ward [228]. This approach assumes that each picture element is the outcome of a random variable Z called which is the product of two independent positive random variables, denoted by X and Y. Random variables X and Y model the terrain backscatter and the speckle noise, respectively.

Let $f_X(x;\boldsymbol{\theta}_X)$ and $f_Y(y;\boldsymbol{\theta}_Y)$ be the densities associated to X and Y with parameter vectors $\boldsymbol{\theta}_X$ and $\boldsymbol{\theta}_Y$, respectively. The pdf of $Z = X \cdot Y$ is given by the following expression [123]:

$$f_Z(z;\boldsymbol{\theta}_X,\boldsymbol{\theta}_Y) = \int_0^\infty x^{-1} f_Y(z/x;\boldsymbol{\theta}_Y) f_X(x;\boldsymbol{\theta}_X)\,\mathrm{d}x. \tag{5.1}$$

A particular successful distribution derived from the multiplicative model is the $\mathcal{G}$ distribution [123]. Among the particular sub-models of the $\mathcal{G}$ distribution, the $\mathcal{G}^0$ [213] and the $\mathcal{K}$ [211] distributions occupy central positions.

In the following subsections, we review the probability distributions usually employed for modeling the speckle noise and the backscattering. By simplicity of notation, we suppressed the argument parameters of densities.

5.2.1 *Speckle noise*

Resulting from the coherent record of the returned echoes, SAR images are strongly contaminated by speckle noise. The speckle Y is often described by the gamma distribution, $Y \sim \Gamma(L, \mu^{-1}\mathcal{L})$ [123], with pdf given by

$$f_Y(y) = \frac{\mathcal{L}^{\mathcal{L}}}{\mu\Gamma(\mathcal{L})}\left(\frac{y}{\mu}\right)^{\mathcal{L}-1}\exp\left[-\mathcal{L}\frac{y}{\mu}\right]\mathbb{I}_{(0,\infty)}(y), \tag{5.2}$$

where $\mathcal{L} > 0$ represents the number of looks and $\mathbb{I}_{(0,\infty)}(\cdot)$ represents the indicator function with respect to the set $(0,\infty)$. In practice, when the parameter $\mathcal{L}$ (which, is often considered an integer number) is estimated from real data, one seeks to obtain the equivalent number of looks (ENL). A detailed discussion about estimation of this quantity is presented in [229].

5.2.2 *Intensity backscatter*

The backscatter contains the relevant information of the mapped area [223]. Indeed, it depends on target physical properties such as moisture and relief. In practice, low backscatter values are represented by dark areas whereas high backscatter values are associated to bright areas. For example, the reflected radiation in smooth regions, such as a lake or crop fields, is often projected away from the sensor. Since moisture spreads the reflected signal, forest regions present a higher degree of backscattering, when compared to homogeneous areas. On the other hand, urban regions are extremely heterogeneity and, thus, allow that the sensor captures much of the reflected energy.

The backscatter is often described according to the following expression:

$$\xi = \sum_{k=1}^{n} A_k \exp(j\theta_k),$$

where A_k, $k = 1,2,\ldots,n$, are real non-negative random variables representing amplitudes, θ_k, $k = 1,2,\ldots,n$, represent the phase of the kth component, $j = \sqrt{-1}$ is the imaginary unit, and n is the number of scatters number per resolution cell. The random variables $A_1, A_2, \ldots, A_n$ and $\theta_1, \theta_2, \ldots, \theta_n$ are assumed to be independent and identically distributed. Whereas the amplitudes A_k, $k = 1,2,\ldots,n$, may follow any distribution of identical finite variance σ^2; the phases are uniformly distributed over the interval $[0, 2\pi)$ [175,209]. For large values of n, the variable ξ is circular

Gaussian distributed [175] with independent components of zero mean and variance $n\sigma^2$. Based on this result, the amplitude relative to ξ follows the Rayleigh distribution, while its intensity is exponentially distributed parameter $(n\sigma^2)^{-1}$.

Several works have indicated the following distributions as representative models for characterizing intensity data: (i) the Dirac distribution (used to model homogeneous areas) [20], (ii) the gamma distribution (used to model heterogeneous areas) with pdf given by

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp[-\lambda x] \mathbb{I}_{(0,\infty)}(x),$$

where $\alpha, \lambda > 0$ [211]; and (iii) the reciprocal gamma distribution (used to model extremely heterogeneous areas) whose pdf is furnished by [123]:

$$f_X(x) = \frac{1}{\gamma^\alpha \Gamma(-\alpha)} x^{\alpha-1} \exp\left[-\frac{\gamma}{x}\right] \mathbb{I}_{(0,\infty)}(x),$$

where $-\alpha, \mu > 0$.

In order to unify the different statistical models, the generalized inverse Gaussian (GIG) distribution, was proposed in [123]. Denoted by $X \sim \mathcal{N}^{-1}(\alpha, \gamma, \lambda)$, its associate random variable is equipped with the following pdf:

$$f_X(x) = \frac{(\lambda/\gamma)^{\alpha/2}}{2K_\alpha(2\sqrt{\lambda\gamma})} x^{\alpha-1} \exp\left[-\frac{\gamma}{x} - \lambda x\right] \mathbb{I}_{(0,\infty)}(x),$$

where $K_\alpha(\cdot)$ is the Bessel function with order α, $-\infty < \alpha < \infty$, and $\gamma, \lambda > 0$.

5.3 Intensity return: new models and their properties

In this subsection, we present expressions for densities from the proposed distributions as well as their associated statistical properties, the LCs and covariance matrix Mellin-based. It is made for the most general case of this approach and for two of its particular cases which extend the $\mathcal{K}$ and $\mathcal{G}^0$ distributions.

5.3.1 *Generalized Speckle Model*

In this chapter, we consider a more general model for describing the speckle noise. To that end, we suggest as an appropriate model a reparametrization of the generalized gamma (GΓ) distribution, $Y \sim \mathrm{G}\Gamma(\mathcal{L},\mu^{-1}\mathcal{L}^{1/\theta},\theta)$, whose associate pdf is given by:

$$f_Y(y) = \theta\frac{\mathcal{L}^{\mathcal{L}}}{\mu\Gamma(\mathcal{L})}\left(\frac{y}{\mu}\right)^{\theta\mathcal{L}-1}\exp\left[-\mathcal{L}\left(\frac{y}{\mu}\right)^{\theta}\right]\mathbb{I}_{(0,\infty)}(y). \tag{5.3}$$

Li *et al.* [208] suggested this distribution as an empirical-statistical modeling of intensity SAR images. In that case, the parameters θ, $\mathcal{L}$, and μ were considered as the power, shape, and scale parameters, respectively. Because (5.3) generalizes (5.2), it can be considered as an extended model for the speckle noise. In particular, the parameters θ, $\mathcal{L}$, and μ are the model correction (to the intensification of the texture presence), ENL, and intensity mean parameters, respectively.

5.3.2 *Generalized Intensity Backscatter Model*

We propose the extended generalized inverse Gaussian distribution (EGIG) as a more general model for the intensity backscatter. According to Ralph *et al.* [230], this distribution can be derived as a power transformation of the GIG law. In that case, if $X \sim \mathrm{EGIG}(\alpha,\gamma,\lambda,\theta)$, then its pdf is expressed by:

$$f_X(x) = \frac{\theta(\lambda/\gamma)^{\theta\alpha/2}}{2K_\alpha(2\sqrt{(\lambda\gamma)^\theta})}x^{\theta\alpha-1}\times\exp\Big[-(\gamma/x)^\theta-(\lambda x)^\theta\Big]\mathbb{I}_{(0,\infty)}(x), \tag{5.4}$$

where $\theta > 0$. By specifying $\alpha > 0$ and $\gamma \to 0$, the above distribution collapses into the generalized gamma law, $X \sim \mathrm{G}\Gamma(\alpha,\lambda,\theta)$, with pdf given by:

$$f_X(x) = \frac{\theta\lambda^{\theta\alpha}}{\Gamma(\alpha)}x^{\theta\alpha-1}\exp[-(\lambda x)^\theta]\mathbb{I}_{(0,\infty)}(x). \tag{5.5}$$

The above distribution is flexible enough to encompass several other distributions as particular subcases: (i) exponential ($\theta=\alpha=1$), (ii) gamma ($\theta=1$), (iii) Weibull ($\alpha=1$), (iv) log-normal ($\alpha\to\infty$), (v) generalized normal ($\theta=2$), (vi) half-normal ($\alpha=1/2$ and $\lambda^2=2\sigma^2$), (vii) Rayleigh ($\alpha=1$ and $\lambda^2=2\sigma^2$), (viii) Maxwell-Boltzman ($\alpha=3/2$), and (ix) χ^2 ($\alpha=k/2$, for $k=1,2,\ldots$) [231].

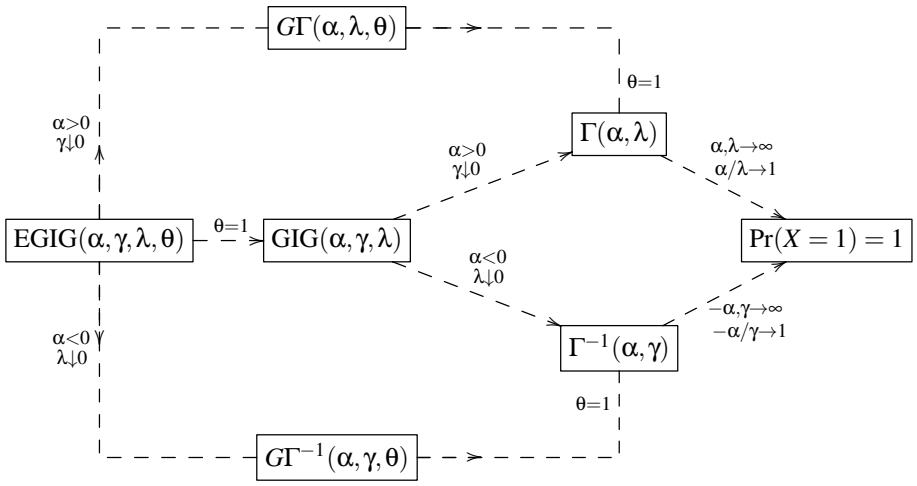


Figure 5.1: Relationships among the probabilistic models for the intensity backscatter.

On the other hand, considering $\alpha < 0$ and $\lambda \to 0$, the pdf in (5.4) is reduced to:

$$f_X(x) = \frac{\theta}{\gamma^{\theta\alpha}\Gamma(\alpha)} x^{\theta\alpha-1} \exp\Big[-\Big(\frac{\gamma}{x}\Big)^{\theta}\Big] \mathbb{I}_{(0,\infty)}(x). \tag{5.6}$$

Such distribution is the reciprocal generalized gamma, $X \sim \mathrm{G}\Gamma^{-1}(\alpha,\lambda,\theta)$.

In Fig. 5.1, we summarize the relationships among the above-mentioned probabilistic models for the intensity backscatter.

5.3.3 *The G-$\mathcal{G}$ distribution*

Let $X \sim \mathrm{EGIG}(\alpha,\gamma,\lambda,\theta)$ and $Y \sim \mathrm{G}\Gamma(\mathcal{L},\mu^{-1}\mathcal{L}^{1/\theta},\theta)$. The application of (5.3) and (5.4) in (5.1) furnishes the return variable $Z = X \cdot Y$, whose pdf is given by:

$$f_Z(z) = \frac{\theta^2 \mathcal{L}^{\mathcal{L}} (\lambda/\gamma)^{\frac{\alpha\theta}{2}} (z/\mu)^{\theta\mathcal{L}-1}}{\mu\Gamma(\mathcal{L}) K_\alpha(2(\gamma\lambda)^{\theta/2})} \left[\frac{\mathcal{L}(z/\mu)^\theta + \gamma^\theta}{\lambda^\theta}\right]^{\frac{\alpha-\mathcal{L}}{2}} K_{\alpha-\mathcal{L}}\Big(2\sqrt{\lambda^\theta[\mathcal{L}(z/\mu)^\theta + \gamma^\theta]}\Big) \mathbb{I}_{(0,\infty)}(z),$$

where $\alpha,\gamma,\lambda,\theta,\mu,\mathcal{L} > 0$. The variable Z is said to follow the G-$\mathcal{G}$ distribution denoted by $Z \sim$ G-$\mathcal{G}(\alpha,\gamma,\lambda,\mathcal{L},\theta,\mu)$. In what follows, we derive some mathematical properties of the G-$\mathcal{G}$ distribution. These properties can be used for extracting features in SAR images, such as coefficient of variation, skewness, and kutosis as we aim at demonstrating in experimental section of this chapter.

Theorem 6 *Let $Z \sim G\text{-}\mathcal{G}(\alpha,\gamma,\lambda,\mathcal{L},\theta,\mu)$. The non-central moments of Z are given by:*

$$\mathrm{E}(Z^k) = \frac{\mu^k}{\mathcal{L}^{\frac{k}{\theta}}}\left(\frac{\gamma}{\lambda}\right)^{k/2}\frac{\Gamma(\mathcal{L}+\frac{k}{\theta})}{\Gamma(\mathcal{L})}\frac{K_{\alpha+\frac{k}{\theta}}(2(\gamma\lambda)^{\theta/2})}{K_\alpha(2(\gamma\lambda)^{\theta/2})}.$$

In particular, the variance of Z is given by:

$$\mathrm{Var}(Z) = \frac{\mu^2}{\mathcal{L}^2\Gamma(\mathcal{L})K_\alpha(2(\gamma\lambda)^{\theta/2})}\frac{\gamma}{\lambda}\left[\Gamma\left(\mathcal{L}+\frac{2}{\theta}\right)K_{\alpha+\frac{2}{\theta}}(2(\gamma\lambda)^{\theta/2}) - \frac{\Gamma^2\left(\mathcal{L}+\frac{1}{\theta}\right)}{\Gamma(\mathcal{L})}\frac{K^2_{\alpha+\frac{1}{\theta}}(2(\gamma\lambda)^{\theta/2})}{K_\alpha(2(\gamma\lambda)^{\theta/2})}\right].$$

The proof of Theorem 6 derives from (i) the non central moment expressions of $X \sim$ EGIG and $Y \sim \mathrm{G}\Gamma$ and (ii) the assumption of independence of X and Y in $Z = X \times Y$.

First particular case: the G-$\mathcal{K}$ distribution

Let $X \sim \mathrm{G}\Gamma(\alpha,\lambda,\theta)$ and $Y \sim \mathrm{G}\Gamma(\mathcal{L},\mu^{-1}\mathcal{L}^{1/\theta},\theta)$. By applying (5.3) and (5.5) in (5.1), the pdf of the return variable Z is expressed by:

$$f_Z(z) = \frac{2\theta\mathcal{L}^{1/\theta}\lambda}{\mu\Gamma(\alpha)\Gamma(\mathcal{L})}\left(\frac{\mathcal{L}^{1/\theta}\lambda z}{\mu}\right)^{\theta\frac{\alpha+\mathcal{L}}{2}-1}K_{\alpha-\mathcal{L}}\left(2\sqrt{\mathcal{L}\lambda^\theta(z/\mu)^\theta}\right)\mathbb{I}_{(0,\infty)}(z),$$

where $\alpha,\lambda,\theta,\mu,\mathcal{L} > 0$. This distribution is referred to as the G-$\mathcal{K}$ distribution, denoted as $Z \sim$ G-$\mathcal{K}(\alpha,\lambda,\mathcal{L},\theta,\mu)$.

By setting $\alpha > 0$ and $\gamma \downarrow 0$ in Theorem 6, the following corollary is obtained.

Corollary 2 *Let $Z \sim G\text{-}\mathcal{K}(\alpha,\lambda,\mathcal{L},\theta,\mu)$. Its non-central moments are expressed by*

$$\mathrm{E}(Z^k) = \left(\frac{\mu}{\mathcal{L}^{1/\theta}\lambda}\right)^k\frac{\Gamma(\frac{k}{\theta}+\mathcal{L})\Gamma(\frac{k}{\theta}+\alpha)}{\Gamma(\mathcal{L})\Gamma(\alpha)}.$$

Accordingly, the associated variance is:

$$\mathrm{Var}(Z) = \left(\frac{\mu}{\mathcal{L}^{1/\theta}\lambda}\right)^2\frac{1}{\Gamma(\mathcal{L})\Gamma(\alpha)}\left[\Gamma\left(\frac{2}{\theta}+\mathcal{L}\right)\Gamma\left(\frac{2}{\theta}+\alpha\right) - \frac{\Gamma^2\left(\frac{1}{\theta}+\mathcal{L}\right)\Gamma^2\left(\frac{1}{\theta}+\alpha\right)}{\Gamma(\mathcal{L})\Gamma(\alpha)}\right].$$

Second particular case: the G-$\mathcal{G}^0$ distribution

Let $X \sim \mathrm{G}\Gamma^{-1}(\alpha,\gamma,\theta)$ and $Y \sim \mathrm{G}\Gamma(\mathcal{L},\mu^{-1}\mathcal{L}^{1/\theta},\theta)$. Combining (5.3) and (5.6) into (5.1) yields the pdf of the return variable Z as follows:

$$f_Z(z) = \frac{\theta\mathcal{L}^{\mathcal{L}}\gamma^{-\theta\alpha}\Gamma(\mathcal{L}-\alpha)}{\mu\Gamma(\mathcal{L})\Gamma(-\alpha)}\left(\frac{z}{\mu}\right)^{\theta\mathcal{L}-1}\left[\gamma^\theta + \mathcal{L}\left(\frac{z}{\mu}\right)^\theta\right]^{\alpha-\mathcal{L}}\mathbb{I}_{(0,\infty)}(z),$$

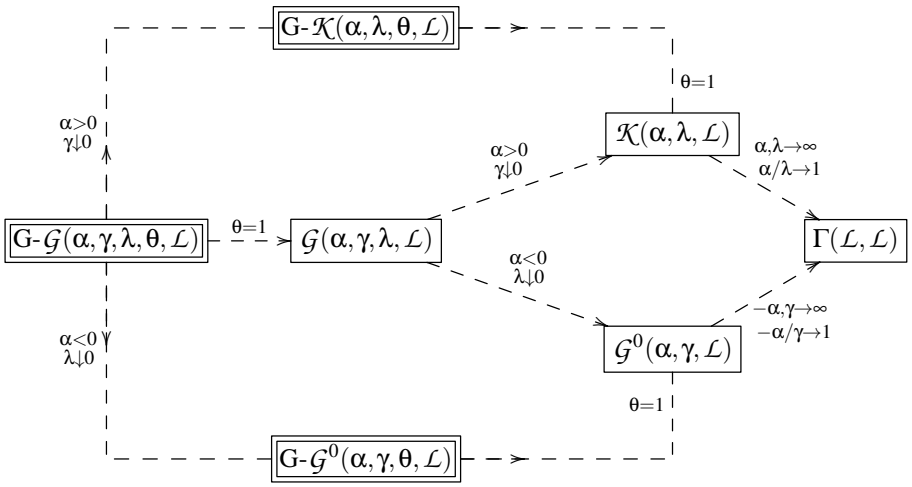


Figure 5.2: The new distributions are emphasized in double-line boxes

where $\alpha,\gamma,\lambda,\theta,\mu,\mathcal{L}>0$. Such distribution is termed the G-$\mathcal{G}^0$ distribution; and we may write $Z\sim$ G-$\mathcal{G}^0(\alpha,\gamma,\mathcal{L},\theta,\mu)$. If we let $\alpha<0$ and $\lambda\downarrow 0$ in Theorem 6, the following corollary holds true.

Corollary 3 *Let $Z\sim$ G-$\mathcal{G}^0(\alpha,\gamma,\mathcal{L},\theta,\mu)$. Its non-central moment is expressed by:*

$$\mathrm{E}(Z^k)=\Big(\frac{\gamma\mu}{\mathcal{L}^{1/\theta}}\Big)^k\frac{\Gamma(\frac{k}{\theta}+\mathcal{L})\Gamma(-\alpha-\frac{k}{\theta})}{\Gamma(\mathcal{L})\Gamma(-\alpha)}.$$

In particular, the variance of this distribution is given by:

$$\mathrm{Var}(Z)=\Big(\frac{\gamma\mu}{\mathcal{L}^{1/\theta}}\Big)^2\frac{1}{\Gamma(\mathcal{L})\Gamma(-\alpha)}\Big[\Gamma\Big(\frac{2}{\theta}+\mathcal{L}\Big)\Gamma\Big(-\alpha-\frac{2}{\theta}\Big)-\frac{\Gamma^2\big(\frac{1}{\theta}+\mathcal{L}\big)\Gamma^2\big(-\alpha-\frac{1}{\theta}\big)}{\Gamma(\mathcal{L})\Gamma(-\alpha)}\Big].$$

The above described models are related as shown in the diagram in Fig. 5.2. The proposed models are pictorially represented in double boxes.

ML Estimation for the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ distributions

Let $\boldsymbol{z}=(z_1,z_2,\ldots,z_N)$ be a possible outcome of a random sample from the G-$\mathcal{G}^0$ distribution with unknown parameter vector $\boldsymbol{\delta}=(\alpha,\gamma,\theta,\mathcal{L})$, where N is the number of pixels in a selected

sample. The log-likelihood function $\ell(\boldsymbol{\delta}) = \ell(\boldsymbol{\delta}; \boldsymbol{z})^\top$ is given by:

$$\begin{aligned}\frac{\ell(\boldsymbol{\delta})}{N} &= \log\theta + \mathcal{L}\log\mathcal{L} - \alpha\theta\log\gamma + \log\Gamma(\mathcal{L}-\alpha) - \log\mu - \log\Gamma(\mathcal{L}) - \log\Gamma(-\alpha) \\ &\quad + (\theta\mathcal{L}-1)\sum_{i=1}^{N}\log\frac{z_i}{\mu} + (\alpha-\mathcal{L})N^{-1}\sum_{i=1}^{N}\log\left(\gamma^\theta + \mathcal{L}\left(\frac{z_i}{\mu}\right)^\theta\right).\end{aligned} \tag{5.7}$$

Analogously, for $\boldsymbol{\xi} = (\alpha, \lambda, \theta, \mathcal{L})^\top$, if $\boldsymbol{z}$ is G-$\mathcal{K}$ distributed, then we have:

$$\begin{aligned}\frac{\ell(\boldsymbol{\delta})}{N} &= \log 2 + \log\theta + \frac{\alpha+\mathcal{L}}{2}\log\mathcal{L} + \theta\frac{\alpha+\mathcal{L}}{2}\log\lambda - \theta\frac{\alpha+\mathcal{L}}{2}\log\mu - \log\Gamma(\alpha) \\ &\quad - \log\Gamma(\mathcal{L}) - N^{-1}\left(\theta\frac{\alpha+\mathcal{L}}{2} - 1\right)\sum_{i=1}^{N}\log z_i + N^{-1}\sum_{i=1}^{N}\log\mathcal{K}_{\alpha-\mathcal{L}}(\varphi_i(\boldsymbol{\xi})),\end{aligned} \tag{5.8}$$

where $\varphi_i(\boldsymbol{\delta}) = 2\sqrt{\mathcal{L}\lambda^\theta (z_i/\mu)^\theta}$, for $i = 1, 2, \ldots, N$. The score function associated to (5.7) and (5.8) are generically given by $\boldsymbol{U}_1(\boldsymbol{\delta}) = (U_\alpha, U_\gamma, U_\theta, U_\mathcal{L})^\top$ and $\boldsymbol{U}_2(\boldsymbol{\delta}) = (U_\alpha, U_\lambda, U_\theta, U_\mathcal{L})^\top$, respectively, where

$$(U_\alpha, U_\gamma, U_\lambda, U_\theta, U_\mathcal{L}) = \left(\frac{\mathrm{d}\ell(\boldsymbol{\delta})}{\mathrm{d}\alpha}, \frac{\mathrm{d}\ell(\boldsymbol{\delta})}{\mathrm{d}\gamma}, \frac{\mathrm{d}\ell(\boldsymbol{\delta})}{\mathrm{d}\lambda}, \frac{\mathrm{d}\ell(\boldsymbol{\delta})}{\mathrm{d}\theta}, \frac{\mathrm{d}\ell(\boldsymbol{\delta})}{\mathrm{d}\mathcal{L}}\right),$$

were derived and are shown in Fig. 5.3. Finally, the ML estimators, $(\widehat{\alpha}, \widehat{\gamma}, \widehat{\theta}, \widehat{\mathcal{L}})$ and $(\widehat{\alpha}, \widehat{\lambda}, \widehat{\theta}, \widehat{\mathcal{L}})$, are defined as the solutions of the non-linear system linked to $\boldsymbol{U}_1(\boldsymbol{\delta}) = \boldsymbol{0}$ and $\boldsymbol{U}_2(\boldsymbol{\delta}) = \boldsymbol{0}$, respectively.

5.3.4 *Mellin Transform: Second kind statistic for the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models*

Mellin Transform and LCs

The most successful tool in the signal processing is the Fourier transform. In statistics, it is defined as the cf on which it is possible to obtain the Pearson diagram [18].

But, the cf is not always tractable. To solve this issue, Colombo [100] has suggested the MT as an alternative tool for characterizing models. Nicolas [22] derived and applied the MT in the SAR image processing context. In what follows, we present some MT results for the G-$\mathcal{G}^0$ and G-$\mathcal{K}$ models.

Let $X \in \mathbb{R}_+$ be a positive random variable with pdf $f(x)$, thereby its MT is given by

$$\phi_X(s) = \int_0^\infty x^{s-1} f(x)\mathrm{d}x = \mathrm{E}(X^{s-1}),$$

a) For the G-$\mathcal{G}^0$ model:

$$
\begin{aligned}
\frac{U_\alpha}{N} =& -\theta\log\gamma - \psi^{(0)}(\mathcal{L}-\alpha) + \psi^{(0)}(-\alpha) + N^{-1}\sum_{i=1}^{N}\log\left(\gamma^\theta + \mathcal{L}\left(\frac{z_i}{\mu}\right)^\theta\right),\\
\frac{U_\gamma}{N} =& -\frac{\alpha\theta}{\gamma} + \theta(\alpha-\mathcal{L})\gamma^{\theta-1}N^{-1}\sum_{i=1}^{N}\left\{\frac{1}{\gamma^\theta + \mathcal{L}(z_i/\mu)^\theta}\right\},\\
\frac{U_\theta}{N} =& \frac{1}{\theta} - \alpha\log\gamma + \mathcal{L}N^{-1}\sum_{i=1}^{N}\log\frac{z_i}{\mu} + \mathcal{L}(\alpha-\mathcal{L})N^{-1}\sum_{i=1}^{N}\left\{\frac{\log(z_i/\mu)(z_i/\mu)^\theta}{\gamma^\theta + \mathcal{L}(z_i/\mu)^\theta}\right\}\\
&+ (\alpha-\mathcal{L})\gamma^\theta\log\gamma N^{-1}\sum_{i=1}^{N}\left\{\frac{1}{\gamma^\theta + \mathcal{L}(z_i/\mu)^\theta}\right\},\\
\frac{U_{\mathcal{L}}}{N} =& \log\mathcal{L} + 1 + \psi^{(0)}(\mathcal{L}-\alpha) - \psi^{(0)}(\mathcal{L}) + \theta N^{-1}\sum_{i=1}^{N}\log\frac{z_i}{\mu} - N^{-1}\sum_{i=1}^{N}\log(\gamma^\theta + \mathcal{L}(z_i/\mu)^\theta)\\
&+ (\alpha-\mathcal{L})N^{-1}\sum_{i=1}^{N}\left\{\frac{(z/\mu)^\theta}{\gamma^\theta + \mathcal{L}(z/\mu)^\theta}\right\}.
\end{aligned}
$$

b) For the G-$\mathcal{K}$ model:

$$
\begin{aligned}
\frac{U_\alpha}{N} =& \frac{\log\mathcal{L}}{2} + \frac{\theta}{2}\log\lambda - \frac{\theta}{2}\log\mu - \psi(\alpha) + N^{-1}\frac{\theta}{2}\sum_{i=1}^{N}\log z_i + N^{-1}\sum_{i=1}^{N}\left[\frac{\partial K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi}))}{\partial\alpha}\right]/K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi})),\\
\frac{U_\gamma}{N} =& \theta\frac{\alpha+\mathcal{L}}{2\lambda} + N^{-1}\sum_{i=1}^{N}\left[\frac{\partial K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi}))}{\partial\lambda}\right]/K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi})),\\
\frac{U_\theta}{N} =& \frac{1}{\theta} + \frac{\alpha+\mathcal{L}}{2}\log\frac{\lambda}{\mu} + N^{-1}\frac{\alpha+\mathcal{L}}{2}\sum_{i=1}^{N}\log z_i + N^{-1}\sum_{i=1}^{N}\left[\frac{\partial K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi}))}{\partial\lambda}\right]/K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi})),\\
\frac{U_{\mathcal{L}}}{N} =& \frac{\log\mathcal{L}}{2} + \frac{\alpha+\mathcal{L}}{2\mathcal{L}} + \frac{\theta}{2}\log\frac{\lambda}{\mu} - \psi(\mathcal{L}) + N^{-1}\frac{\theta}{2}\sum_{i=1}^{N}\log z_i + N^{-1}\sum_{i=1}^{N}\sum_{i=1}^{N}\left[\frac{\partial K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi}))}{\partial\mathcal{L}}\right]/K_{\alpha-\mathcal{L}}(\varphi(z_i,\boldsymbol{\xi})).
\end{aligned}
$$

Figure 5.3: G-$\mathcal{G}^0$ and G-$\mathcal{K}$ score functions.

where $s = a + ib \in \mathbb{C}$ is a complex variable [22].

It is known the kth moment of the MT is defined as

$$\widetilde{m}_k = \frac{\mathrm{d}^k \, \phi_X(s)}{\mathrm{d}\, s^k}\bigg|_{s=1} = \int_{\mathbb{R}_+} (\log x)^k f(x)\, \mathrm{d}\, x, \quad k \in \mathbb{N}.$$

The second kind mean (known as log-mean) given by Nicolas [22], say $\widetilde{m}_1$, is

$$\log(\widetilde{\mu}) = \widetilde{m}_1 \iff \widetilde{\mu} = \mathrm{e}^{\widetilde{m}_1}. \tag{5.9}$$

The tailored cumulative generator function is given

$$\psi_X(s) = \log\left(\phi_X(s)\right).$$

The second kind cumulants or LCs of order k are obtained from the nth derivative of $\psi_X(s)$ evaluated at $s = 1$:

$$\widetilde{\kappa}_k = \frac{\mathrm{d}^k \psi_X(s)}{\mathrm{d} s^k}\bigg|_{s=1}, \qquad k \in \mathbb{N}.$$

For proposed models in the previous section, the following MT expressions hold.

Corollary 4 *If $Z \sim G\text{-}\mathcal{K}(\alpha, \lambda, \mathcal{L}, \theta, \mu)$ or $Z \sim G\text{-}\mathcal{G}^0(\alpha, \gamma, \mathcal{L}, \theta, \mu)$ their MT, are expressed by*

$$\phi_Z(s) = \begin{cases} \left(\frac{\mu}{\mathcal{L}^{1/\theta}\lambda}\right)^{s-1} \frac{\Gamma\left(\frac{s-1}{\theta}+\mathcal{L}\right)\Gamma\left(\frac{s-1}{\theta}+\alpha\right)}{\Gamma(\mathcal{L})\Gamma(\alpha)}, & Z \sim G\text{-}\mathcal{K}; \\ \\ \left(\frac{\mu\gamma}{\mathcal{L}^{1/\theta}}\right)^{s-1} \frac{\Gamma\left(\frac{s-1}{\theta}+\mathcal{L}\right)\Gamma\left(-\frac{s-1}{\theta}-\alpha\right)}{\Gamma(\mathcal{L})\Gamma(-\alpha)}, & Z \sim G\text{-}\mathcal{G}^0. \end{cases}$$

To proof the Corollary 4 derive from (i) the assumption of independence of X and Y in $Z = X \times Y$, (ii) apply $\phi_Z(s) = \phi_X(s)\phi_Y(s)$, and (iii) use the function $\Gamma_x(\delta) = \int_0^\infty x^{\delta-1}\, \mathrm{e}^{-u}\, \mathrm{d}\, u$. To obtain the classical moments, we can make $s = k + 1$ in the previous corollary. Based on Corollary 4, the associated LCs follow.

Corollary 5 *If $Z \sim G\text{-}\mathcal{K}(\alpha, \lambda, \mathcal{L}, \theta, \mu)$ or $Z \sim G\text{-}\mathcal{G}^0(\alpha, \gamma, \mathcal{L}, \theta, \mu)$, the LCs for $k > 1$ are given by*

$$\widetilde{\kappa}_k = \begin{cases} \frac{\psi^{(k-1)}(\mathcal{L}) + \psi^{(k-1)}(\alpha)}{\theta^k}, & Z \sim G\text{-}\mathcal{K}, \\ \\ \frac{\psi^{(k-1)}(\mathcal{L}) + (-1)^k\, \psi^{(k-1)}(-\alpha)}{\theta^k}, & Z \sim G\text{-}\mathcal{G}^0, \end{cases}$$

being that $\psi^{(k)}(x) = {}^{\partial^{k+1}\log\Gamma(x)}/_{\partial x^{k+1}}$ and $\psi(x)$ is the digamma function [112, p. 260].

Note that G-$\mathcal{K}$ LCs with order higher than two do not depend of μ and λ, whilst the G-$\mathcal{G}^0$ LCs independent of μ and γ for these orders. From (5.9) and Corollary 4, for $Z \sim$ G-$\mathcal{K}(\alpha,\lambda,\mathcal{L},\theta,\mu)$, we have

$$\widetilde{\mu} = \frac{\mu}{\mathcal{L}^{1/\theta}\lambda}\exp\left\{\frac{\Psi(\mathcal{L})+\Psi(\alpha)}{\theta}\right\} \underset{\theta\to\infty}{\longrightarrow} \frac{\mu}{\lambda},$$

and for $Z \sim$ G-$\mathcal{G}^0(\alpha,\gamma,\mathcal{L},\theta,\mu)$

$$\widetilde{\mu} = \frac{\mu\gamma}{\mathcal{L}^{1/\theta}}\exp\left\{\frac{\Psi(\mathcal{L})-\Psi(-\alpha)}{\theta}\right\} \underset{\theta\to\infty}{\longrightarrow} \mu\gamma.$$

Delignon *et al.* [20] discussed the Pearson diagram based on skewness and kurtosis measures for SAR imagery. This map has been used as an important GoF measure. Nicolas [22] presented evidence that this technique may be analytically intractable. This author introduced the $(\widetilde{\kappa}_3,\widetilde{\kappa}_2)$ diagram, which represents the analogue to the Pearson diagram, using $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$ instead of skewness and kurtosis measures. They showed that the $(\widetilde{\kappa}_3,\widetilde{\kappa}_2)$ diagram is a suitable alternative for classifying SAR images. Li *et al.* [226] also considered the MT-based diagram for pdf classification from SAR imagery data.

Now we introduce the $(\widetilde{\kappa}_3,\widetilde{\kappa}_2)$ diagram for the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models using LCs in Corollary 5. Fig. 5.4 exhibits the G-$\mathcal{K}$, G-$\mathcal{G}^0$, $\mathcal{K}$, and $\mathcal{G}^0$ regions in the $(\widetilde{\kappa}_3,\widetilde{\kappa}_2)$ diagram. In this graph, the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models are represented by surface and $\mathcal{G}^0$ and $\mathcal{K}$ models by curves, whose dimensions depend on the texture parameter number in the associated distribution. The shape of the $\mathcal{G}^0$ and $\mathcal{K}$ models are driven by one texture parameter and, therefore, represented by curves (black and dotted, respectively); i.e., having one-dimensional manifolds. The G-$\mathcal{G}^0$ and G-$\mathcal{K}$ models have three parameters and are represented by surfaces (shading lines) [60], i.e., two-dimensional manifolds.

5.3.5 *GoF for the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models*

This section aims to present GoF measures for the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models. This is the way to combine the asymptotic behavior of LC estimators with the Hotelling T^2 statistics [232, 233].

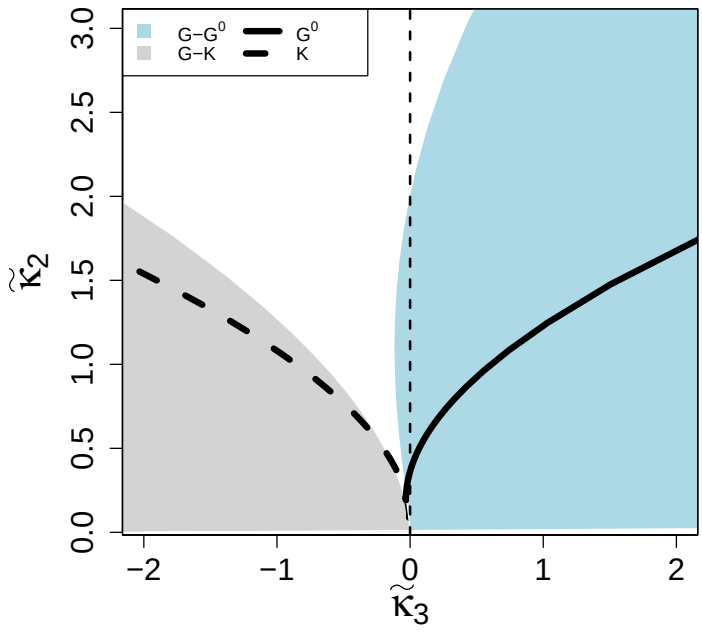


Figure 5.4: Diagram of the LCs $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ for G-$\mathcal{K}$, G-$\mathcal{G}^0$, $\mathcal{K}$ and $\mathcal{G}^0$ models.

Covariance matrix of LCEs

It is known the following result holds: Let

$$\widetilde{\boldsymbol{m}} = (\widetilde{m}_1, \ldots, \widetilde{m}_p)^\top \quad \text{and} \quad \widehat{\widetilde{\boldsymbol{m}}} = (\widehat{\widetilde{m}}_1, \ldots, \widehat{\widetilde{m}}_p)^\top.$$

The central limit theorem and the Cramér-Wold theorem [162] proved that

$$\sqrt{n}(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{M}_p), \tag{5.10}$$

where $\boldsymbol{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution with mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$, and

$$\boldsymbol{M}_p = \mathrm{E}\left\{(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}})(\widehat{\widetilde{\boldsymbol{m}}} - \widetilde{\boldsymbol{m}})^\top\right\} = \{\widetilde{m}_{r+\nu} - \widetilde{m}_r \widetilde{m}_\nu\}\Big|_{r,\nu=1}^{p}.$$

Let $g_\nu : \mathbb{R}^p \to \mathbb{R}$, be a moment-to-cumulant transformation function (which are continuously differentiable), thus we get

$$\widetilde{\boldsymbol{\kappa}}_\nu = g_\nu(\widetilde{m}_1, \ldots, \widetilde{m}_p) \quad \text{and} \quad \widehat{\widetilde{\boldsymbol{\kappa}}}_\nu = g_\nu(\widehat{\widetilde{m}}_1, \ldots, \widehat{\widetilde{m}}_p).$$

Using the multivariate delta method [162] and (5.10), together with the Cramér-Wold theorem, we have

$$\sqrt{n}(\widehat{\widetilde{\boldsymbol{\kappa}}}-\widetilde{\boldsymbol{\kappa}}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_p(\mathbf{0},\, \boldsymbol{K}_p),$$

where

$$\boldsymbol{K}_\nu = \boldsymbol{J}_\nu^\top \cdot \boldsymbol{M}_\nu \cdot \boldsymbol{J}_\nu,$$

$$\boldsymbol{\nabla} g_\nu = \left[\frac{\partial}{\partial \widetilde{m}_1} g_\nu(\widetilde{\boldsymbol{m}}), \ldots, \frac{\partial}{\partial \widetilde{m}_\nu} g_\nu(\widetilde{\boldsymbol{m}})\right]^\top, \text{ for } \nu = 1, \ldots, r,$$

and

$$\boldsymbol{J}_\nu = \left[\boldsymbol{\nabla} g_1 \middle| \cdots \middle| \boldsymbol{\nabla} g_\nu\right]^\top.$$

Thence, the covariance matrices of LCEs vector $[\kappa_2, \kappa_3]^\top$ is agiven by

$$\boldsymbol{K}_{\text{G-}\mathcal{K}} = \begin{bmatrix} \widetilde{\kappa}_{22} & \widetilde{\kappa}_{23} \\ \\ \widetilde{\kappa}_{32} & \widetilde{\kappa}_{33} \end{bmatrix},$$

where

$$\begin{aligned}
\widetilde{\kappa}_{22} &= \frac{1}{\theta^4}\left\{\psi^{(3)}(\mathcal{L}) + \psi^{(3)}(\alpha) + 2[\psi(\mathcal{L}) + \psi(\alpha)]^2\right\}, \\
\widetilde{\kappa}_{23} = \widetilde{\kappa}_{32} &= \frac{1}{\theta^5}\left\{\psi^{(4)}(\mathcal{L}) + \psi^{(4)}(\alpha) + +6[\psi^{(2)}(\mathcal{L}) + \psi^{(2)}(\alpha)][\psi(\mathcal{L}) + \psi(\alpha)]\right\}, \\
\widetilde{\kappa}_{33} &= \frac{1}{\theta^6}\left\{\psi^{(5)}(\mathcal{L}) + \psi^{(5)}(\alpha) + +9[\psi^{(3)}(\mathcal{L}) + \psi^{(3)}(\alpha)][\psi(\mathcal{L}) + \psi(\alpha)] + \right. \\
&\quad \left. +9[\psi^{(2)}(\mathcal{L}) + \psi^{(2)}(\alpha)]^2 + 6[\psi(\mathcal{L}) + \psi(\alpha)]^3\right\};
\end{aligned}$$

and for the another model we have:

$$\boldsymbol{K}_{\text{G-}\mathcal{G}^0} = \begin{bmatrix} \widetilde{\kappa}_{22} & \widetilde{\kappa}_{23} \\ \\ \widetilde{\kappa}_{32} & \widetilde{\kappa}_{33} \end{bmatrix},$$

where

$$\widetilde{\kappa}_{22} = \frac{1}{\theta^4}\left\{\psi^{(3)}(\mathcal{L}) + \psi^{(3)}(-\alpha) + 2[\psi(\mathcal{L}) + \psi(-\alpha)]^2\right\},$$
$$\widetilde{\kappa}_{23} = \widetilde{\kappa}_{32} = \frac{1}{\theta^5}\left\{\psi^{(4)}(\mathcal{L}) - \psi^{(4)}(-\alpha) + +6[\psi^{(2)}(\mathcal{L}) - \psi^{(2)}(-\alpha)][\psi(\mathcal{L}) + \psi(-\alpha)]\right\},$$
$$\widetilde{\kappa}_{33} = \frac{1}{\theta^6}\left\{\psi^{(5)}(\mathcal{L}) + \psi^{(5)}(-\alpha) + +9[\psi^{(3)}(\mathcal{L}) + \psi^{(3)}(-\alpha)][\psi(\mathcal{L}) + \psi(-\alpha)] + \right.$$
$$\left. +9[\psi^{(2)}(\mathcal{L}) - \psi^{(2)}(-\alpha)]^2 + 6[\psi(\mathcal{L}) + \psi(-\alpha)]^3\right\}.$$

It is known the statistic:

$$T^2 = n\left(\begin{bmatrix}\widehat{\widetilde{\kappa}}_2\\ \widehat{\widetilde{\kappa}}_3\end{bmatrix} - \begin{bmatrix}\widetilde{\kappa}_2\\ \widetilde{\kappa}_3\end{bmatrix}\right)^{\top} \widehat{\boldsymbol{K}}^{-1}\left(\begin{bmatrix}\widehat{\widetilde{\kappa}}_2\\ \widehat{\widetilde{\kappa}}_3\end{bmatrix} - \begin{bmatrix}\widetilde{\kappa}_2\\ \widetilde{\kappa}_3\end{bmatrix}\right),$$

where $\widehat{\boldsymbol{K}}^{-1} = \left\{\widehat{\widetilde{\kappa}}^{r,v}\right\}_{r,v=2,3}$, follows the chi-square distribution. Thus, the probability of random vector to belong to the contour is

$$P\left(T^2 \leq \chi^2_{2,\eta}\right) = 1 - \eta.$$

From brief algebraic manipulation one can provide GoF measures for G-$\mathcal{K}$, and G-$\mathcal{G}^0$ models, similarly [232].

Proposition 12 *Let Z be a random variable, in which Z can assume G-$\mathcal{K}(\alpha,\lambda,\mathcal{L},\theta,\mu)$ or G-$\mathcal{G}^0(\alpha,\gamma,\mathcal{L},\theta,\mu)$ distribution, then the T^2 statistics is given by*

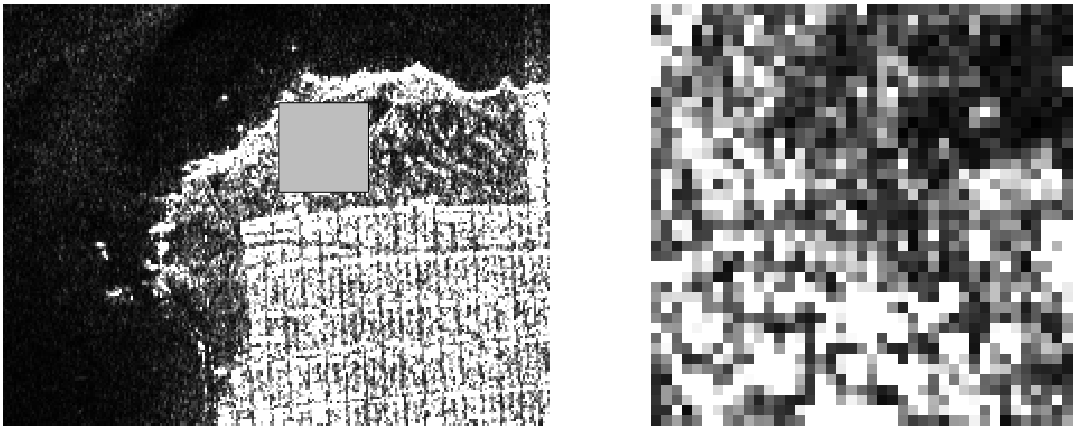
$$T^2 = n\left[\widehat{\widetilde{\kappa}}^{22}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)^2 + \widehat{\widetilde{\kappa}}^{33}\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)^2 2\widehat{\widetilde{\kappa}}^{23}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)\right],$$

where $\widehat{\widetilde{\kappa}}^{22}$, $\widehat{\widetilde{\kappa}}^{33}$, $\widehat{\widetilde{\kappa}}^{23}$, $\widehat{\widetilde{\kappa}}_2$, and $\widehat{\widetilde{\kappa}}_3$, are the estimators of $\widetilde{\kappa}^{22}$, $\widetilde{\kappa}^{33}$, $\widetilde{\kappa}^{23}$, $\widetilde{\kappa}_2$, and $\widetilde{\kappa}_3$, respectively;

In particular its asymptotic confidence interval (ACI) is given by

$$\text{ACI}\left(\widetilde{\kappa}_n,\, 100(1-\gamma)\%\right) = \left(\widehat{\widetilde{\kappa}}_n \pm z_{\gamma/2}\sqrt{\widehat{\widetilde{\kappa}}^{n,n}}\right),$$

for $n = 2,3$, and $z_{\gamma/2}$ is the quantile $(1-\gamma/2)$ of the standard normal distribution.



(a) Slected HH SAR image (b) Selected region

Figure 5.5: SAR image of the selected forest region.

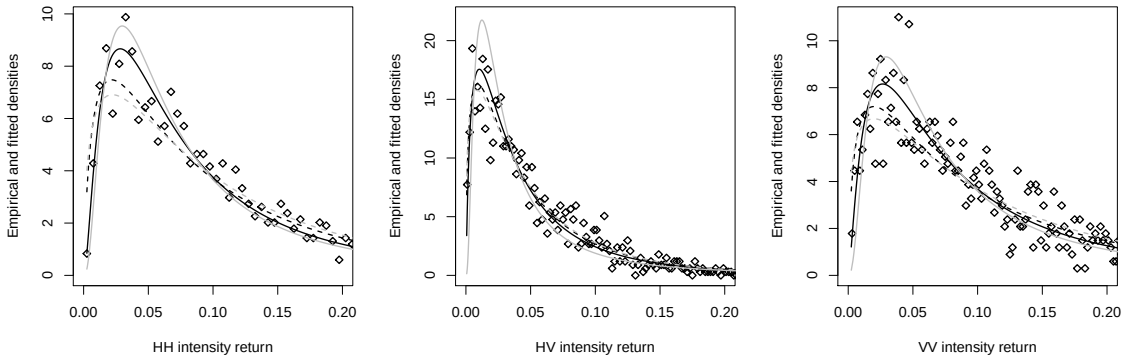
5.4 Application to real SAR data

In this section, we analyze actual intensity SAR data. In particular, we consider the forest region of the San Francisco Bay image obtained by the AIRSAR sensor in L-band with four nominal looks [1]. Fig. 5.5 illustrates the selected SAR image with details.

We aim at submitting the intensity data from the HH, HV, and VV polarization channels of the selected image to the discussed models G-$\mathcal{G}^0$, G-$\mathcal{K}$, $\mathcal{G}^0$, and $\mathcal{K}$, according to the discussed ML estimation.

We focused our attention on the G-$\mathcal{G}^0$ and G-$\mathcal{K}$ models, because they are solutions requiring fewer parameters when compared with G-$\mathcal{G}$ distribution. Moreover, we do not include the $\mathcal{G}$ model in this study because of the following rationale. The $\mathcal{G}$ model collapses to the $\mathcal{G}^0$ and $\mathcal{K}$ distributions after a limiting process subject to: (a) $\alpha < 0$ and $\lambda \downarrow 0$ and (b) $\alpha > 0$ and $\gamma \downarrow 0$, respectively. However, the conditions imposed on γ and λ are difficult to check in practice. It is a statistically non-trivial task—in the hypothesis test theory sense—to verify whether $(\mathcal{G}, \mathcal{G}^0)$ and $(\mathcal{G}, \mathcal{K})$. differ. On the other hand, the $\mathcal{G}^0$ and $\mathcal{K}$ models are implicit in the new G-$\mathcal{G}^0$ and G-$\mathcal{K}$ distributions under the condition $\theta = 1$, which can be easily checked.

Fig. 5.6 shows the obtained fitted densities for each case. Qualitative visual inspection suggests that the $\mathcal{G}$ family suitably models the HH and HV intensity data in the form of $\mathcal{G}^0$ and $\mathcal{K}$ distri-



(a) forest HH channel (b) forest HV channel (c) forest VV channel

Figure 5.6: Fitted densities for G-$\mathcal{G}^0$ (solid black), G-$\mathcal{K}$ (dashed black), $\mathcal{G}^0$ (solid gray), and $\mathcal{K}$ (dashed gray) distributions.

butions, respectively. However, for the VV intensity data, it is less successful. For quantitative analysis, we consider two GoF measures: (i) the corrected version of Akaike information criterion (AICc) and (ii) Bayesian information criterion (BIC). The computational results for the AIC returned identical values to the values associated to the corrected AIC. This is expected because AIC and AICc behave differently only for small and moderate sample sizes [234]. Table 5.1 shows ML estimates and GoF measurements. In all cases the proposed G-$\mathcal{G}^0$ and G-$\mathcal{K}$ outperformed the competing models.

Table 5.2 presents the mean ($\bar{Z}$) and variance Var(Z) on both real data and for fitted models according to Corollaries 2 and 3. The proposed models describe the sample mean and variance better than the traditional $\mathcal{G}^0$ and $\mathcal{K}$ distributions. The variance by means of $\mathcal{G}^0$ fits are undefined; however this difficulty is overcome by the proposed G-$\mathcal{G}^0$ model.

Now, we aim at quantifying the discrimination capability between the elements in the following pairs: ($\mathcal{G}^0$,G-$\mathcal{G}^0$) and ($\mathcal{K}$,G-$\mathcal{K}$), at neighborhoods of SAR intensity data. To that end, it was employed two ratio likelihood statistics to test $\mathcal{H}_0 : \theta = 1$, which have chi-squared asymptotic distribution with one degree of freedom and expressions based on G-$\mathcal{G}^0$ and G-$\mathcal{K}$ models given by, respectively:

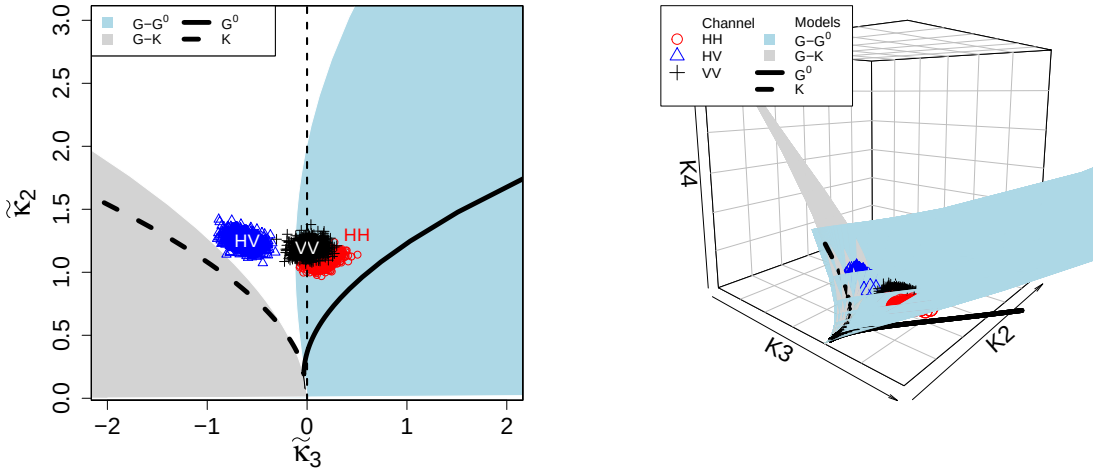
$$S_{\text{G-}\mathcal{G}^0} = 2\sum_{i=1}^{W}\Big[\log f_{\text{G-}\mathcal{G}^0}(z_i) - \log f_{\mathcal{G}^0}(z_i)\Big], \quad \text{and} \quad S_{\text{G-}\mathcal{K}} = 2\sum_{i=1}^{W}\Big[\log f_{\text{G-}\mathcal{K}}(z_i) - \log f_{\mathcal{K}}(z_i)\Big],$$

Table 5.1: Estimates of parameters and GoF measures

Models	Estimates ($\widehat{\xi}$)				Goodness-of-fit ($\times 10^3$)	
	$\widehat{\alpha}$	$\widehat{\gamma}/\widehat{\lambda}$	$\widehat{\theta}$	$\mathcal{L}$	AICc	BIC
Results for the channel HH						
G-$\mathcal{G}^0$	−3.120	0.317	0.775	4	**−3.702**	**−3.685**
G-$\mathcal{K}$	2.673	30.394	0.769	4	−3.595	−3.579
$\mathcal{G}^0$	−1.679	0.107	n/a	4	−3.676	−3.665
$\mathcal{K}$	1.456	11.296	n/a	4	−3.511	−3.500
Results for the channel HV						
	−10.763	2.761	0.559	4	**−6.292**	**−6.276**
	1.745	33.522	0.901	4	−6.288	−6.272
	−1.519	0.042	n/a	4	−6.077	−6.066
	1.380	23.897	n/a	4	−6.280	−6.269
Results for the channel VV						
	−3.859	0.546	0.698	4	**−3.448**	**−3.432**
	2.300	22.739	0.786	4	−3.365	−3.349
	−1.555	0.100	n/a	4	−3.393	−3.382
	1.361	9.912	n/a	4	−3.294	−3.284

Table 5.2: Comparison between empirical and descriptive properties

Model/Data	Channel	$\bar{Z}$	Var(Z)
Data	HH	0.1317362	0.0531517
	HV	0.0557846	0.0066813
	VV	0.1423261	0.0531961
G-$\mathcal{G}^0$	HH	0.1389117	0.0951966
	HV	0.0592943	0.0059955
	VV	0.1471238	0.0827471
$\mathcal{G}^0$	HH	0.1575847	∞
	HV	0.0809249	∞
	VV	0.1801802	∞
G-$\mathcal{K}$	HH	0.1321899	0.0234291
	HV	0.0579247	0.0041409
	VV	0.1414152	0.0282010
$\mathcal{K}$	HH	0.1288952	0.0184169
	HV	0.0577478	0.0038544
	VV	0.1373083	0.0220293



(a) (b)

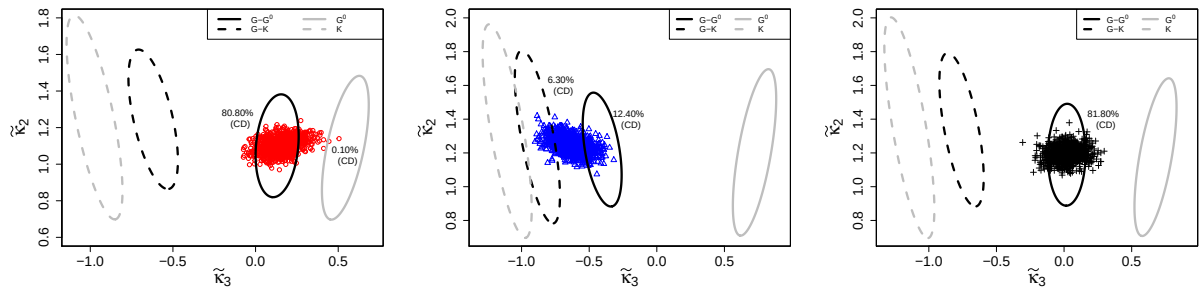
Figure 5.7: In left panel we have the diagram of the LCs $\widetilde{\kappa}_3 \sim \widetilde{\kappa}_2$ and in right panel the LCs cube, both plots are in respect to the G-$\mathcal{K}$ and G-$\mathcal{G}^0$ models.

where $W = 7 \times 7$ pixels is the window size. In Fig. 5.9, we present the p-value map associated with both statistics. In general, results provide evidence that the model pair (G-$\mathcal{K}$, $\mathcal{K}$) presents the highest difference. From results presented by $S_{G-\mathcal{K}}$ for HH and VV channels, higher differences are found at regions with more pronounced textures. On the other hand, differences seem to be uniformly distributed in the maps provided by $S_{G-\mathcal{G}^0}$, specially at ocean regions.

Fig. 5.7 exhibits the $(\widetilde{\kappa}_3, \widetilde{\kappa}_2)$ diagram and overlapping data. To build it, multiple points have been obtained for each set by means of the bootstrap method [235]. In particular 1,000 replicates were considered.

For the HH and VV channels, the majority of over lapping points provide are covered by the G-$\mathcal{G}^0$ model at the Fig. 5.7(a). For the HV channel, discussed models do not present good performance, although our proposals have furnished the best results. We also provide the 3D diagram in the Fig. 5.7(b), called LC cube [28], for the two proposed models. From this approach we observe that data of three channels are covered the G-$\mathcal{G}^0$ region.

Besides of LC diagram and LC cube, we provide confidence ellipses (with 90% confidence) their coverage degrees, ACI, and the T^2 statistics (see Table 5.3) for each considered model. The



(a) (b) (c)

Figure 5.8: Confidence ellipses about LC diagram for each data set aforementioned by bootstrap with 1,000 replications.

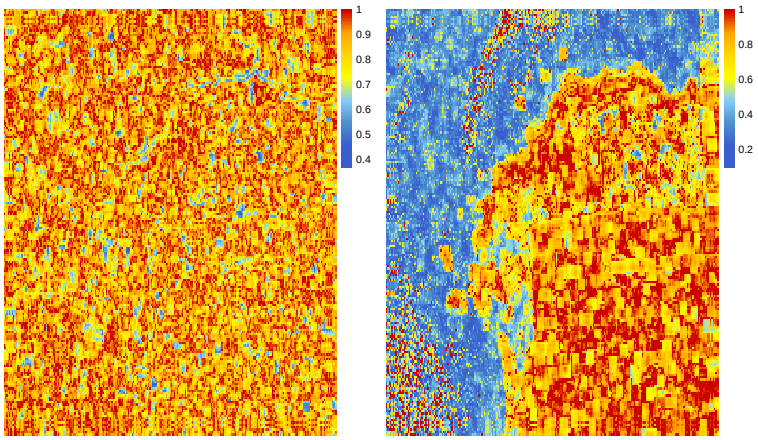
ellipses are obtained through of estimated cumulants $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$ and covariance matrix of LCs of the models studied according to Anfinsen and Elfort [60] (see Fig. 5.8). Results, as already discussed present evidence in favor of the G-$\mathcal{G}^0$ model

5.5 Conclusion

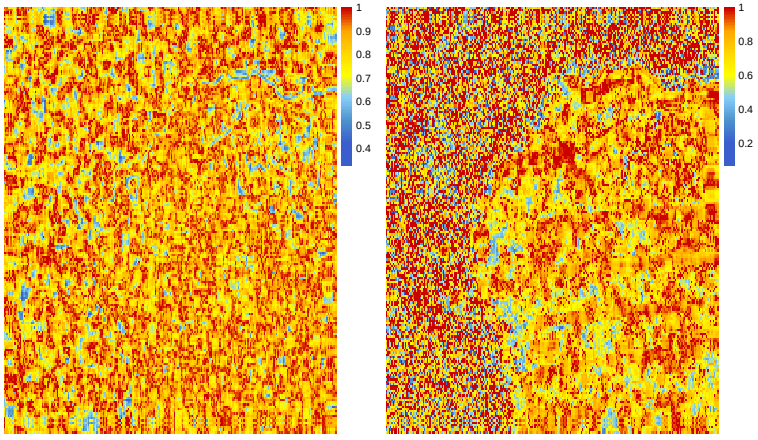
Considering the multiplicative modeling, a new class of distributions for speckled data has been proposed: the G-$\mathcal{G}$ distribution. Among its particular cases, we have two new distributions which generalizes the classical $\mathcal{G}^0$ and $\mathcal{K}$ distributions. We derived statistical properties of the new distributions and related them according to variation of their parameters. Two diagrams were presented to illustrate their relationships. The LCs for the new models and its 2D and 3D diagrams were proposed. Applications to real data have been considered. As figures of merit, we selected three GoF measures (AIC, corrected AIC, and BIC) as well as ones based on LCs and likelihood ratio statistics. We demonstrate that the proposed models are capable of outperforming the $\mathcal{G}^0$ and $\mathcal{K}$ distributions.

Table 5.3: ACI and p-value of the T^2 statistics

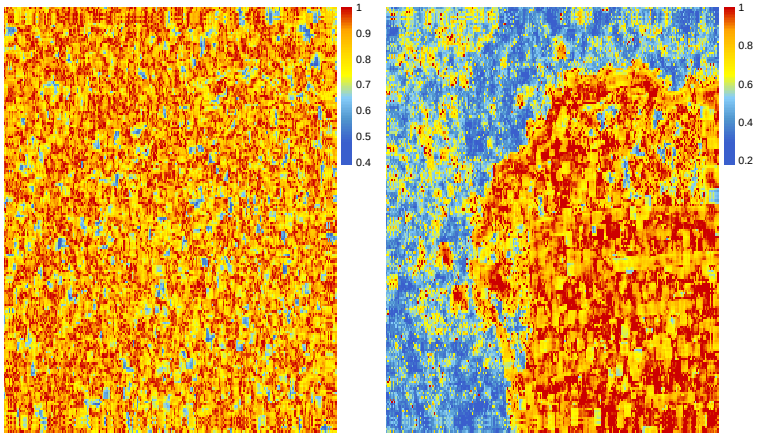
Models	LCs	ACI	p-value
	Results for the channel HH		
G-$\mathcal{G}^0$	$\widetilde{\kappa}_2$	$(-2.1997, 4.4010)$	**0.5763**
	$\widetilde{\kappa}_3$	$(-6.9893, 7.2501)$	
G-$\mathcal{K}$	$\widetilde{\kappa}_2$	$(-2.5434, 5.0340)$	0.0000
	$\widetilde{\kappa}_3$	$(-10.2761, 9.0370)$	
$\mathcal{G}^0$	$\widetilde{\kappa}_2$	$(-2.4983, 4.6782)$	0.0038
	$\widetilde{\kappa}_3$	$(-9.4092, 10.4955)$	
$\mathcal{K}$	$\widetilde{\kappa}_2$	$(-3.0267, 5.5398)$	0.0000
	$\widetilde{\kappa}_3$	$(-15.1112, 13.1624)$	
	Results for the channel HV		
		$(-2.4230, 4.8627)$	**0.1094**
		$(-8.9429, 8.1349)$	
		$(-2.9302, 5.5189)$	0.0005
		$(-13.8510, 12.0847)$	
		$(-2.8540, 5.2603)$	0.0000
		$(-11.7602, 13.2054)$	
		$(-3.2670, 5.9257)$	0.0000
		$(-17.1269, 14.9137)$	
	Results for the channel VV		
		$(-2.3336, 4.7120)$	**0.2132**
		$(-7.6207, 7.6602)$	
		$(-2.8121, 5.4872)$	0.0000
		$(-12.2778, 10.7615)$	
		$(-2.7645, 5.1146)$	0.0000
		$(-11.1508, 12.5033)$	
		$(-3.3334, 6.0318)$	0.0000
		$(-17.6987, 15.4107)$	



(a) G-$\mathcal{G}^0$ p-value map for HH (b) G-$\mathcal{K}$ p-value map for HH



(c) G-$\mathcal{G}^0$ p-value map for HV (d) G-$\mathcal{K}$ p-value map for HV



(e) G-$\mathcal{G}^0$ p-value map for VV (f) G-$\mathcal{K}$ p-value map for VV

Figure 5.9: G-$\mathcal{K}$ and G-$\mathcal{G}^0$ p-value maps for the ratio likelihood statistic of HH, HV, and VV channels from the San Francisco AIRSAR image.

6 $\mathcal{G}_I^0$ Regression Model for Speckled Data

Abstract

Synthetic aperture radar (SAR) is an efficient and widely used remote sensing tool. However, data extracted from SAR images are contaminated with speckle noise, which precludes their modeling directly. Intensities are important SAR features which have a non-additive nature and various distributions obtained from the multiplicative approach have been proposed to describe them. The $\mathcal{G}_I^0$ model is one of the most successful among them. Several inferential procedures have been proposed for $\mathcal{G}_I^0$ parameters, but–from a literature review we made–there are not works which tackle a regression structure for this model. This chapter proposes a $\mathcal{G}_I^0$ regression model in order to consider the influence of other intensities (present in the polarimetric SAR data) in the modeling of intensities due to one particular polarization channel. We derive some theoretical properties for the new model: Fisher information matrix, residual-kind measures, and influential tools. Point and interval estimation procedures via maximum likelihood are proposed and assessed by Monte Carlo experiments. Results from synthetic and real studies point out the new model may be useful for speckled data.

Keywords: Speckled data. $\mathcal{G}_I^0$ regression model. Influence measure. Conditional intensity.

6.1 Introduction

SAR is a remote sensing powerful technique. SAR systems operate under the dynamic: electromagnetic pulses are sent towards a target and their returns are recorded. The intensity of the echoed signal is a powerful information source since it carries physical properties due to the target surface. Modeling intensities is therefore sought in SAR image processing.

SAR images are obtained using coherent illumination and, consequently, strongly contaminated by a particular interference pattern, called speckle noise [223]. This noise imposes deviations from the classical properties of additivity and Gaussian distribution and, therefore, tailored modeling and analysis techniques are mandatory in SAR imagery processing. The multiplicative model (MulM) is one of the most successful approaches for describing speckled data [207]. Unlike other methods, this modeling has a phenomenological nature which is closely tied to the physics of the image formation.

A particular case of the MulM is the $\mathcal{G}_I^0$ distribution, which was pioneered by Frery *et al.* [123]. The $\mathcal{G}_I^0$ model depends on three parameters, which indicate roughness (α), brightness (γ), and the number of looks ($\mathcal{L}$). The last parameter represents a quantity which is related to the signal-to-noise ratio in SAR images (high values of $\mathcal{L}$ indicate smoother images). The $\mathcal{G}_I^0$ model has received great attention by characterizing areas with different degrees of texture (from mild to extreme). This model is also known for its heavy tail, what give itself advantage at describing urban scenarios (which impose more difficult to work with) [236]. Several works have proposed advances for the $\mathcal{G}_I^0$ distribution. The estimation of its parameters is often based on maximum likelihood (ML) [214], moment method [215], robust procedures [215–218], and bias correction [219, 219–221]. Additionally, parameter hypothesis tests based on classical inference [222], information theory [223], and non-parametric method [213, 224, 225] have been suggested. All these works have assumed intensities from a unconditional perspective; i.e., intensities at one polarization channel are not influenced by ones of other channels in the same scene.

In recent years, the interest at understanding such type of imagery in a multidimensional and multilook perspective has increased. Basically, polarimetric SAR measurements record the amplitude and phase of backscattered signals for possible combinations of linear reception and transmission polarizations: HH, HV, VH, and VV (H for horizontal and V for vertical polarization). Evidence from practical situations support the assumption that HV and VH polarizations are identical. Thus, for each pixel of a SAR image, one has three intensities associated to the HH, HV, and VV channels. In this chapter, we discuss about how to model SAR intensities given the influence of ones.

This chapter proposes a new regression model for distributed $\mathcal{G}_I^0$ response variables, which

are related to SAR imagery pixels. Beyond the challenge of deriving theoretical properties of regression of a widely used model, our proposal is motivated by the fact: given a polarimetric SAR return, any particular intensity defined in the former is related to ones. We propose a formulation of a $\mathcal{G}_I^0$ regression model as well as the derivation of some of its properties: Moments for the reciprocal $\mathcal{G}_I^0$ model, Fisher information matrix, two residual-kind and four influential-kind measures. Point and interval estimation procedures are also designed and assessed via a Monte Carlo simulation study. Finally, our proposal is applied to real SAR data and, comparatively to the performance of exponential, gamma (Γ), inverse gamma (Γ^{-1}), normal ($\mathcal{N}$), inverse normal ($\mathcal{N}^{-1}$), Weibull, power exponential, and exponential generalized beta type 2 (EGB2) regression models, shows to be a useful tool in SAR imagery processing.

The chapter unfolds as follows. Section 6.2 presents the main properties of the $\mathcal{G}_I^0$ model. In Section 6.3, we suggest a new regression model and discuss some of its properties. A simulation study involving the $\mathcal{G}_I^0$ regression parameters is presented in Section 6.4. An application to real data is made in Section 6.5. Section 6.6 addresses the main conclusions.

6.2 The $\mathcal{G}_I^0$ distribution

According to the MM approach [223], SAR intensities, say $Z \in \mathbb{R}_+$, are defined as the product of two independent positive random variables: one describes the speckle noise, say Y, and other models the terrain backscatter, say X; i.e.,

$$Z = X \times Y,$$

where X and Y are independent. Assuming $X \sim \Gamma^{-1}(\alpha, \gamma)$ and $Y \sim \Gamma(\mathcal{L}, \mathcal{L})$, Frery *et al.* [123] proposed the $\mathcal{G}_I^0$ distribution having pdf given by

$$f_Z(z; \alpha, \gamma, \mathcal{L}) = \frac{\mathcal{L}^{\mathcal{L}}\Gamma(\mathcal{L}-\alpha)}{\gamma^{\alpha}\Gamma(-\alpha)\Gamma(\mathcal{L})} z^{\mathcal{L}-1}[\gamma + \mathcal{L}z]^{\alpha-\mathcal{L}}, \tag{6.1}$$

where $-\alpha > 0$, $\gamma = \mu(-\alpha - 1) > 0$, and $\mathcal{L} > 0$ (which assumes natural values when obtained from the SAR system) represent roughness, brightness, and the number of looks, respectively. This situation is denoted by $Z \sim \mathcal{G}_I^0(\alpha, \gamma, \mathcal{L})$. In what follows, we present some results which will be used to develop the $\mathcal{G}_I^0$ regression model.

Frery *et al.* [123] proposed the *n*th ordinary moment of the $\mathcal{G}_I^0$ distribution as follows.

Corollary 6 *Let $Z \sim \mathcal{G}_I^0(\alpha, \gamma, \mathcal{L})$. Its non-central moments are given by*

$$\mathrm{E}(Z^k) = \left(\frac{\gamma}{\mathcal{L}}\right)^k \frac{B(\mathcal{L}+k, -\alpha-k)}{B(\mathcal{L}, -\alpha)},$$

if $\alpha < -k$. Otherwise, it is infinite.

Proof 2 *Let $Z \sim \mathcal{G}_I^0(\alpha,\gamma,\mathcal{L})$, then*

$$\mathrm{E}(Z^k) = \int_0^\infty z^k \, f_Z(\boldsymbol{\theta}) \, \mathrm{d}z = \int_0^\infty \frac{\mathcal{L}^{\mathcal{L}}\Gamma(\mathcal{L}-\alpha)}{\gamma^\alpha\Gamma(-\alpha)\Gamma(\mathcal{L})} z^{\mathcal{L}-1+k}[\gamma+\mathcal{L}z]^{\alpha-\mathcal{L}} \, \mathrm{d}z.$$

With this, we make the change of variable $u = \gamma + \mathcal{L}z$ and we compute

$$\mathrm{E}(Z^k) = \frac{\Gamma(\mathcal{L}-\alpha)}{\mathcal{L}^k\gamma^\alpha\Gamma(-\alpha)\Gamma(\mathcal{L})} \int_\gamma^\infty \left(1-\frac{\gamma}{u}\right)^{\mathcal{L}+k-1} u^{\alpha+k-1} \, \mathrm{d}u.$$

Again let $r = \gamma/u$, then

$$\mathrm{E}(Z^k) = \left(\frac{\gamma}{\mathcal{L}}\right)^k \frac{\Gamma(\mathcal{L}-\alpha)}{\Gamma(-\alpha)\Gamma(\mathcal{L})} \int_0^1 (1-r)^{\mathcal{L}+k-1} r^{-\alpha-k-1} \, \mathrm{d}r.$$

As $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} \, \mathrm{d}x$, we acquire:

$$\mathrm{E}(Z^k) = \left(\frac{\gamma}{\mathcal{L}}\right)^k \frac{B(\mathcal{L}+k, -\alpha-k)}{B(\mathcal{L}, -\alpha)}.$$

■

From Corollary 6, the $\mathcal{G}_I^0$ variance is

$$\mathrm{Var}(Z) = \mu^2 \left[\left(\frac{\alpha+1}{\alpha+2}\right)\frac{\mathcal{L}+1}{\mathcal{L}} - 1\right], \tag{6.2}$$

if $\alpha < -2$. Otherwise, it is infinite. We assume the last constraint in all the chapter. It is noticeable the individual increasing of α or μ indicates SAR scenes with higher variability; whereas the increasing of $\mathcal{L}$ implies more homogeneous areas. Two next results are contributions of this chapter.

Corollary 7 *Let $Z \sim \mathcal{G}_I^0(\alpha, \gamma, \mathcal{L})$. Then, the kth inverse moment of $T = (\gamma + \mathcal{L}Z)$ is given by*

$$\mathrm{E}\left[\frac{1}{T^k}\right] = \frac{1}{\gamma^k}\prod_{i=0}^{k-1}\left(\frac{-\alpha+i}{-\alpha+\mathcal{L}+i}\right).$$

■

The proof these Corollary is in the Appendix F.

Lemma 2 *The $\mathcal{G}_I^0$ distribution is a scale family.*

Proof 3 *Let $Z \sim \mathcal{G}_I^0(\alpha, \gamma, \mathcal{L})$ and c a non-negative constant. Then if $Z_1 = cZ$, we have by definition of cumulative distribution function:*

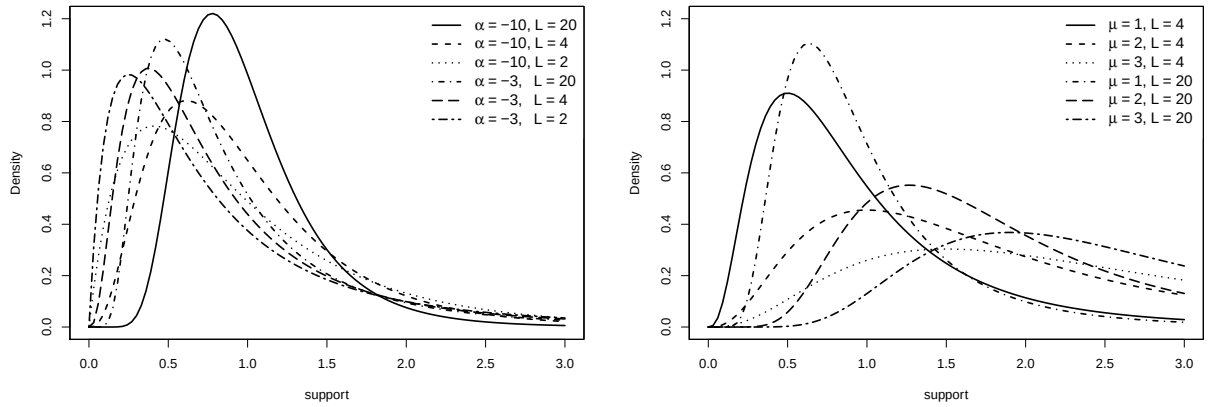
$$F_{Z_1}(z_1) = P(Z \leq z_1/c) \quad i.e. \quad f_{Z_1}(z_1) = \frac{1}{c} f_Z(z_1/c). \tag{6.3}$$

Then, using (6.1) *in* (6.3) *and rearranging the expression, we have:*

$$f_{Z_1}(z_1) = \frac{\mathcal{L}^{\mathcal{L}}\Gamma(\mathcal{L}-\alpha)}{(\gamma c)^{\alpha}\Gamma(-\alpha)\Gamma(\mathcal{L})} {z_1}^{\mathcal{L}-1}[(\gamma c) + \mathcal{L} z_1]^{\alpha-\mathcal{L}},$$

therefore $Z_1 \sim \mathcal{G}_I^0(\alpha, \gamma c, \mathcal{L})$, this is, the $\mathcal{G}_I^0$ distribution is a scale family.

■



(a) Varying α and $\mathcal{L}$ for $\mu = 1$

(b) Varying μ and $\mathcal{L}$ for $\alpha = -5$

Figure 6.1: $\mathcal{G}_I^0$ densities.

Figs 6.1(a) and 6.1(b) show the $\mathcal{G}_I^0$ density curves for $\alpha \in \{-10, -5, -3\}$, $\mu \in \{1, 2, 3\}$, and $\mathcal{L} \in \{2, 4, 20\}$. Evidence are given in sense: (i) the increasing of μ increases location and variability at the same time (what is expected from (6.2)) and (ii) high values of α yield curves having weightier tails.

6.3 $\mathcal{G}_I^0$ Regression Model

When modeling features at a polarization channel extracted from SAR images (like intensity), there are always other features–associated to other channels, which work in parallel. For instance, if one aims to model HH channel features, they have any correlation degree with features due to HV and VV channels. In this section, we provide a mathematical treatment to study linear dependency among intensities of polarization channels.

From Lemma 2, one can define a linear model for the $\mathcal{G}_I^0$ distribution. Let Z_k be a random variable such that

$$Z_k = \mathrm{e}^{\boldsymbol{x}_k^\top \boldsymbol{\beta}}\, \varepsilon_k, \quad \text{for} \quad k = 1, \ldots, n, \tag{6.4}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$ is a vector of regression coefficients, $\boldsymbol{x}_k = (1, x_{k1}, \ldots, x_{kp})^\top$ is the vector of $p < n$ explanatory variables (assumed fixed and known), and $\varepsilon_k \sim \mathcal{G}_I^0(\alpha, -(\alpha+1), \mathcal{L})$. Thus, it follows from Lemma 2 that $[Z_k|\boldsymbol{x}_k] \sim \mathcal{G}_I^0(\alpha, \gamma_k, \mathcal{L})$, where $\gamma_k = \mu_k(-\alpha-1)$, $\mu_k = g^{-1}(\eta_k)$, $\eta_k = \boldsymbol{x}_k^\top \boldsymbol{\beta}$, and $g(\bullet)$ is a positive value, strictly monotonic, and twice differentiable link function. It is known that there are several possible link functions; for instance [237–240]:

$$g(\mu_k) = \begin{cases} \log\{\mu_k/(1-\mu_k)\}, & \text{logit}, \\ \Phi^{-1}(\mu_k), & \text{probit}, \\ \log\{-\log(1-\mu_k)\}, & \text{complementary log-log}, \\ -\log\{-\log(\mu_k)\}, & \text{log-log}, \\ \tan\{\pi(\mu_k - 0.5)\}, & \text{Cauchy}, \\ \log(\mu_k), & \text{log}, \\ \mu_k^\lambda, & \text{power}, \end{cases}$$

where $\Phi(\bullet)$ is the cumulative distribution function of the standard normal distribution. By analytical tractability issues, we choose $g(x) = \log(x)$.

Thus, we assume n SAR intensities at a region of a given channel are described by random variables $Z_1,\ldots,Z_n$ such that $Z_k \sim \mathcal{G}_I^0(\alpha,\gamma_k,\mathcal{L})$ with mean $\mu_k = e^{\boldsymbol{x}_k^\top \boldsymbol{\beta}}$, where $\boldsymbol{x}_k$ is the vector of features known which may be associated to Z_k; e.g., biomass index (as in [241]), other features at the same channel (as phase), and same features at the other channels. In this chapter, we work in the last branch: mapping the linear dependence/effect of the HV channel over HH and VV channels. To illustrate the use of the regression model, we consider a real SAR image (Fig. 6.2(a) in RGB, adopting red, green, and blue as the intensities of HH, HV, and VV, respectively) obtained from the AIRSAR sensor over an region of San Francisco (USA) under four number of looks. This image is characterized by three regions: ocean, forest, and urban. In order to assess the effect of kind of regions over the estimates of parameters α,γ, and μ, we plot maps of estimates obtained over around of 7×7 windows for each pixel in Figs 6.2(b)-6.2(c). Note that maps for α and μ are more informative than that for γ, since only formers reproduce the behavior of the considered SAR image. Finally, Fig. 6.2(e) illustrates the use of the proposed $\mathcal{G}_I^0$ regression model.

6.3.1 *Estimation by maximum likelihood*

Consider $(Z_1\,|\,\boldsymbol{x}_1),\ldots,(Z_n\,|\,\boldsymbol{x}_n)$ be a n-points random sample such that $Z_k \sim \mathcal{G}_I^0(\alpha,\gamma_k,\mathcal{L})$ and $\gamma_k = e^{\boldsymbol{x}_k^\top\boldsymbol{\beta}}(-\alpha-1)$. From (6.1) and (6.4), the log-likelihood function at $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha, \mathcal{L})^\top$ is given as

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^{n} \ell_k(\boldsymbol{\theta}), \tag{6.5}$$

where

$$\begin{aligned}\ell_k(\boldsymbol{\theta}) :=& \mathcal{L}\log(\mathcal{L}) + \log\Gamma(\mathcal{L}-\alpha) - \alpha\log(\mu_k) - \alpha\log(-\alpha-1) - \log\Gamma(-\alpha)\\ &- \log\Gamma(\mathcal{L}) + (\mathcal{L}-1)\log(z_k) + (\alpha-\mathcal{L})\log(t_k)\end{aligned}$$

and $t_k = \gamma_k + \mathcal{L}z_k$ such that z_k is an outcome from Z_k.

Grounded on (6.5), the associated score function,

$$U(\boldsymbol{\theta}) := (U_{\boldsymbol{\beta}}, U_\alpha, U_{\mathcal{L}}) = \left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}^\top}, \frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha}, \frac{\partial\ell(\boldsymbol{\theta})}{\partial\mathcal{L}}\right)^\top,$$

is determined by:

$$U_{\boldsymbol{\beta}} = \alpha \boldsymbol{X}^\top \boldsymbol{E}(\boldsymbol{T}^* - \boldsymbol{\mu}^*), \tag{6.6}$$

where $\boldsymbol{X}$ is an $n \times p$ matrix whose kth row is $\boldsymbol{x}_k^\top$, $\boldsymbol{E} \coloneqq \operatorname{diag}\{1/g'(\mu_1), \ldots, 1/g'(\mu_k), \ldots, 1/g'(\mu_n)\}$, $g'(x) = \mathrm{d}\, g(x)/\mathrm{d}x$, $\boldsymbol{T}^* = (T_1^*, \ldots, T_k^*, \ldots, T_n^*)^\top$, with $T_k^* = {}^{-c_1}/_{T_k}$, $c_1 = (\mathcal{L} - \alpha)(-\alpha - 1)$,
and $\boldsymbol{\mu}^* = (1/\mu_1, \ldots, 1/\mu_k, \ldots, 1/\mu_n)^\top$.

$$U_\alpha = n U_1(\alpha, \mathcal{L}) + \sum_{k=1}^{n} \log\left(\frac{T_k}{\mu_k}\right) - (\alpha - \mathcal{L}) \sum_{k=1}^{n} \frac{\mu_k}{T_k}, \tag{6.7}$$

where

$$U_1(\alpha, \mathcal{L}) = -\Psi(\mathcal{L} - \alpha) + \Psi(-\alpha) - \log(-\alpha - 1) + \frac{\alpha}{(-\alpha - 1)},$$

and $\Psi(x) = \mathrm{d}\log\Gamma(x)/\mathrm{d}x$ is the digamma function.

$$U_{\mathcal{L}} = n U_2(\alpha, \mathcal{L}) + \sum_{k=1}^{n} \log\left(\frac{z_k}{T_k}\right) + (\alpha - \mathcal{L}) \sum_{k=1}^{n} \frac{z_k}{T_k}, \tag{6.8}$$

where

$$U_2(\alpha, \mathcal{L}) = 1 + \log(\mathcal{L}) + \Psi(\mathcal{L} - \alpha) - \Psi(\mathcal{L}).$$

In this case, the Fisher information matrix is given by

$$\boldsymbol{K}(\boldsymbol{\theta}) = \mathrm{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top \partial \boldsymbol{\theta}}\right] = \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{K}_{\boldsymbol{\beta}\alpha} & \boldsymbol{K}_{\boldsymbol{\beta}\mathcal{L}} \\ \bullet & \boldsymbol{K}_{\alpha\alpha} & \boldsymbol{K}_{\alpha\mathcal{L}} \\ \bullet & \bullet & \boldsymbol{K}_{\mathcal{L}\mathcal{L}} \end{bmatrix}, \tag{6.9}$$

where $\boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \alpha \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X}$, $\boldsymbol{K}_{\boldsymbol{\beta}\alpha} = c_2 \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{\mu}^*$, $\boldsymbol{K}_{\alpha\alpha} = n c_3$, $\boldsymbol{K}_{\mathcal{L}\mathcal{L}} = n c_4$, $\boldsymbol{K}_{\boldsymbol{\beta}\mathcal{L}} = \boldsymbol{K}_{\alpha\mathcal{L}} = \boldsymbol{0}$, $\boldsymbol{W} \coloneqq \operatorname{diag}\{\omega_1, \ldots, \omega_t, \ldots, \omega_n\}$, and

$$\omega_k = \left(\frac{\mathcal{L}}{\mathcal{L} - \alpha + 1}\right) \frac{1}{\mu_k^2} \frac{1}{\{g'(\mu_k)\}^2}.$$

A detailed discussion about the elements of $\boldsymbol{K}(\boldsymbol{\theta})$ and c_2, c_3, and c_4 is presented in Appendix G. We observe that the parameters $\boldsymbol{\beta}$ and α are not orthogonal (in sense that θ_i and θ_j in $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^\top$

are orthogonal, termed by "$\theta_i \perp \theta_j$", if only if $\mathrm{E}[\partial^2 \ell(\boldsymbol{\theta})/\partial\theta_i\partial\theta_j] = 0$ for all $i \neq j$ [242]), which is the opposite to what is verified in the class of generalized linear regression models (see McCullagh and Nelder [239]). On the other hand, $\boldsymbol{\beta} \perp \mathcal{L}$ and $\alpha \perp \mathcal{L}$.

For a sufficiently large sample and under the usual regularity conditions [110], it follows that

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_{p+1}\left(\mathbf{0},\, \boldsymbol{K}^{-1}(\boldsymbol{\theta})\right),$$

where $\widehat{\boldsymbol{\theta}}$ is the MLE for $\boldsymbol{\theta}$, $\xrightarrow[n\to\infty]{\mathcal{D}}$ represents *convergence in distribution*, and $\mathcal{N}_p(\boldsymbol{\mu},\, \boldsymbol{\Sigma})$ indicates the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. In this chapter, we obtain a closed-form expression for $\boldsymbol{K}^{-1}(\boldsymbol{\theta})$ of the $\mathcal{G}_I^0$ regression model:

$$\boldsymbol{K}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{K}^{\beta\beta} & \boldsymbol{K}^{\beta\alpha} & \boldsymbol{K}^{\beta\mathcal{L}} \\ \bullet & \boldsymbol{K}^{\alpha\alpha} & \boldsymbol{K}^{\alpha\mathcal{L}} \\ \bullet & \bullet & \boldsymbol{K}^{\mathcal{L}\mathcal{L}} \end{bmatrix},$$

with elements given in Appendix G.

The asymptotic multivariate normal distribution can be also used to construct asymptotic confidence intervals, which are useful for testing significance of some sub-models and comparing some special sub-models using the likelihood ratio, score and Wald statistics. In particular, the asymptotic confidence interval at $100\,(1-\varepsilon)\%$ for the ith component of $\boldsymbol{\theta}$, θ_i, is given by

$$\mathrm{ACI}_\varepsilon\left(\theta_j\right) = \left(\widehat{\theta}_j - z_{\varepsilon/2}\sqrt{\widehat{k}^{j,j}}\,;\, \widehat{\theta}_j + z_{\varepsilon/2}\sqrt{\widehat{k}^{j,j}}\right),$$

where $\widehat{k}^{j,j}$ denotes the jth element of the main diagonal of $\boldsymbol{K}^{-1}(\boldsymbol{\theta})$ and $z_{\varepsilon/2}$ is the $(1-\varepsilon/2)$ quantile of the standard normal distribution.

We obtain the MLE for α, $\mathcal{L}$, and $\boldsymbol{\beta}$ from the system $(U_{\boldsymbol{\beta}}, U_\alpha, U_{\mathcal{L}}) = (0,0,0)$. Hereby, they need to be obtained by numerically maximizing the log-likelihood function using iterative methods; for instance, BFGS (Broyden-Fletcher-Goldfarb-Shanno), Newton Raphson, etc [159]. When using the optimization algorithms we need to determine initial estimates for parameter vector $\boldsymbol{\beta}$. A suggestion is to use the ordinary least squares estimates for this parameter obtained from a linear regression model of the transformed responses $g(z_1), g(z_2), \ldots, g(z_n)$; i.e., $(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top \log(\boldsymbol{z})$,

where $\log(\boldsymbol{z}) := [\log(z_1), \ldots, \log(z_k), \ldots, \log(z_n)]^\top$. For preliminary estimation of α fixed the value $\mathcal{L}$, equaling score components from (6.1) to the null vector, it follows

$$U_\alpha = n[\Psi(-\alpha) - \Psi(L-\alpha) - \log(\gamma)] + \sum_{k=1}^{n} \log(T_k) = 0$$

and

$$U_\gamma = -\frac{n\alpha}{\gamma} + (\alpha - \mathcal{L}) \sum_{k=1}^{n} \frac{1}{T_k} = 0. \tag{6.10}$$

Through Equation (6.10), we have the initial guess of α given by $\alpha = \mathcal{L}\frac{S}{S-n/\gamma}$, where $S = \sum_{k=1}^{n} T_k^{-1}$, and γ can be obtained by the nonlinear equation

$$\frac{1}{n}\sum_{k=1}^{n} \log\left(\frac{T_K}{\gamma}\right) = \Psi\left(\mathcal{L}\frac{S}{S-n/\gamma}\right) - \Psi\left(\mathcal{L}\left(1 + \frac{S}{S-n/\gamma}\right)\right).$$

6.3.2 *Residual analysis*

An important step to propose a new regression model is the analysis of its assumptions based on residual-based quantities. To that end, we introduce two kinds of residuals for our proposal.

Residual analysis is often made in terms of ordinary residuals, standardized variants, or deviance residuals and has a twofold goal: identifying outliers (this step is named as sensibility or influential study) and/or checking assumptions of the model. In this case, the kth standardized residual is defined as:

$$r_k = \frac{Z_k - \widehat{\mu}_k}{\sqrt{\widehat{\mathrm{Var}}(Z_k)}},$$

where $\widehat{\mu}_k = g^{-1}(x_k^\top \widehat{\boldsymbol{\beta}})$ and

$$\widehat{\mathrm{Var}}(Z_k) = \widehat{\mu}_k^2 \left[\left(\frac{\widehat{\alpha}+1}{\widehat{\alpha}+2}\right)\frac{\mathcal{L}+1}{\mathcal{L}} - 1\right].$$

A second residual is the deviance residual according to [243, 244], which defined–adapting to the $\mathcal{G}_I^0$ regression– $\mathbf{D}_{dk}(\boldsymbol{z};\boldsymbol{\mu},\alpha,\mathcal{L}) = \mathrm{sign}(z_k - \mu_k)\sqrt{|r_{dk}|}$ as a residual, where

$$r_{dk} = 2[\ell(\boldsymbol{z};\boldsymbol{z},\alpha,\mathcal{L}) - \ell(\boldsymbol{z};\boldsymbol{\mu},\alpha,\mathcal{L})], = 2\left[\alpha\log\left(\frac{z_k}{\mu_k}T_k^\diamond\right) - \mathcal{L}\log\left(T_k^\diamond\right)\right],$$

with $T_k^{\diamond} = {}^{T_k}/_{z_k(-\alpha-1+\mathcal{L})}$ and $\text{sign}(z_k-\mu_k)$ is the signum function defined as

$$\text{sign}(\bullet) = \begin{cases} +, & \text{if } z_k > \mu_k, \\ 0, & \text{if } z_k = \mu_k, \\ -, & \text{if } z_k < \mu_k. \end{cases}$$

Hence, the standardized deviance residuals are defined by [245, p. 49]

$$SR_k = \frac{\mathbf{D}_{dk}(\boldsymbol{z};\boldsymbol{\mu},\alpha,\mathcal{L})}{S\sqrt{(1-h_{kk})}},$$

where $S = \sum_{k=1}^{n} {}^{\mathbf{D}_{dk}(\boldsymbol{z};\boldsymbol{\mu},\alpha,\mathcal{L})^2}/_{(n-p)}$.

When the distribution of the residuals is not known, normal plots with simulated envelopes are a helpful diagnostic tool (Atkinson [240], Section 4.2; Neter *et al.* [246], Section 14.6). The main aim is to make this graph as good as possible and, so for this, simulated envelopes can be used to decide if the observed residuals are consistent with the fitted model.

To build the envelope plot, we need of minimum and maximum values of the ν order statistics. We follow Atkinson [240, p. 36], who suggested $\nu = 19$ and the probability that a given absolute residual falls beyond the upper band-provided by the envelope-is approximately equal to 0.05. As rule of decision, if some absolute residuals are outside the limits provided by the simulated envelope, we will have to investigate that (these) residue(s). On the other hand, if a great amount of points is out of this range, we have evidence against the suitability of the fitted model.

6.3.3 *Influential measures*

This section tackles the proposal of influential tools for the $\mathcal{G}_I^0$ regression model. After writing about kind of residuals, it is required to define measures for identifying influential observations. In what follows, we derive four influential measures: a generalized leverage quantity, projection-based (known as hat) matrix, Cook's distance, and Difference in Fits (DFFITS).

Wei *et al.* [247] developed the generalized leverage, which is defined as

$$\mathbf{GL}(\widetilde{\boldsymbol{\theta}}) = \frac{\partial \widetilde{\boldsymbol{z}}}{\partial \boldsymbol{z}^{\top}}, \tag{6.11}$$

where $\boldsymbol{\theta}$ is a parameter vector such that $\mathrm{E}(\boldsymbol{z}) = \boldsymbol{\mu}(\boldsymbol{\theta})$, $\widetilde{\boldsymbol{\theta}}$ is an estimator of $\boldsymbol{\theta}$, and $\widetilde{\boldsymbol{z}} = \mu(\widetilde{\boldsymbol{\theta}})$. Let $\ell(\boldsymbol{\theta})$ be the $\mathcal{G}_I^0$ log-likelihood function having second order continuous derivatives with respect to $\boldsymbol{\theta}$ and $\boldsymbol{z}$. Let also $\widehat{\boldsymbol{\theta}}$ be the MLE for $\boldsymbol{\theta}$, assuming its existence and uniqueness follow, and $\widehat{\boldsymbol{z}}$ the predicted response vector. Wei *et al.* [247] showed the generalized leverage $n \times n$ matrix in (6.11) may be expressed by

$$\mathbf{GL}(\boldsymbol{\theta}) = D_{\boldsymbol{\theta}} \left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right]^{-1} D_{\boldsymbol{\theta} z},$$

evaluated at $\widehat{\boldsymbol{\theta}}$, where $D_{\boldsymbol{\theta}} = \dfrac{\partial \mu}{\partial \boldsymbol{\theta}^\top}$ and $D_{\boldsymbol{\theta} z} = \dfrac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{z}^\top}$.

Considering α as nuisance parameter and $\mathcal{L}$ fixed, then we obtain a closed-form expression for $\mathbf{GL}(\boldsymbol{\beta})$ in the $\mathcal{G}_I^0$ regression model. We have

(i)

$$D_{\boldsymbol{\beta}} = \boldsymbol{EX},$$

(ii)

$$-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \alpha \boldsymbol{X}^\top \boldsymbol{Q} \boldsymbol{X},$$

where $\boldsymbol{Q} = \operatorname{diag}\{q_1, \ldots, q_k, \ldots, q_n\}$ with

$$q_t = -\left\{ \left[\frac{1}{\mu_k^2} + \frac{c_1(-\alpha - 1)}{\alpha T_k^2} \right] + \left[\frac{1}{\mu_k} + \frac{c_1}{\alpha T_k} \right] \frac{g''(\mu_k)}{g'(\mu_k)} \right\} \frac{1}{\{g'(\mu_k)\}^2}$$

and

(iii)

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{z}^\top} = \alpha \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{T}^*,$$

where

$$\boldsymbol{T}^* = \frac{\mathcal{L} c_1}{\alpha} \operatorname{diag}\left\{ {}^1\!/_{T_1^2}, \ldots, {}^1\!/_{T_k^2}, \ldots, {}^1\!/_{T_n^2} \right\}.$$

Therefore, the following identity holds:

$$\mathbf{GL}(\boldsymbol{\beta}) = \boldsymbol{EX}(\boldsymbol{X}^\top \boldsymbol{Q} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{T}^*. \tag{6.12}$$

More details are presented in Appendix H.

Now, replacing the observed information matrix by the Fisher information matrix, the expression (6.12) collapses in

$$\mathbf{GL}_{\text{A}}(\boldsymbol{\beta}) = \boldsymbol{EX}(\boldsymbol{X}^\top \boldsymbol{WX})^{-1}\boldsymbol{X}^\top \boldsymbol{ET}^*.$$

Hence, the (k,k)-entry of $\mathbf{GL}_{kk}$ can be expressed as

$$\mathbf{GL}_{kk}(\boldsymbol{\beta}) = \omega_k x_k^\top (\boldsymbol{X}^\top \boldsymbol{WX})^{-1} x_k,$$

where

$$\omega_k = \frac{\partial^2 \ell_k(\boldsymbol{\theta})}{\partial \mu_k^2}\left(\frac{\partial \mu_k}{\partial T_k}\right)^2 = \left[\frac{\alpha}{\mu_k^2} + \frac{c_1(-\alpha-1)}{T_k^2}\right]\frac{1}{\{g'(\mu_k)\}^2}. \tag{6.13}$$

Pregibon [248] extended the projection/hat matrix from linear regression $\widehat{z}$ versus X under weights to generalized linear models, given as

$$\mathbf{H} = \boldsymbol{W}^{1/2}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{WX})^{-1}\boldsymbol{X}^\top \boldsymbol{W}^{1/2}.$$

The literature has suggested the use of the (k,k)-entry of $\mathbf{H}$, say h_{kk}, for detecting the presence of leverage points in linear generalized regression models [247, 249]. It is worth mentioning that–for large samples–$\mathbf{GL}$ and $\mathbf{H}$ coincide and, in the case of canonical bonding, this equality holds for any sample size.

We suppose that the parameters α and $\mathcal{L}$ are unknown and hence $\boldsymbol{\theta}^\top = (\boldsymbol{\beta}^\top, \alpha, \mathcal{L})$. Let $D_\theta = [(\mathbf{EX})^\top \quad \mathbf{0}^\top \quad \mathbf{0}^\top]^\top$, where $\mathbf{0}$ is an n-vector of zeros. Moreover,

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{z}^\top} = \begin{bmatrix} \alpha \boldsymbol{X}^\top \boldsymbol{ET}^* \\ \boldsymbol{a}^\top \\ \boldsymbol{b}^\top \end{bmatrix},$$

where $\boldsymbol{a} = (a_1, \ldots, a_k, \ldots, a_n)^\top$ with $a_k = {}^{\mathcal{L}[T_k + (\alpha - \mathcal{L})\mu_k]}/_{T_k^2}$, and $\boldsymbol{b} = (b_1, \ldots, b_k, \ldots, b_n)^\top$ with $b_t = {}^{1}/_{z_k} - {}^{\mathcal{L}}/_{T_k} + \frac{(\alpha-\mathcal{L})}{T_k}\left(1 - \frac{\mathcal{L} z_k}{T_k}\right)$. Finally, to determine $-{}^{\partial^2 \ell(\boldsymbol{\theta})}/_{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}$, we follow the same idea on which

was built the expression (6.9), and we have

$$-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{bmatrix} \alpha \boldsymbol{X}^\top \boldsymbol{Q} \boldsymbol{X} & \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{M} & \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{N} \\ \bullet & \boldsymbol{R} & \boldsymbol{P} \\ \bullet & \bullet & \boldsymbol{S} \end{bmatrix}, \tag{6.14}$$

where $\boldsymbol{M} = \text{diag}\{m_1, \ldots, m_k, \ldots, m_n\}$ with

$$m_k = \frac{1}{\mu_k} + \frac{1}{T_k}\left(2\alpha - \mathcal{L} + 1 + \frac{\mu_t c_1}{T_k}\right),$$

$\boldsymbol{N} = \text{diag}\{n_1, \ldots, n_k, \ldots, n_n\}$ with

$$n_k = \frac{(-\alpha - 1)}{T_k}\left[1 + \frac{(\alpha - \mathcal{L}) z_k}{T_k}\right],$$

$\boldsymbol{P} = \text{tr}(P^*)$, where $P^* = \text{diag}\{p_1^*, \ldots, p_k^*, \ldots, p_n^*\}$ with

$$p_k^* = \Psi^{(1)}(\mathcal{L} - \alpha) + \frac{1}{T_k}(z_k + \mu_k) + \frac{(\alpha - \mathcal{L})\mu_k z_k}{T_k^2},$$

$\boldsymbol{R} = \text{tr}(R^*)$, where $R^* = \text{diag}\{r_1^*, \ldots, r_k^*, \ldots, r_n^*\}$ with

$$r_k^* = -U_1^{(1)}(\alpha, \mathcal{L}) + \frac{\mu_k}{T_k}\left[2 + (\alpha - \mathcal{L})\frac{\mu_k}{T_k}\right],$$

and $\boldsymbol{S} = \text{tr}(S^*)$, where $S^* = \text{diag}\{s_1^*, \ldots, s_k^*, \ldots, s_n^*\}$ with

$$s_k^* = -\Psi^{(1)}(\mathcal{L} - \alpha) + \Psi^{(1)}(\mathcal{L}) - \frac{1}{\mathcal{L}} + \frac{z_k}{T_k}\left[2 + (\alpha - \mathcal{L})\frac{z_k}{T_k}\right].$$

Faced with this, using the strategy that was used to find the inverse $\boldsymbol{K}(\boldsymbol{\theta})$, and partitioning matrix (6.14) as

$$-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{bmatrix} \boldsymbol{A}^\star & \boldsymbol{B}^\star \\ \bullet & \boldsymbol{D}^\star \end{bmatrix},$$

we have that its inverse is given by

$$\left(-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right)^{-1} = \begin{bmatrix} \boldsymbol{A}^* & \boldsymbol{B}^* & \boldsymbol{C}^* \\ \bullet & \boldsymbol{D}^* & \boldsymbol{E}^* \\ \bullet & \bullet & \boldsymbol{F}^* \end{bmatrix},$$

where $\boldsymbol{A}^* = \left[(\alpha \boldsymbol{X}^\top \boldsymbol{Q}\boldsymbol{X})^{-1} + \boldsymbol{\zeta}^* \boldsymbol{\vartheta}^{*^{-1}} \boldsymbol{\zeta}^{*^\top}\right] \boldsymbol{\Upsilon}^* \boldsymbol{\Phi}^{*^{-1}} \boldsymbol{\Upsilon}^{*^\top}$, $\boldsymbol{B}^* = \left[-\boldsymbol{\zeta}^* \boldsymbol{\vartheta}^{*^{-1}}\right] \boldsymbol{\Upsilon}^* \boldsymbol{\Phi}^{*^{-1}} \boldsymbol{\Upsilon}^{*^\top}$,
$\boldsymbol{D}^* = \boldsymbol{\vartheta}^{*^{-1}} \boldsymbol{\Upsilon}^* \boldsymbol{\Phi}^{*^{-1}} \boldsymbol{\Upsilon}^{*^\top}$, $\boldsymbol{C}^* = \boldsymbol{\Upsilon}^* \boldsymbol{\Phi}_1^{*^{-1}} \boldsymbol{\Upsilon}^{*^\top}$, $\boldsymbol{E}^* = \boldsymbol{\Upsilon}^* \boldsymbol{\Phi}_2^{*^{-1}} \boldsymbol{\Upsilon}^{*^\top}$, $\boldsymbol{F}^* = \boldsymbol{\Phi}^{*^{-1}}$,
$\boldsymbol{\vartheta} = \boldsymbol{R} - (\boldsymbol{X}^\top \boldsymbol{E}\boldsymbol{N})^\top (\alpha \boldsymbol{X}^\top \boldsymbol{Q}\boldsymbol{X})^{-1} (\boldsymbol{X}^\top \boldsymbol{E}\boldsymbol{N})$, $\boldsymbol{\zeta}^* = (\alpha \boldsymbol{X}^\top \boldsymbol{Q}\boldsymbol{X})^{-1} (\boldsymbol{X}^\top \boldsymbol{E}\boldsymbol{N})$, $\boldsymbol{\Upsilon}^* = \boldsymbol{A}^{\star^{-1}} \boldsymbol{B}^\star$, $\boldsymbol{\Phi}^* = \boldsymbol{D}^\star -$
$\boldsymbol{B}^{\star^\top} \boldsymbol{\Upsilon}^*$, $\boldsymbol{\Phi}^* = [\boldsymbol{\Phi}_1^* \quad \boldsymbol{\Phi}_2^*]$. Therefore, it can be shown that

$$\mathbf{GL}(\boldsymbol{\beta}, \alpha, \mathcal{L}) = \mathbf{GL}(\boldsymbol{\beta}) + \boldsymbol{E}\boldsymbol{X}\boldsymbol{B}^* \boldsymbol{a}^\top + \boldsymbol{E}\boldsymbol{X}\boldsymbol{C}^* \boldsymbol{b}^\top,$$

where $\mathbf{GL}(\boldsymbol{\beta})$ is given in equation (6.12). We can observe that for a large α $\mathbf{GL}(\boldsymbol{\beta}, \alpha, \mathcal{L}) \to \mathbf{GL}(\boldsymbol{\beta})$.

Cook [250] developed an influence measure for each observation. This distance quantifies the effect of the tth observation by means of its square distance between $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\beta}}_{(t)}$ (where $\widehat{\boldsymbol{\beta}}_{(t)}$ is the parameter estimate without the tth observation), given by

$$\mathbf{D}_c(\boldsymbol{\theta}) = \frac{1}{p} (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(k)})^\top \boldsymbol{X}\boldsymbol{W}\boldsymbol{X} (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(k)}).$$

We can use the usual approximation to Cook distance given as

$$\mathbf{D}_c^*(\boldsymbol{\theta}) = \frac{h_{kk} r_k^2}{p(1-h_{kk})^2},$$

which avoid us to fit the model (n+1) times and it combines leverage and residuals (see Cook [251]).

We also determine a DFFITS diagnostic measure proposed by White [252], that serves to show how influential the point is in a linear regression. This measure is defined as

$$\mathbf{DFFITS}(\boldsymbol{\theta}) = (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(k)}) \sqrt{\left(\frac{h_{kk}}{1-h_{kk}}\right)}.$$

In order to illustrate proposed influential measures, we compute them for distributed $\mathcal{G}_I^0$ data following the configuration: $(\boldsymbol{\beta}^\top, \alpha, \mathcal{L}) = (1, 1, -50, 4)$ and $n = 100$. Results are presented in Fig. 6.3. In the Fig. 6.3(e), the plot of studentized residuals vs. predicted values are displayed. There are not patterns, as expected from a good adjustment. Fig. 6.3(b) shows the Cook distances vs. predicted values. We also compute values of (h_{kk}) vs. predicted values in Fig. 6.3(c). The graph of DFFITS vs. observed values is made in Fig. 6.3(d). Finally, the envelope plot of studentized residuals is done in Fig. 6.3(f). As expected, it indicates that the model is well fitted.

6.4 A simulation study

MLEs for $\mathcal{G}_I^0$ regression parameters given in (6.6), (6.7), and (6.8) do not have closed-form expressions and, therefore, iterative numerical methods are required. To that end, we did a pilot study to choose an iterative method. In this initial study, we elect four procedures: BFGS (for `Broyden-Fletcher-Goldfarb-Shanno`), CG (for `Conjugate Gradients`), NM (for `Nelder-Mead`) and SANN (for `Simulated ANNealing`). We consider $\boldsymbol{\beta} \in \{0.01, 0.01, 0.01\}$, $\alpha \in \{-50, -10, -5, -3\}$, $\mathcal{L} = 2$, and sample sizes $n \in \{20, 50, 100, 500\}$. For each combination $(\boldsymbol{\beta}, \alpha, \mathcal{L}, n)$, we generate 100 Monte Carlo replications on which proposed estimators are assessed. All computational manipulations were made using software `R` (see R Core Team [253]). Using root mean square error (RMSE) as figure of merit, the `CG` method obtained the best results. Detailed numerical results are presented in Appendix I. From now on, we use the `CG` procedure.

Now, we are in position of making a performance study. To reach this goal, we use 1,000 Monte Carlo replications on which distributed $\mathcal{G}_I^0(\alpha, e^{\beta_0+\beta_1 x_1+\beta_2 x_2}(-\alpha-1), \mathcal{L})$ data were generated under the following specifications: $n \in \{20, 50, 100, 500\}$, $\beta_0 = \beta_1 = \beta_2 \in \{0.01, 1, 2\}$, $\alpha \in \{-15, -10, -5, -3\}$, $\mathcal{L} \in \{1, 4, 8\}$, and $x_k \sim \mathcal{G}_I^0(\alpha, (-\alpha-1), \mathcal{L})$.

We use five comparison criteria: the absolute bias (Abias), root mean square error (RMSE), Akaike information criterion (AIC) and its corrected version AICc, and Bayesian information criterion (BIC). Results are presented in Fig. 6.4 and Table 6.1. In general, as expected, the quality of estimates improves with the increase of sample sizes. It is known that the increasing of $\mathcal{L}$ diminishes the effect of speckle noise over generated data. It is possible to observe that better results are associated to high values of $\mathcal{L}$. On the other hand, harder fits are related to smaller values of α for single look data ($\mathcal{L} = 1$); whereas, harder fits are function of higher values of α for multillok ($\mathcal{L} > 1$) data. Results of first and third scenarios are in the Appendix I.

6.5 Application to SAR imagery

In this section we apply the proposed $\mathcal{G}_I^0$ regression model to SAR imagery processing and compare its fit with those due to other seven models: exponential, Γ, Γ^{-1}, $\mathcal{N}$, $\mathcal{N}^{-1}$, Weibull,

Table 6.1: Parameter estimates using the model $\mu_k = e^{\beta_0+\beta_1 x_{1,k}+\beta_2 x_{2,k}}$, where $x_k \sim \mathcal{G}_I^0(\alpha, (-\alpha-1), \mathcal{L})$ with $\boldsymbol{\beta} = (1,1,1)$

	Parameter	n	$\alpha=-3$			$\alpha=-5$			$\alpha=-10$			$\alpha=-15$		
			Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE
$\mathcal{L}=1$	β_0	20	0.9924	0.0076	0.5568	0.9918	0.0082	0.4913	0.9946	0.0054	0.4656	1.0011	0.0011	0.4577
		50	1.0154	0.0154	0.3088	1.0152	0.0152	0.2816	1.0260	0.0260	0.2696	1.0062	0.0062	0.2737
		100	1.0073	0.0073	0.2174	1.0055	0.0055	0.1937	1.0041	0.0041	0.1840	1.0097	0.0097	0.1809
		500	1.0024	0.0024	0.0934	1.0082	0.0082	0.0842	1.0069	0.0069	0.0801	1.0028	0.0028	0.0793
	β_1	20	0.9640	0.0360	0.3559	0.9633	0.0367	0.3358	0.9648	0.0352	0.3229	0.9694	0.0306	0.3182
		50	0.9728	0.0272	0.1662	0.9789	0.0211	0.1576	0.9783	0.0217	0.1651	0.9816	0.0184	0.1668
		100	0.9927	0.0073	0.1006	0.9908	0.0092	0.1088	0.9996	0.0004	0.1047	0.9899	0.0101	0.1068
		500	0.9971	0.0029	0.0402	0.9953	0.0047	0.0437	0.9959	0.0041	0.0467	0.9977	0.0023	0.0462
	β_2	20	0.9752	0.0248	0.3474	0.9761	0.0239	0.3174	0.9644	0.0356	0.3159	0.9429	0.0571	0.3022
		50	0.9826	0.0174	0.1643	0.9849	0.0151	0.1610	0.9747	0.0253	0.1654	0.9825	0.0175	0.1676
		100	0.9898	0.0102	0.1067	0.9925	0.0075	0.1121	0.9866	0.0134	0.1096	0.9877	0.0123	0.1106
		500	0.9972	0.0028	0.0398	0.9978	0.0022	0.0434	0.9980	0.0020	0.0439	0.9988	0.0012	0.0454
	α	20	−6.2772	3.2772	4.9687	−8.2102	3.2102	5.0661	−11.0354	1.0354	3.8719	−15.1020	0.1020	3.0831
		50	−4.4825	1.4825	2.7875	−6.2796	1.2796	2.8566	−9.6713	0.3287	2.6133	−14.3900	0.6100	1.7358
		100	−3.9924	0.9924	2.1181	−6.0977	1.0977	2.6556	−9.6499	0.3501	2.7932	−14.2794	0.7206	2.0741
		500	−3.1718	0.1718	0.6680	−5.2236	0.2236	1.2107	−9.4019	0.5981	1.6048	−14.0003	0.9997	1.5528
$\mathcal{L}=4$	β_0	20	0.9640	0.0360	0.4204	0.9725	0.0275	0.3942	0.9959	0.0041	0.3689	1.0122	0.0122	0.3879
		50	0.9867	0.0133	0.2330	0.9780	0.0220	0.2240	1.0030	0.0030	0.2256	1.0045	0.0045	0.2321
		100	0.9859	0.0141	0.1642	0.9949	0.0051	0.1515	1.0038	0.0038	0.1569	0.9960	0.0040	0.1585
		500	1.0012	0.0012	0.0708	0.9978	0.0022	0.0650	1.0045	0.0045	0.0664	1.0050	0.0050	0.0640
	β_1	20	0.9928	0.0072	0.2853	1.0019	0.0019	0.2770	1.0089	0.0089	0.2669	0.9961	0.0039	0.2714
		50	1.0023	0.0023	0.1414	1.0036	0.0036	0.1434	0.9987	0.0013	0.1489	0.9941	0.0059	0.1490
		100	1.0029	0.0029	0.0860	0.9959	0.0041	0.0957	0.9990	0.0010	0.1062	1.0013	0.0013	0.1042
		500	1.0007	0.0007	0.0342	0.9996	0.0004	0.0394	1.0008	0.0008	0.0430	0.9990	0.0010	0.0438
	β_2	20	0.9940	0.0060	0.2883	0.9988	0.0012	0.2629	0.9853	0.0147	0.2498	0.9828	0.0172	0.2702
		50	0.9985	0.0015	0.1383	1.0117	0.0117	0.1402	0.9962	0.0038	0.1532	0.9992	0.0008	0.1526
		100	1.0030	0.0030	0.0845	1.0013	0.0013	0.0943	0.9971	0.0029	0.1052	1.0008	0.0008	0.1023
		500	0.9993	0.0007	0.0350	1.0020	0.0020	0.0397	0.9977	0.0023	0.0425	0.9962	0.0038	0.0442
	α	20	−4.9700	1.9700	3.3480	−7.4023	2.4023	4.0255	−10.7984	0.7984	3.5743	−14.8559	0.1441	2.8710
		50	−3.6625	0.6625	1.5581	−6.1744	1.1744	2.6286	−10.3427	0.3427	2.8598	−14.5587	0.4413	2.6102
		100	−3.3191	0.3191	0.8637	−5.5130	0.5130	1.5094	−9.5457	0.4543	1.4945	−14.1127	0.8873	1.2772
		500	−3.0427	0.0427	0.3141	−5.1051	0.1051	0.6645	−9.3749	0.6251	0.9552	−14.0128	0.9872	1.0874
$\mathcal{L}=8$	β_0	20	0.9257	0.0743	0.3826	0.9880	0.0120	0.3754	1.0113	0.0113	0.3826	1.0245	0.0245	0.3699
		50	0.9800	0.0200	0.2170	0.9938	0.0062	0.1997	0.9995	0.0005	0.2142	1.0167	0.0167	0.2217
		100	0.9958	0.0042	0.1509	0.9962	0.0038	0.1459	0.9920	0.0080	0.1501	1.0121	0.0121	0.1541
		500	0.9975	0.0025	0.0647	0.9977	0.0023	0.0611	1.0000	0.0000	0.0644	1.0047	0.0047	0.0671
	β_1	20	1.0117	0.0117	0.2583	0.9968	0.0032	0.2621	0.9877	0.0123	0.2673	0.9882	0.0118	0.2593
		50	1.0029	0.0029	0.1295	0.9998	0.0002	0.1287	0.9970	0.0030	0.1470	0.9906	0.0094	0.1534
		100	1.0019	0.0019	0.0856	1.0014	0.0014	0.0926	1.0050	0.0050	0.1008	0.9889	0.0111	0.1051
		500	1.0010	0.0010	0.0320	1.0018	0.0018	0.0375	1.0006	0.0006	0.0434	0.9979	0.0021	0.0466
	β_2	20	1.0175	0.0175	0.2586	0.9959	0.0041	0.2557	0.9944	0.0056	0.2565	0.9802	0.0198	0.2598
		50	1.0008	0.0008	0.1296	0.9981	0.0019	0.1359	1.0042	0.0042	0.1420	0.9941	0.0059	0.1513
		100	0.9963	0.0037	0.0813	1.0005	0.0005	0.0916	1.0022	0.0022	0.0984	1.0006	0.0006	0.1051
		500	1.0016	0.0016	0.0324	1.0009	0.0009	0.0378	1.0002	0.0002	0.0417	0.9984	0.0016	0.0431
	α	20	−4.6538	1.6538	2.9152	−7.0135	2.0135	3.4979	−11.0715	1.0715	3.2874	−14.9771	0.0229	2.8080
		50	−3.5156	0.5156	1.1705	−5.8053	0.8053	1.8376	−11.0096	1.0096	3.4952	−15.2506	0.2506	4.1899
		100	−3.2240	0.2240	0.6444	−5.3719	0.3719	1.1582	−9.9394	0.0606	1.6493	−14.3275	0.6725	1.5228
		536	−3.0425	0.0425	0.2523	−5.0635	0.0635	0.4597	−9.6555	0.3445	0.8068	−14.1784	0.8216	1.0722

power exponential, and EGB2 regression models. . We describe intensities due to HH and VV channels in terms of HV intensities. From initial descriptive analysis, HV data are well correlated with HH and VV data: 0.5486 and 0.4465, respectively. We compose two simple linear regression models under the link functions: (i) $\mu_{HH}(x_k) = \exp(\beta_0 + \beta_1 x_k)(HH \sim HV)$ and (ii) $\mu_{VV}(x_k) = \exp(\beta_0 + \beta_1 x_k)(VV \sim HV)$. In what follows, we aim to describe the SAR intensities obtained by the AIRSAR [205] sensor of scenes of San Francisco (USA) regions. Fig. 6.5 shows an intensity HH map of a San Francisco SAR region. To this framework, the $\mathcal{G}_I^0$ regression model is equipped by the following link function:

$$\log(\mu_k) = \beta_0 + \beta_1 x_k,$$

where the response variable is $Z_k \sim \mathcal{G}_I^0(\alpha, \mu_k(-\alpha - 1))$ and the roughness α is a nuisance parameter. To quantify the performance of considered fits we use three measures: mean absolute biases (MAB) and root mean squared error (RMSE) given by

$$\frac{1}{n}\sum_{k=1}^{n} |Z_k - \widehat{Z}_k|, \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^{n} (Z_k - \widehat{Z}_k)^2,$$

respectively. Table 6.3 shows values for MAB and RMSE, which indicate the $\mathcal{G}_I^0$ regression model may overcome considered competing models. The last panel of Fig. 6.7 shows standardized residuals of considered models. Our proposal represents the unique acceptable model by this criterion and the GoF measure given in Table 6.2.

MLEs for involved parameters and their standard errors are presented in Table 6.4. All models were well adjusted and presented a positive impact of HV over HH or VV. The fitted models are given by

$$\begin{aligned}(HH \sim HV): &\quad \widehat{\mu}_{HH}(x) = \exp(-2.524 + 17.207x)\\ (VV \sim HV): &\quad \widehat{\mu}_{VV}(x) = \exp(-2.551 + 14.608x),\end{aligned}$$

where x is the HV polarization channel. Then, if the HV channel is 0.01, we have the expected value for HH and VV given by 0.09519 and 0.09027, respectively.

Fig. 6.6 and 6.7 displays fitted line plot under the data set, a standardized residuals, Cook's distance, and the normal plot of deviance residuals with a simulated envelope (according to Atkinson (1985)). Results indicate the new model presents the best results.

Table 6.2: GoF measures for SAR image models based on actual data

	$HH \sim HV$				$VV \sim HV$			
Model	GD	AIC	CAIC	SBC	GD	AIC	CAIC	SBC
G_I^0	−462.20	−454.20	−454.00	−440.70	−486.70	−478.70	−478.50	−465.10
Exponential	−372.60	−368.60	−368.60	−361.80	−436.50	−432.50	−432.40	−425.70
Gamma	−460.10	−452.10	−451.90	−438.50	−485.20	−477.20	−477.00	−463.60
Inverse gamma	−429.20	−423.20	−423.10	−413.00	−457.10	−451.10	−450.90	−440.90
Normal	−290.20	−284.20	−284.00	−274.00	−284.30	−278.30	−278.20	−268.10
Inverse normal	−423.50	−417.50	−417.40	−407.30	−474.20	−468.20	−468.10	−458.10
Weibull	−447.70	−441.70	−441.60	−431.50	−476.00	−470.00	−469.80	−459.80
Power Exponential	−359.50	−353.50	−353.40	−343.30	−360.80	−354.80	−354.70	−344.60
EGB2	−390.80	−380.80	−380.50	−363.90	−427.20	−417.20	−416.90	−400.20

Table 6.3: Statistics MAB, MRAB and RMSE of the regression models

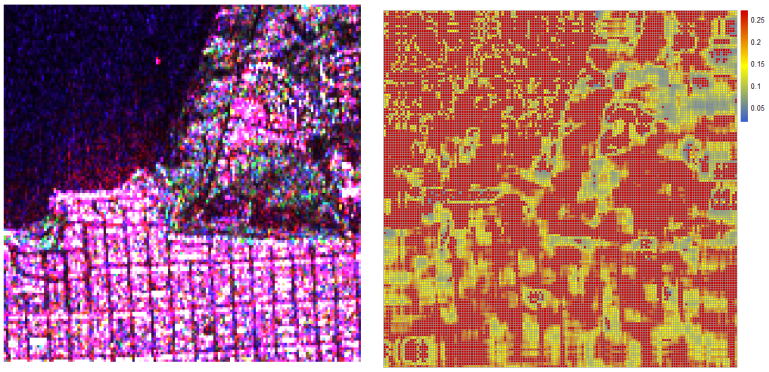
	$HH \sim HV$		$VV \sim HV$	
Model	MAB	RMSE	MAB	RMSE
$\mathcal{G}_I^0$	0.087	0.126	0.087	0.127
Exponential	0.096	0.135	0.096	0.137
Gamma	0.096	0.135	0.096	0.137
Inverse gamma	0.104	0.161	0.102	0.164
Normal	0.096	0.135	0.097	0.137
Inverse normal	0.090	0.131	0.099	0.139
Weibull	0.092	0.139	0.091	0.139
Power exponential	0.092	0.139	0.092	0.134
EGB2	0.131	0.186	0.137	0.192

Table 6.4: Parameter estimates using data on SAR imagery

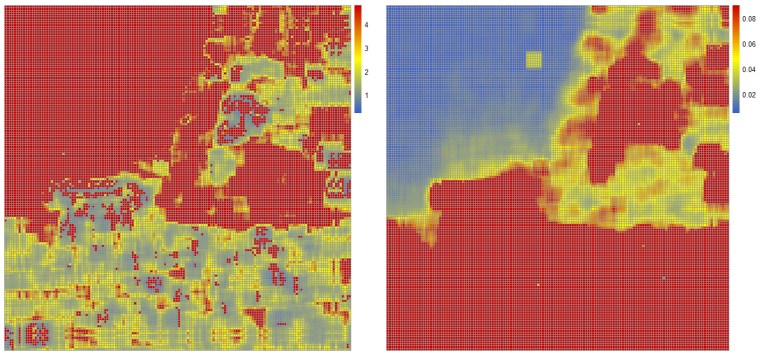
Model	Parameter	Estimate	Std. error	t stat	p-value
		$HH \sim HV$			
$\mathcal{G}_I^0$	β_0	−2.524	0.077	−32.800	0.000
	β_1	17.207	1.613	10.700	0.000
Exponential	β_0	−2.514	0.123	−20.510	0.000
	β_1	16.896	2.592	6.520	0.000
Gamma	β_0	−2.514	0.076	−33.000	0.000
	β_1	16.896	1.610	10.500	0.000
Inverse gamma	β_0	−3.344	0.079	−42.200	0.000
	β_1	17.615	1.512	11.700	0.000
Normal	β_0	−2.414	0.108	−22.290	0.000
	β_1	14.879	1.548	9.610	0.000
Inverse normal	β_0	−2.610	0.094	−27.900	0.000
	β_1	19.845	2.867	6.920	0.000
Weibull	β_0	−2.385	0.078	−30.700	0.000
	β_1	16.592	1.619	10.200	0.000
Power exponential	β_0	−2.649	0.023	−113.600	0.000
	β_1	15.904	0.192	82.800	0.000
EGB2	β_0	−4.265	0.557	−7.660	0.000
	β_1	23.989	5.068	4.730	0.000
		$VV \sim HV$			
$\mathcal{G}_I^0$	β_0	−2.551	0.100	−25.48	0.000
	β_1	14.608	1.985	7.36	0.000
Exponential	β_0	−2.598	0.122	−21.26	0.000
	β_1	15.340	2.579	5.95	0.000
Gamma	β_0	−2.598	0.087	−29.80	0.000
	β_1	15.340	1.841	8.33	0.000
Inverse gamma	β_0	−3.576	0.094	−38.26	0.000
	β_1	13.882	1.815	7.65	0.000
Normal	β_0	−2.510	0.126	−19.99	0.000
	β_1	13.542	1.839	7.36	0.000
Inverse normal	β_0	−2.670	0.105	−25.33	0.000
	β_1	17.501	3.065	5.71	0.000
Weibull	β_0	−2.509	0.089	−28.17	0.000
	β_1	15.637	1.869	8.37	0.000
Power exponential	β_0	−2.884	0.044	−65.00	0.000
	β_1	14.828	0.739	20.10	0.000
EGB2	β_0	−7.570	2.31	−3.28	0.001
	β_1	51.760	24.25	2.13	0.034

6.6 Conclusion

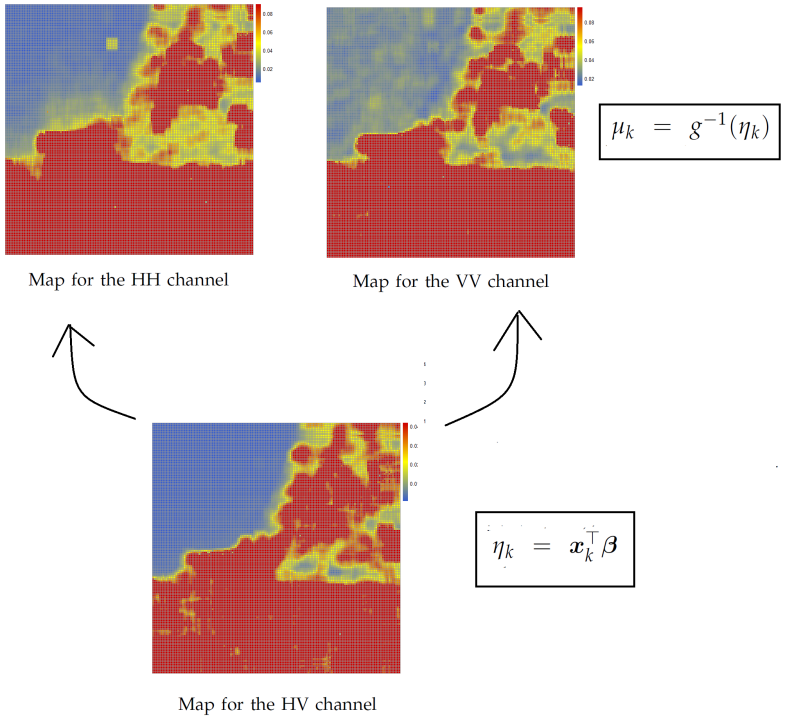
In this chapter, we have proposed a new $\mathcal{G}_I^0$ regression model to describe conditionally SAR intensities. The proposed model showed to be very flexible for several SAR textures, comparatively to exponential, Γ, Γ^{-1}, $\mathcal{N}$, $\mathcal{N}^{-1}$, Weibull, power exponential, and EGB2 regression models. We have used the maximum likelihood procedure for $\mathcal{G}_I^0$ regression parameters and proposed a closed-form expression for the Fisher information matrix. The performance of estimates was assessed by Monte Carlo simulation study. Subsequently, some diagnostic and influential techniques were proposed: a generalized leverage and Cook measures and two kind of residuals. Finally, an application of the $\mathcal{G}_I^0$ regression model to actual SAR data was made. Results pointed out our proposal may be useful in SAR image processing.



(a) San Francisco image (b) Map of γ

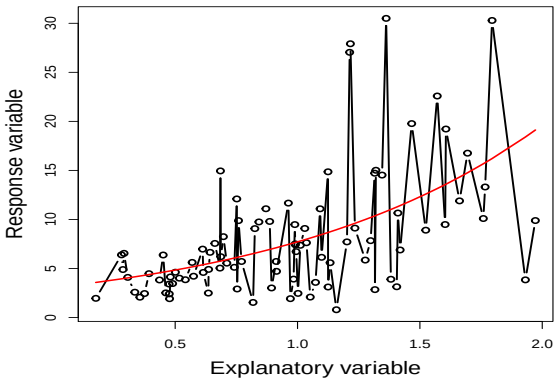


(c) Map of α (d) Map of μ

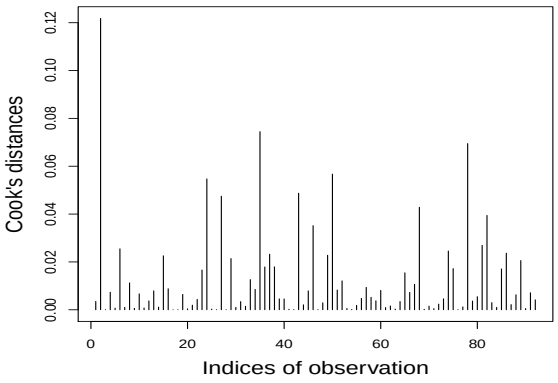


(e) Illustration of regression

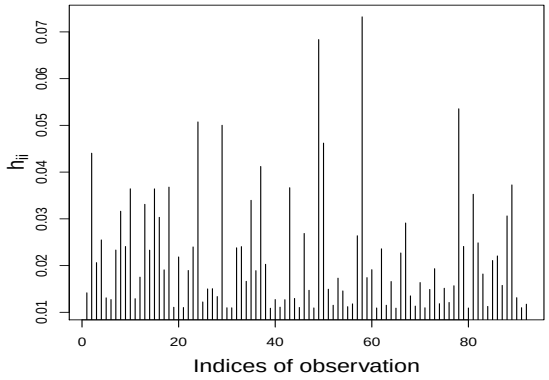
Figure 6.2: The effect of roughness over estimates for (α, γ, μ) and illustration for the regression model.



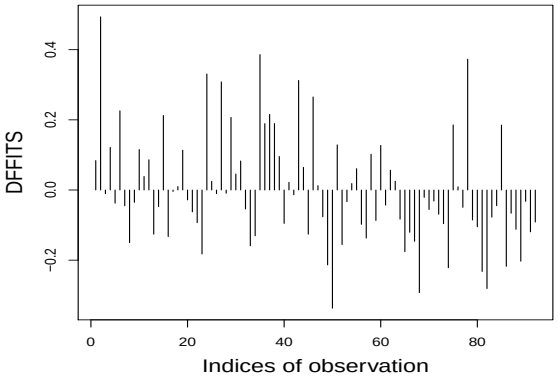
(a) A fitted line plot



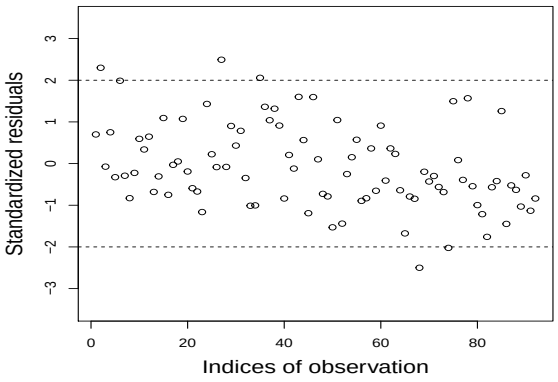
(b) Cook's distance versus predicted values



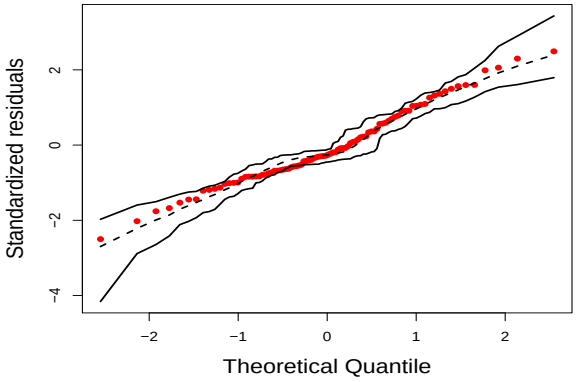
(c) Leverage (h_{kk}) versus predicted values



(d) DFFITS versus indices of observation

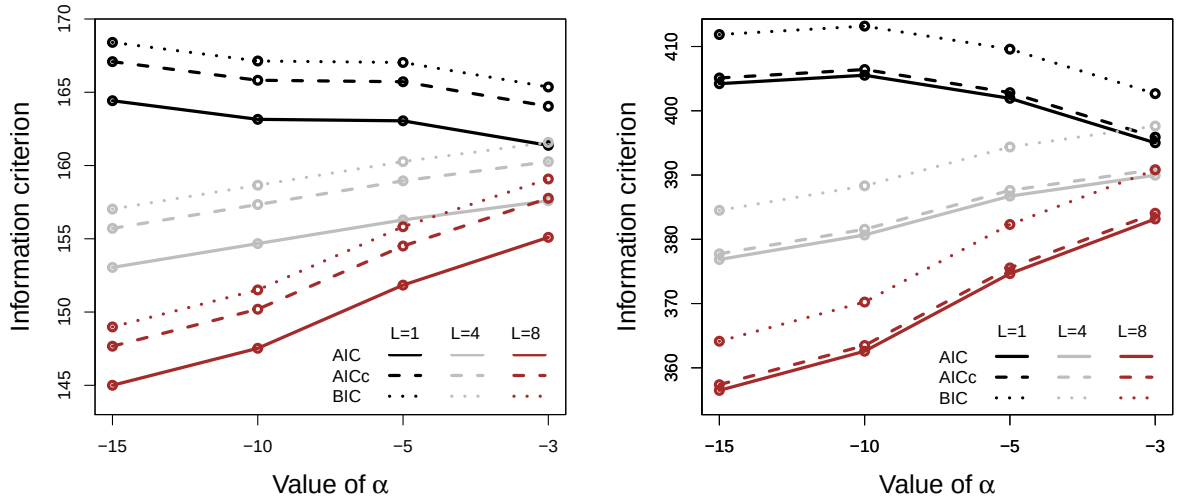


(e) Standardized residuals versus predicted values

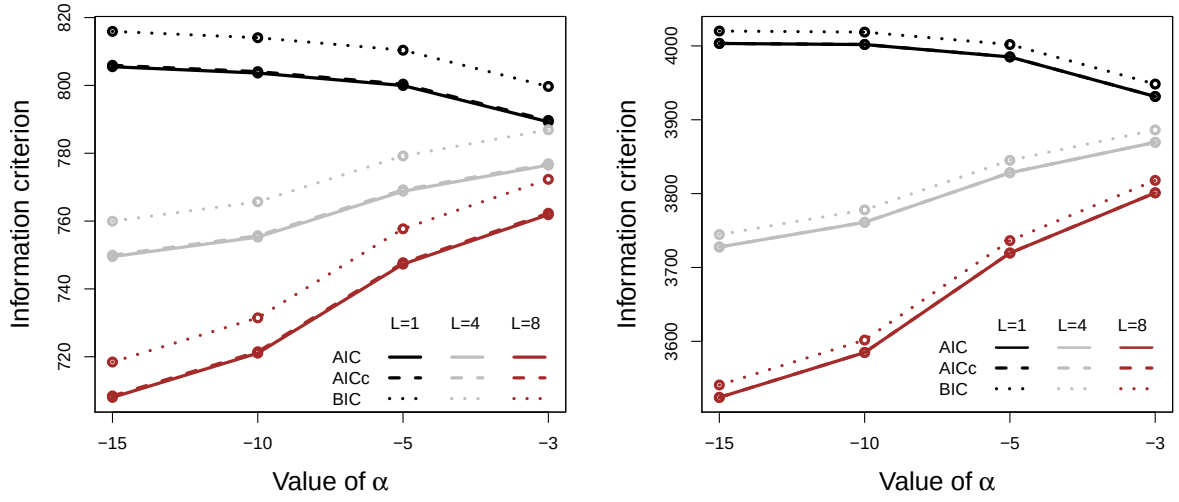


(f) Normal Q-Q Plot

Figure 6.3: The upper (a) is a fitted line plot, (b), (c), and (d) panels plots the Cooks distance, leverage and DFFITS against the predicted. The (e) panels is standardized residuals versus predicted and the another panel displays Normal Q-Q plot of absolute deviance residuals with a simulated envelope.



(a) $n = 20$ (b) $n = 50$



(c) $n = 100$ (d) $n = 500$

Figure 6.4: Measure AIC, AICc and BIC to sample size $n \in \{20, 50, 100, 500\}$, respectively.

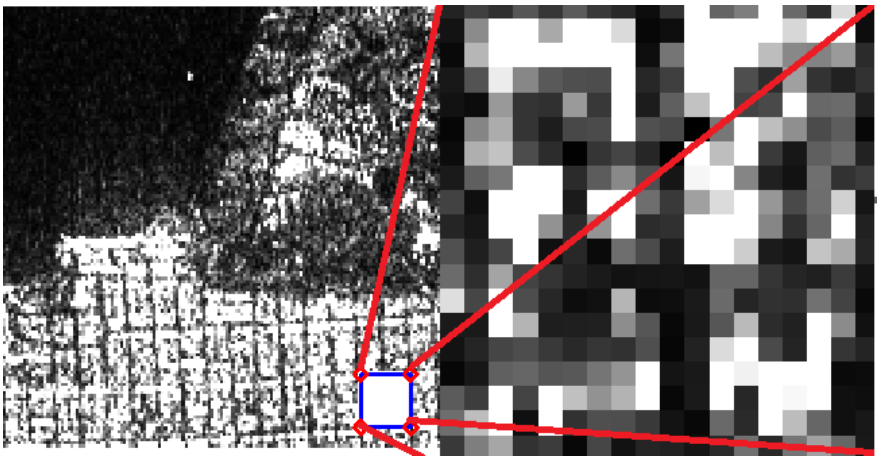


Figure 6.5: PolSAR image with selected region San Francisco (USA).

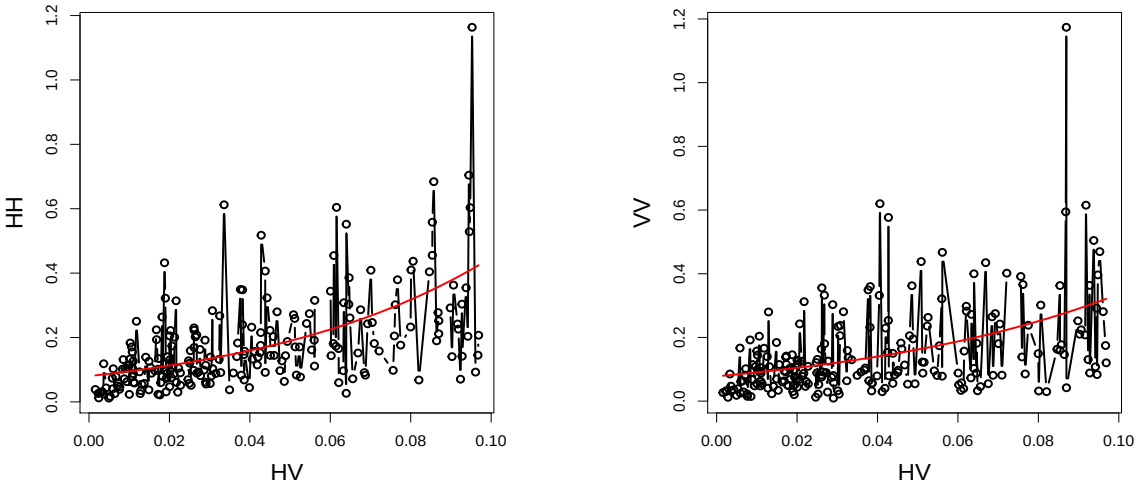


Figure 6.6: A fitted line plot for urban SAR imagery data of the two models (left panels plots are of the model I and right band plots are of the second model).

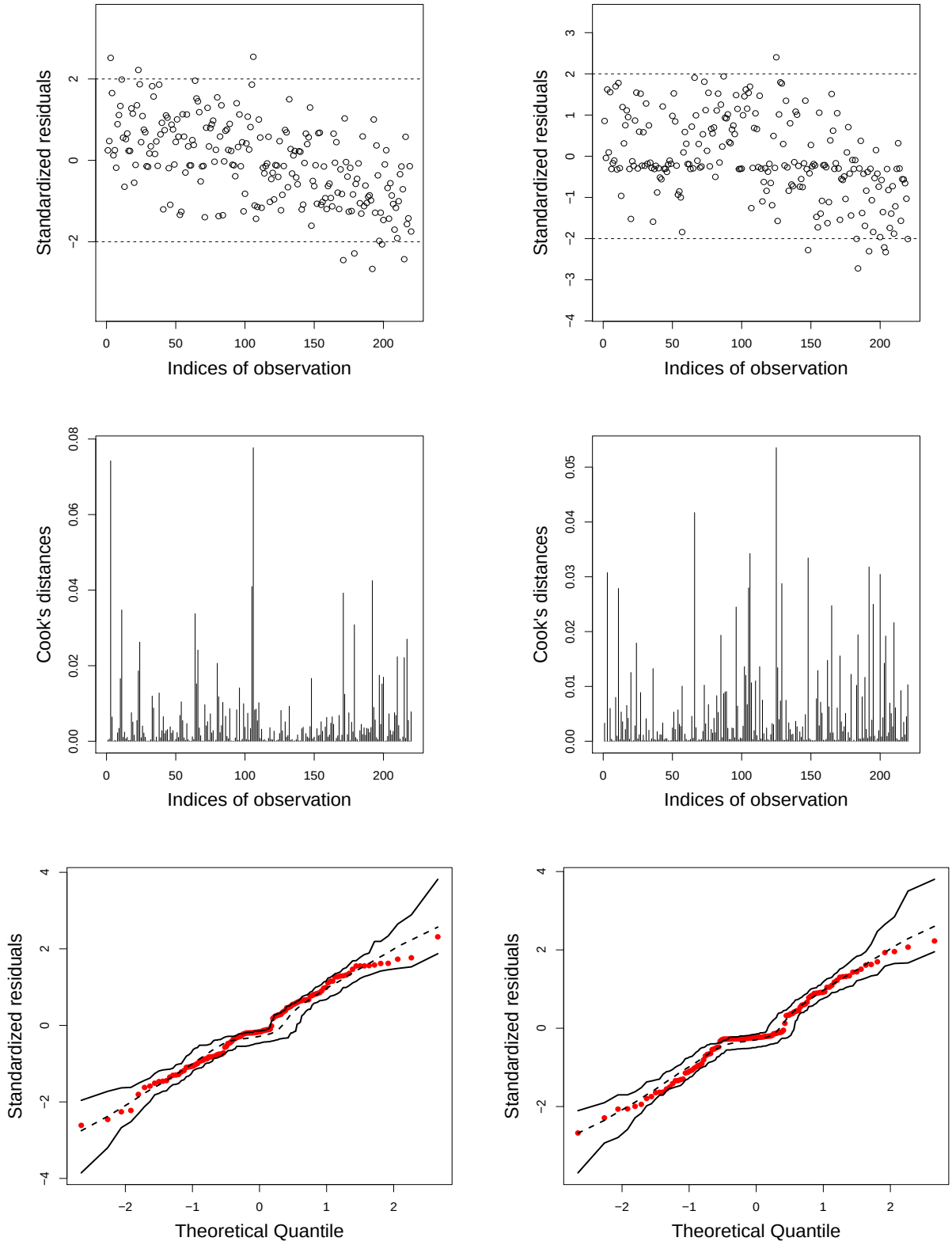


Figure 6.7: The three diagnostic plots for urban SAR imagery data of the two models (left panels plots are of the model I and right band plots are of the second model).

7 Conclusions

In this chapter, we present some concluding remarks and directions for future works.

7.1 Concluding Remarks

In Chapters 2 and 3 of this thesis, we have shown that Mellin-kind statistics can be successfully applied like GoF measures and estimation criteria for beta-G models in the survival analysis context. Expressions for the MT of BW, BF, BKw, and BLL distributions have been derived as some analytical contributions. Subsequently, the LCs related to these expressions have been used as entries to the Hotelling's T^2 statistic and their asymptotic distributions were studied. Our proposals have been illustrated by means of five experiments with actual lifetime data. Moreover, LC-based estimators have been derived for the BW model. These estimators presented results better than those due to the ML method for both synthetic and real data.

In Chapter 4, a model for multimodal SAR intensities has been introduced, called CTPC model. We derived some of its mathematical properties; such as, cf and a distance measure (which was used as a possible GoF tool). Further, we have proposed two estimation methods for CTPC parameters: ML and QD estimators. A Monte Carlo study has indicated that QD estimators have furnished best results. An application to real SAR data was made and results have shown evidence that the CTPC model may outperform other seven classic distributions for modeling SAR intensities: gamma, Weibull, generalized gamma, $\mathcal{K}$, $\mathcal{G}^0$, beta generalized normal distributions, and CTPG.

In Chapter 5, a new class of distributions for speckled data has been proposed: named the G-$\mathcal{G}$ family. Two new distributions which generalize the classical (in sense of SAR imagery) $\mathcal{G}_I^0$ and $\mathcal{K}$ models have been detailed among its special cases. We derived some mathematical properties of these new distributions and proposed their LC-based diagrams for model selection. Applications to real data have been considered and results have demonstrated that the new models are capable

of outperforming the $\mathcal{G}_I^0$ and $\mathcal{K}$ distributions.

In Chapter 6, we proposed a new $\mathcal{G}_I^0$ regression model to describe speckled data from SAR imagery. We derived some of its mathematical properties: Fisher information matrix, residual-kind quantities, and influential measures. Point and interval estimation procedures have been developed for $\mathcal{G}_I^0$ regression model parameters via maximum likelihood. Results from simulated and actual studies have illustrated that our proposal may be useful as conditioned models for SAR intensity data.

7.2 Directions for future researches

This thesis has also raised several points, which require deeper studies. Some of them are presented as follows:

- ★ to revisit the first paper looking for developments for censored data;
- ★ to derive the Mellin transform, log-cumulants diagram, and estimators based in log-cumulants for other families of distributions (as those presented in Fig. 5.4);
- ★ to provide estimators via probability weigth moment and log-cumulants for the beta-log-logistic model;
- ★ to introduce a new regression model with response variables given by the $\mathcal{K}$ and CTPC distributions;
- ★ to develop generalized hypothesis tests for the $\mathcal{G}_I^0$ regression model;

and

- ★ to implement a computational package in `R` for regression models having $\mathcal{G}_I^0$ and $\mathcal{K}$ response variables.

References

[1] NASCIMENTO, A. ao D. C. et al. The gamma extended Weibull family of distributions. *Journal of Statistical Theory and Applications*, v. 13, p. 1–16, 2014.

[2] CORDEIRO, G. M. et al. The gamma extended Weibull distribution. *Journal of Statistical Distributions and Applications*, v. 3, p. 1–19, 2016.

[3] AL-BABTAIN, A. A.; MEROVCI, F.; ELBATAL, I. The McDonald exponentiated gamma distribution and its statistical properties. *Springer*, v. 4, p. 1–22, 2015.

[4] ALIZADEH, M. et al. Odd-Burr generalized family of distributions with some applications. *Journal of Statistical Computation and Simulation*, Taylor and Francis, v. 87, p. 367–389, 2016.

[5] PARANAIBA, P. F. et al. The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Wind Engineering and Industrial Aerodynamics*, v. 83, p. 2117–2143, 2013.

[6] LEE, C.; FAMOYE, F.; OLUMOLADE, O. Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods*, v. 6, p. 173–186, 2007.

[7] MARSHALL, A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, v. 84, p. 641–652, 1997.

[8] CORDEIRO, G. M.; CASTRO, M. de. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, v. 87, p. 1–16, 2010.

[9] CORDEIRO, G. M. et al. The McDonald normal distribution. *Pakistan Journal of Statistics and Operation Research*, v. 8, p. 301–329, 2012.

[10] ALZAATREH, A.; LEE, C.; FAMOYE, F. A new method for generating families of continuous distributions. *METRON*, v. 71, p. 63–79, 2013.

[11] ALJARRAH, M. A.; LEE, C.; FAMOYE, F. On generating T-X family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, v. 1, p. 1–17, 2014.

[12] CORDEIRO, G. M.; ORTEGA, E. M. M.; RAMIRE, T. G. A new generalized Weibull family of distributions: mathematical properties and applications. *Journal of Statistical Distributions and Applications*, v. 2, p. 1–25, 2015.

[13] PAKYARI, R.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test for progressively type-II censored data. *IEEE transactions on reliability*, v. 61, p. 238–244, 2012.

[14] VIVEROS, R.; BALAKRISHNAN, N. Interval estimation of parameters of life from progressively censored data. *Technometrics*, v. 36, p. 84–91, 1994.

[15] CHEN, G.; BALAKRISHNAN, N. A general purpose approximate good-ness-of-fit test. *Journal of Quality Technology*, v. 27, p. 154–161, 1995.

[16] LINHART, H.; ZUCCHINI, W. *Model Selection*. : John Wiley & Sons, 1986.

[17] MCQUARRIE, A. D. R.; TSAI, C.-L. *Regression and time series model selection*. : World Scientific Publishing Co. Pte. Ltd., 1998.

[18] PEARSON, K. Memoir on skew variation in homogeneous material. *Philos. Trans. Roy. Soc. A 186*, p. 343–414, 1895.

[19] JOHNSON, N. L.; KOTZ, S.; N, B. *Continuous Univariate Distributions*. second. : Wiley, 1994. 784p p. ISBN 978-0-471-58495-7.

[20] DELIGNON, Y.; GARELLO, R.; HILLION, A. Statistical modelling of ocean SAR images. *IEE Proceedings Radar, Sonar and Navigation*, v. 144, p. 348–354, 1997.

[21] VOGEL, R. M.; FENNESSEY, N. M. L moment diagrams should replace product moment distribution. *Water resources research*, v. 29, p. 1745–1752, 1993.

[22] NICOLAS, J. M. Introduction aux statistiques de deuxième espèce: applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar. *Traitement du Signal*, v. 19, p. 139–167, 2002.

[23] LI, H.-C. et al. On the empirical-statistical modeling of SAR images with generalized gamma distribution. *IEEE Journal of Selected Topics in Signal Processing*, v. 5, p. 386–397, 2011.

[24] JOHNSON, N. L.; KEMP, A. W.; KOTZ, S. *Univariate Discrete Distributions*. 3. ed. : Wiley-Interscience, 2005.

[25] ANFINSEN, S. N.; DOULGERIS, A. P.; ELTOFT, T. Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, v. 47, p. 3795–3809, 2009. ISSN 0196-2892.

[26] ANFINSEN, S. N.; DOULGERIS, A. P.; ELTOFT, T. Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform. *IEEE Transactions on Geoscience and Remote Sensing*, v. 49, p. 2764–2781, 2011. ISSN 0196-2892.

[27] KRYLOV, V. et al. On the method of logarithmic cumulants for parametric probability density function estimation. *IEEE Signal Processing Society*, v. 22, p. 3791–3806, 2013.

[28] DENG, X.; LÓPEZ-MARTÍNEZ, C. Higher order statistics for texture analysis and physical interpretation of polarimetric SAR data. *IEEE Geoscience and Remote Sensing Letters*, v. 13, p. 912–916, 2016.

[29] DIAS, L. *Ensaios sobre Inferência estatística em imagens de radar e dificuldades numéricas em estimação por máxima verossimilhança*. Tese (Thesis) — Universidade Federal de Pernambuco, February 2013.

[30] KARLIS, D.; XEKALAKI, E. Mixed Poisson distributions. *International Statistical Review*, v. 73, p. 35–58, 2005.

[31] CUADRAS, C. M. Geometrical understanding of the Cauchy distribution. *QUESTIIÓ*, v. 26, p. 283–287, 2002.

[32] COHEN, M. P. Sample means of independent standard Cauchy random variables are standard Cauchy: A new approach. *Mathematical Association of America*, v. 119, p. 240–244, 2012.

[33] MANDELBROT, B. B.; GOMORY, R. E. *Fractals and Scaling in Finance*. 1. ed. : Springer-Verlag, 1997.

[34] JENKINS, F. A.; WHITE, H. E. *Fundamentals of Optics*. 4. ed. : McGraw-Hill, 1981.

[35] CORDEIRO, G. M. et al. The generalized odd half-Cauchy family of distributions: properties and applications. *Communication in Statistics Theory and Methods*, p. 1–24, 2015.

[36] ALSHAWARBEH, E.; FAMOYE, F.; LEE, C. Beta-Cauchy distribution: some properties and applications. *Journal of Statistical Theory and Applications*, v. 12, p. 378–391, 2013.

[37] HUANG, W.-J.; CHEN, Y.-H. Generalized skew-Cauchy distribution. *Statistics & Probability Letters*, v. 77, p. 1137–1147, 2007.

[38] TAHIR, M. H. et al. The Weibull-power Cauchy distribution: Model, properties and applications. *Journal of Mathematics and Statistics*, p. 1–26, 2000.

[39] ALZAATREH, A. et al. The generalized Cauchy family of distributions with applications. *Journal of Statistical Distributions and Applications*, v. 3, p. 1–16, 2016.

[40] BOSCH, P. HCM property and the half-Cauchy distribution. *Probability and Mathematical Statistics*, v. 35, p. 191–200, 2015.

[41] GHOSH, B.; BASU, B.; O´MAHONY, M. Bayesian time-series model for shortterm trafic forecasting. *Journal of Transportation Engineering-Asce*, v. 133, p. 180–189, 2007. ISSN 0003-9519.

[42] CORDEIRO, G. M.; LEMONTE, A. J. The beta-half-Cauchy distribution. *Journal of Probability and Statistics*, 2011.

[43] GRUSHKA, E. Characterization of exponentially modified gaussian peaks in chromatography. *Analytical Chemistry*, v. 44, p. 1733–1738, 1972.

[44] GOLUBEV, A. Exponentially modified Gaussian (EMG) relevance to distributions related to cell proliferation and differentiation. *Journal of Theoretical Biology*, v. 262, p. 257–266, 2010.

[45] TEICH, M. C.; DIAMENT, P. Multiply stochastic representations for K distributions and their Poisson transforms. *Journal of the Optical Society of America A*, v. 6, p. 80–91, 1989.

[46] REVFEIM, K. J. A. An initial model of the relationship between rainfall events and daily rainfalls. *Journal of Hydrology*, v. 75, p. 357–364, 1984.

[47] PANJER, H. H. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin*, v. 12, p. 22–26, 1981.

[48] THOMPSON, C. S. Homogeneity analysis of rainfall series: An application of the use of a realistic rainfall model. *International Journal of Climatology*, v. 4, p. 609–619, 1984.

[49] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. Disponível em: <https://www.R-project.org/>.

[50] DEENICK, E. K.; GETT, A. V.; HODGKIN, P. D. Stochastic model of T cell proliferation: A calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. *The Journal of Immunology*, v. 170, p. 4963–4972, 2003.

[51] CINTRA, R. J. et al. Beta generalized normal distribution with an application for SAR image processing. *Statistics: A Journal of Theoretical and Applied Statistics*, 2012.

[52] CINTRA, R. J.; FRERY, A. C.; NASCIMENTO, A. D. C. Parametric and nonparametric tests for speckled imagery. *Pattern Analysis and Applications*, v. 16, p. 141–161, 2013.

[53] EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, v. 31, p. 497–512, 2002.

[54] GUPTA, R. D.; KUNDU, D. Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, v. 41, p. 173–188, 1999.

[55] ZOGRAFOS, K.; BALAKRISHNAN, N. On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, v. 6, p. 344–362, 2009.

[56] RISTIC, M. M.; BALAKRISHNAN, N. The gamma exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, v. 82, p. 1191–1206, 2012.

[57] ALZAATREH, A.; LEE, C.; FAMOYE, F. A new method for generating families of continuous distributions. *METRON*, v. 71, p. 63–79, 2013.

[58] CHABERT, M.; TOURNERET, J. Y. Bivariate Pearson distributions for remote sensing imagens. *IEEE International Geoscience and Remote Sensing Symposium*, p. 4038–4041, 2011.

[59] NAGAHARA, Y. A method of simulating multivariate nonnormal distributions by the Pearson distribution system and estimation. *Computational Statistics & Data Analysis*, v. 47, p. 1–29, 2004.

[60] ANFINSEN, S. N.; ELTOFT, T. Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 49, p. 2281–2295, 2011. ISSN 0196-2892.

[61] KHAN, S.; GUIDA, R. Application of Mellin kind statistics to polarimetric G distribution for SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 52, p. 3513–3528, 2013.

[62] NICOLAS, J. M.; MARUANI, A. Lower order statistics: a new approach for probability density functions defined on $\mathbb{R}_+$. In: *IEEE Signal Processing Conference*. 2000. p. 1–4.

[63] WEIBULL, W. *A statistical theory of the strength of material*. : The Royal Swedish Royal Institute for Engineering Research, 1939. 45 p.

[64] FRÉCHET, M. Sur la loi de probabilité de l´écart maximum. *Annales de la Société Polonaise de Mathématique*, v. 6, p. 30–38, 1927.

[65] KHAN, M. S.; PASHA, G. R.; PASHA, A. H. Theoretical analysis of inverse Weibull distribution. *WSEAS Transactions on Mathematics*, v. 7, p. 93–116, 2008.

[66] KUMARASWAMY, P. Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology*, v. 46, p. 79–88, 1980.

[67] SHAH, B. K.; DAVE, P. H. A note on log-logistic distribution. *Journal of the MS University of Baroda*, v. 12, p. 15–20, 1963.

[68] FAMOYE, F.; LEE, C.; OLUMOLADE, O. The beta-Weibull distribution. *Journal Statistics Theory Application*, v. 4, p. 121–136, 2005.

[69] NADARAJAH, S.; KOTZ, S. The beta Gumbel distribution. *Mathematical Problems in Engineering*, v. 4, p. 323–332, 2004.

[70] CARRASCO, J. M. F.; FERRARI, S. L. P.; CORDEIRO, G. M. A new generalized Kumaraswamy distribution. 2010.

[71] LEMONTE, A. J. The beta log-logistic distribution. *Brazilian Journal of Probability and Statistics*, v. 28, p. 313–332, 2014.

[72] AKINSETE, A.; FAMOYE, F.; LEE, C. The beta-Pareto distribution. *Statistics*, v. 42, p. 547–563, 2008.

[73] CORDEIRO, G. M. A simple formula based on quantiles for the moments of beta generalized distributions. *Journal of Statistical Computation and Simulation*, v. 83, p. 1932–1943, 2013.

[74] ACHCAR, J. A.; COELHO-BARROS, E. A.; CORDEIRO, G. M. Beta generalized distributions and related exponentiated models: A bayesian approach. *Brazilian Journal of Probability and Statistics*, v. 27, p. 1–19, 2013.

[75] ZOGRAFOS, K. *On some beta generated distributions and their maximum entropy characterization: The beta-Weibull distribution*. New Jersey: Nova Science Publishers, 2008. 237260 p.

[76] GREENWOOD, J. A. et al. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources*, v. 15, p. 1049–1054, 1979.

[77] CAIZA, P. D. T.; UMMENHOFER, T. General probability weighted moments for the three-parameter Weibull distribution and their application in S-N curves modelling. *International Journal of Fatigue*, v. 1, p. 1533–1538, 2011.

[78] MAHDI, S.; ASHKAR, F. Exploring generalized probability weighted moments, generalized moments and maximum likelihood estimating methods in two-parameter Weibull model. *Journal of Hydrology*, v. 285, p. 62–75, 2004.

[79] WHALEN, T. M.; SAVAGE, G. T.; JEONG, G. D. An evaluation of the self-determined probability-weighted moment method for estimating extreme wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, v. 92, p. 219–239, 2004.

[80] MAHDI, S.; CENAC, M. Estimating parameters of Gumbel distribution using the methods of moments, probability weighted moments and maximum likelihood. *Revista de Matemática: Teoría y Aplicaciones*, v. 12, p. 151–156, 2005.

[81] ASHKAR, F.; MAHDI, S. Fitting the log-logistic distribution by generalized moments. *Journal of Hydrology*, v. 328, p. 694–703, 2006.

[82] SINGH, V. P.; GUO, H.; YU, F. X. Parameter estimation for three-parameter log-logistic distribution (LLD3) by Pome. *Stochastic Hydrology and Hydraulics*, v. 7, p. 163–177, 1993.

[83] ASHKAR, F.; MAHDI, S. Comparison of two fitting methods for the log-logistic distribution. *Water Resources Research*, v. 39, p. 1–8, 2003.

[84] SINGH, V. P.; GUO, H. Parameter estimation for two-parameter log-logistic distribution (LLD2) by maximum entropy. *Taylor & Francis*, v. 12, p. 343–357, 2013.

[85] HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. Estimation of the generalized extreme-value distribution by the method of probability weighted moments. *American Statistical Association and the American Society for Quality Control*, v. 27, p. 251–261, 1985.

[86] SOUKISSIAN, T. H.; TSALIS, C. The effect of the generalized extreme value distribution parameter estimation methods in extreme wind speed prediction. *Springer*, v. 1, p. 1777–1809, 2015.

[87] JEON, J.-J.; KIM, Y.-O.; KIM, Y. Expected probability weighted moment estimator for censored flood data. *Advances in Water Resources*, v. 1, p. 933–945, 2011.

[88] DIEBOLT, J. et al. Improving probability-weighted moment methods for the generalized extreme value distribution. *Statistical Journal*, v. 6, p. 33–50, 2008.

[89] RIBEREAU, P.; GUILOU, A.; NAVEAU, P. Estimating return levels from maxima of non-stationary random sequences using the generalized pwm method. *Copernicus Publications on behalf of the European Geosciences Union and the American Geophysical Union*, v. 15, p. 1033–1039, 2008.

[90] RASMUSSEN, P. F. Generalized probability weighted moments: Application to the generalized Pareto distribution. *Water Resources Research*, v. 37, p. 1745–1751, 2001.

[91] CAEIRO, F.; GOMES, M. I. Semi-parametric tail inference through probability-weighted moments. *Journal of Statistical Planning and Inference*, v. 141, p. 937–950, 2011.

[92] HOSKING, J. R. M.; WALLIS, J. J. R. Parameter and quantile estimation for the generalized pareto distribution. *Taylor & Francis, Ltd. on behalf of American Statistical Association and American Society for Quality*, v. 29, p. 339–349, 1987.

[93] BERMUDEZ, P. d. Z.; KOTZ, S. Parameter estimation of the generalized Pareto distributiona-part II. *Journal of Statistical Planning and Inference*, v. 140, p. 1374 – 1388, 2010.

[94] BERMUDEZ, P. d. Z.; KOTZ, S. Parameter estimation of the generalized Pareto distribution-part I. *Journal of Statistical Planning and Inference*, v. 140, p. 1353–1373, 2010.

[95] MERCY, J.; M., K. Estimation of the generalized lambda distribution from censored data. *Brazilian Journal of Probability and Statistics*, v. 24, n. 1, p. 42–56, 2008.

[96] AVERBUCH, A. et al. Fast and accurate polar Fourier transform. *Applied and Computational Harmonic Analysis*, v. 21, p. 145–167, 2006.

[97] OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.

[98] BILINGSLEY, P. *Probability and Measure*. 4. ed. : Wiley, 2016.

[99] CASTELLARES, F.; MONTENEGRO, L. C.; CORDEIRO, G. M. The beta log-normal distribution. *Journal of Statistical Computation and Simulation*, v. 83, p. 203–228, 2013.

[100] COLOMBO, S. *Les transformations de Mellin et de Hankel: applications a la physique mathematique*. : Centre national de la recerche scientifique, 1972. (Monographies du Centre d'etudes mathematiques en vue des applications. B, Methodes de calcul).

[101] HAKTANIR, T.; BOZDUMAN, A. A study on sensitivity of the probability-weighted moments method on the choice of the plotting position formula. *Journal of Hydrology*, v. 168, p. 265–281, 1995.

[102] DENG, J.; PANDEY, M. D. Cross entropy quantile function estimation from censored samples using partial probability weighted moments. *Journal of Hydrology*, v. 363, p. 18–31, 2008.

[103] DENG, J.; PANDEY, M. D. Estimation of minimum cross-entropy quantile function using fractional probability weighted moments. *Probabilistic Engineering Mechanics*, v. 24, p. 43–50, 2007.

[104] DENG, J.; PANDEY, M. D.; GU, D. Extreme quantile estimation from censored sample u-sing partial cross-entropy and fractional partial probability weighted moments. *Structural Safety*, v. 31, p. 43–54, 2009.

[105] WANG, Q. J. Unbiased estimation of probability weighted moments and partial probability weighted moments from systematic and historical flood information and their application to estimating the GEV distribution. *Journal of hydrology*, v. 120, p. 115–124, 1990.

[106] HAKTANIR, T.; BOZDUMAN, A. A study on sensitivity of the probability-weighted moments method on the choice of the plotting position formula. *Journal of Hydrology*, v. 168, p. 265–281, 1995.

[107] ASHKAR, F.; BOBéE, B. The generalized method of moments as applied to problems of flood frequency analysis: Some practical results for the log-Pearson type 3 distribution. *Journal of Hydrology*, v. 90, p. 199–217, 1987.

[108] BOBéE, B.; ASHKAR, F. The generalized method of moments applied to the LP3 distribution. *Journal of Hydraulic Engineering*, v. 114, p. 899–909, 1988.

[109] BOBéE, B.; ASHKAR, F. *The Gamma Family and Derived Distributions Applied in Hydrology*. : Water Resources Publications, 1991. 203 p.

[110] BICKEL, P. J.; DOKSUM, K. A. *Mathematical Statistics: Basic Ideas and Selected Topics*. 2. ed. : Prentice-Hall. 532 p.

[111] POULARIKAS, A. D. *The Transforms and Applications Handbook*. Third edition. : CRC Press LLC, 2000. 911 p. (Electrical Engineering Handbook).

[112] ABRAMOWITZ, M.; STEGUN, I. A. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. : Dover, 1964.

[113] HARDY, G. H.; LITTLEWOOD, J. E. Contributions to the theory of the Riemann Zeta-function and the theory of the distribution of primes. *Acta Mathematica*, v. 41, p. 119–196, 1916.

[114] STUART, A.; ORD, K. *Kendall´s Advanced Theory of Statistics, Distribution Theory*. 6. ed. : Distribution Theory, 2010. 700 p.

[115] LEHMANN, E. L.; CASELLA, G. *Theory of Point Estimation*. 2. ed. : Springer, 1998.

[116] MOSER, G.; ZERUBIA, J.; SERPICO, S. B. SAR amplitude probability density function estimation based on a generalized Gaussian model. *IEEE Transactions on image processing*, v. 15, p. 1429–1442, 2006.

[117] TISON, C. et al. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 42, p. 2046–2057, 2004. ISSN 0196-2892.

[118] ELTOFT, T.; ANFINSEN, S. N.; DOULGERIS, A. P. A multitexture model for multilook polarimetric radar data. *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2011.

[119] NICOLAS, J. M. *Application de la transformée de Mellin: Étude des lois statistiques de l´imagerie cohérente*. Paris, France, 2017. 238 p.

[120] RODRIGUES, F. A. A. et al. SAR image segmentation using the roughness information. *IEEE Geoscience and Remote Sensing Letters*, v. 13, p. 132–136, 2016.

[121] GAO, G.; QIN, X.; ZHOU, S. Modeling SAR images based on a generalized gamma distribution for texture component. *Progress In Electromagnetics Research*, v. 137, p. 669–685, 2013.

[122] CHENG, J. et al. An improved scheme for parameter estimation of G0 distribution model in high-resolution SAR images. *Progress In Electromagnetics Research*, v. 134, p. 23–46, 2013.

[123] FRERY, A. C. et al. A model for extremely heterogeneous clutter. *IEEE transactions on geoscience and remote sensing*, v. 35, p. 648–659, 1997.

[124] CUI, S.; SCHWARZ, G.; DATCU, M. A comparative study of statistical models for multilook SAR images. *IEEE Geoscience and Remote Sensing Letters*, v. 11, p. 1752–1756, 2014.

[125] OLIVER, C.; QUEGAN, S. *Understanding Synthetic Aperture Radar Images*. 2. ed. : SciTech, 2005.

[126] GAMBINI, J. et al. Parameter estimation in SAR imagery using stochastic distances and asymmetric kernels. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 8, p. 365–375, 2015.

[127] ADEPOJU, K. A.; CHUKWU, A. U.; WANG, M. The beta power exponential distribution. *Journal of Statistical Science and Application*, v. 2, p. 37–46, 2014.

[128] ANDERSON, T. W. *An introduction to multivariate statistical analysis*. : Wiley, 2003. 752 p. (Wiley series in probability and statistics).

[129] CORDEIRO, G. M.; LEMONTE, A. J. The beta Laplace distribution. *Statistics and Probability Letters*, v. 81, p. 973–982, 2011.

[130] NADARAJAH, S.; GUPTA, A. K. The beta Fréchet distribution. *Far East Journal of Theoretical Statistics*, v. 14, p. 15–24, 2004.

[131] BUSTOS, O. H.; LUCINI, M. M.; FRERY, A. C. M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, v. 1, p. 105–114, 2002.

[132] CASELLA, G.; BERGER, R. L. *Statistical Inference*. : Duxbury Press, 2002.

[133] CORDEIRO, G. M.; SILVA, G. O.; ORTEGA, E. M. M. The beta-Weibull geometric distribution. *Statistics*, Taylor & Francis, v. 47, p. 817–834, 2013.

[134] CORDEIRO, G. M.; LEMONTE, A. J. The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis*, v. 1, p. 1445–1461, 2011.

[135] LINDSEY, J. K. *Parametric Statistical Inference*. : Oxford Science Publications, 1996.

[136] CORDEIRO, G. M.; CRIBARI-NETO, F. *An Introduction to Bartlett Correction and Bias Reduction*. : SpringerBriefs in Statistics, 2014.

[137] HARVILLE, D. A. *Matrix Algebra From a Statistician's Perspective*. : Springer, 1997.

[138] RENCHER, A. C.; SCHAALJE, G. B. *Linear Models in Statistics*. 2. ed. : John Wiley & Sons, 2008. 688 p.

[139] RAO, C.; TOUTENBURG, H.; SHALABH HEUMANN, C. *Linear Models and Generalizations: Least Squares and Alternatives*. : Springer, 2008.

[140] SEBER, G. A. F. *A Matrix Handbook for Statisticians*. 1st. ed. New York, NY, USA: Wiley-Interscience, 2007.

[141] HÄRDLE, W. K.; SIMAR, L. *Applied Multivariate Statistical Analysis*. : Springer, 2007.

[142] NICHOLS, M. D.; PADGETT, W. J. A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, v. 22, p. 141–151, 2006.

[143] CORDEIRO, G. M.; PESCIM, R. R.; ORTEGA, E. M. M. The Kumaraswamy generalized Half-normal distribution for skewed positive data. *Journal of Data Science*, v. 10, p. 195–224, 2012.

[144] JØ RGENSEN, B. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. 1982. (Lecture Notes in Statistics, v. 9).

[145] ALVEN, W. H. V. *Reliability engineering by ARINC*. 1964. 821–826 p.

[146] CORDEIRO, G. M.; NADARAJAH, S.; ORTEGA, E. M. M. The Kumaraswamy Gumbel distribution. *Statistical Methods and Applications*, v. 21, p. 139–168, 2012.

[147] CORDEIRO, G. M.; ORTEGA, E. M.; NADARAJAH, S. The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 2010.

[148] YONG, T. *Extended weibull distributions in reliability engineering*. Tese (Doutorado) — National University of Singapore, 2004.

[149] ALZAATREH, A.; LEE, C.; FAMOYE, F. A new method for generating families of continuous distributions. *METRON*, v. 71, p. 63–79, 2013.

[150] HIROSE, H. Maximum likelihood parameter estimation in the extended Weibull distribution and its applications to breakdown voltage estimation. *IEEE Transactions on Dielectrics and Electrical Insulation*, v. 9, p. 524–536, 2002.

[151] SONG, K. Globally convergent algorithms for estimating generalized gamma distributions in fast signal and image processing. *IEEE Transactions on Image Processing*, v. 17, p. 1233–1250, 2008.

[152] LAWLESS, J. F. Inference in the generalized gamma and log gamma distribution. *Technometrics*, v. 22, p. 409–419, 1980.

[153] WINGO, D. R. Computing maximum-likelihood parameter estimates of the generalized gamma distribution by numerical root isolation. *IEEE Transactions on Reliability*, v. 36, p. 586–590, 1987.

[154] HAGER, H. W.; BAIN, L. J. Theory and method inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, v. 65, p. 1601–1609, 1970.

[155] RHIND, A. Tables to facilitate the computation of the probable errors of the chief constants of skew frequency distributions. *Biometrika*, v. 7, p. 127–147, 1909.

[156] PEARSON, K. Mathematical contributions to the theory of evolution, xix: Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society A*, v. 216, p. 429–457, 1916.

[157] CORDEIRO, G. M.; SIMAS, A. B.; STOVSI'C, B. D. Closed form expressions for moments of the beta Weibull distribution. *Annals of the Brazilian Academy of Sciences*, v. 83, p. 357–373, 2011.

[158] CORDEIRO, G. M.; NADARAJAH, S.; ORTEGA, E. M. M. General results for the beta Weibull distribution. *Journal of Statistical Computation and Simulation*, v. 83, p. 1082–1114, 2013.

[159] NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. New York: Springer, 1999.

[160] NADARAJAH, S.; GUPTA, A. K. The beta fréchet distribution. *Far East Journal of Theoretical Statistics*, v. 14, p. 15–24, 2004.

[161] VASCONCELOS, J. M.; CINTRA, R. J.; NASCIMENTO, A. D. C. Goodness-of-fit measures based on the Mellin transform for beta generalized lifetime data. *(Submitted)*.

[162] BILODEAU, M.; BRENNER, D. *Theory of Multivariate Statistics*. New York: Springer, 1999.

[163] KENDALL, M. G. *Kendall´s Advanced Theory of Statistics, Distribution Theory*. 5. ed. 1987.

[164] HOTELLING, H. *Multivariate Quality Control in Techniques of Statistical Analysis*. : McGraw-Hill, 1947.

[165] LEHMANN, E. L. *Elements of Large-Sample Theory*. : Springer Texts in Statistics, 1999.

[166] MEEKER, W. Q.; ESCOBAR, L. A. *Statistical Methods for Reliability Data*. : John Wiley, 1998.

[167] MURTHY, D. N. P.; XIE, M.; JIANG, R. *Weibull Models*. : John Wiley & Sons, 2004.

[168] SMITH, R. L.; NAYLOR, J. C. A comparison of maximum likelihood and bayesian estimators for the three-parameter Weibull distributuion. *Applied Statistics*, v. 36, p. 358–369, 1987.

[169] ALSHAWARBEH, E.; FAMOYE, F.; LEE, C. Beta-Cauchy distribution: some properties and applications. *Journal of Statistical Theory and Applications*, v. 12, p. 378–391, 2013.

[170] TAHIR, M. H. et al. The Weibull-power Cauchy distribution: Model, properties and applications. *Journal of Mathematics and Statistics*, v. 0, p. 1–26, 2000.

[171] ALZAATREH, A. et al. The generalized cauchy family of distributions with applications. *Journal of Statistical Distributions and Applications*, v. 3, p. 1–16, 2016.

[172] GHOSH, I. The Kumaraswamy- half-Cauchy distribution: Properties and applications. *Journal of Statistical Theory and Applications*, v. 13, p. 1–14, 2013.

[173] CORDEIRO, G. M. Bias correction. In: LOVRIC, M. (Ed.). *International Encyclopedia of Statistical Science*. : Springer Berlin Heidelberg, 2011. p. 148–152. ISBN 978-3-642-04898-2.

[174] TEICH, M. C.; DIAMENT, P. Multiply stochastic representations for k distributions and their Poisson transforms. *Journal of the Optical Society of America A*, v. 6, p. 80–91, 1989.

[175] DELIGNON, Y.; PIECZYNSKI, W. Modeling non-Rayleigh speckle distribution in SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 40, p. 1430–1435, 2002.

[176] YI-XIANGLU et al. Directionlet-based bayesian filter for SAR image despeckling. *Procedia Engineering*, v. 15, p. 2788–2792, 2011.

[177] GUOZHONG, C.; XINGZHAO, L. Cauchy pdf modelling and its application to SAR image despeckling. *Journal of Systems Engineering and Electronics*, v. 19, p. 717–721, 2008.

[178] GAO, Q. et al. Directionlet-based denoising of SAR images using a Cauchy model. *Signal Processing*, v. 93, p. 1056–1063, 2013.

[179] LI, Y.; LI, J. Oil spill detection from SAR intensity imagery using a marked point process. *Remote Sensing of Environment*, v. 114, p. 15902–1601, 2010.

[180] ZHAO, Q. H.; LI, Y.; WANG, Y. SAR image segmentation with unknown number of classes combined voronoi tessellation and RJMMC algorithm. In: ISPRS ANNALS OF THE PHOTOGRAMMETRY. 2016. III-7, p. 12–19.

[181] STACY, E. W. A generalization of the gamma distribution. *Annals of Mathematical Statistics*, v. 33, p. 1187–1192, 1962.

[182] JAKEMAN, E.; PUSEY, P. N. A model for non-Rayleigh sea echo. *IEEE Transactions on Antennas and Propagation*, v. 24, p. 806 – 814, 1976.

[183] LEE, J. S. et al. K-distribution for multi-look processed polarimetric SAR imagery. *Geoscience and Remote Sensing Symposium, 1994. IGARSS 94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, v. 4, p. 2179–2181, 1994.

[184] SILVA, J. W. A. *Distribuição Composta Poisson Truncada-Gama: propriedades e métodos de estimação*. Dissertação (Dissertação) — Universidade Federal de Pernambuco, 2016. Dissertação.

[185] FINE, T. L. *Probability and Probabilistic Reasoning for Electrical Engineering*. 1. ed. : Pearson, 2005. 688 p. (Prentice Hall).

[186] ATHREYA, K. B.; LAHIRI, S. N. *Measure Theory and Probability Theory*. 1. ed. : Springer, 2006. 625 p.

[187] ROUSSAS, G. G. *A Course in Mathematical Statistics*. 2. ed. : Academic Press, 1997. 593 p.

[188] NAVARRETE, V. E. d. L. S. *Um novo método para compor distribuições: uma análise da classe G-Poisson*. Dissertação (Dissertação) — Universidade Federal de Pernambuco, 2013.

[189] ARAúJO, R. L. B. d. *Família Composta Poisson-Truncada: Propriedades e aplicações*. Dissertação (Dissertação) — Universidade Federal de Pernambuco, 2015.

[190] DAVID, F. N.; JOHNSON, N. L. The truncated Poisson. *Biometrics*, v. 8, p. 275–285, 1952.

[191] GROPARU-COJOCARU, I.; DORAY, L. G. Inference for the generalized normal laplace distribution. *Communications in Statistics - Simulation and Computation*, v. 42, p. 1989–1997, 2013.

[192] PRESS, S. j. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, v. 67, p. 842–846, 1972.

[193] PAULSON, A. S.; HOLCOMB, E. W.; LEITCH, R. A. The estimation of the parameters of the stable laws. *Biometrika*, v. 62, p. 163–170, 1975.

[194] KNIGHT, J. L.; YU, J. Empirical characteristic function in time series estimation. *Econometric Theory*, v. 18, p. 691–721, 2002.

[195] BESBEAS, P.; MORGAN, B. J. T. Integrated squared error estimation of Cauchy parameters. *Statistics & Probability Letters*, v. 55, p. 397–401, 2001.

[196] MATSUI, M.; TAKEMURA, A. Empirical characteristic function approach to goodness-of-fit tests for the Cauchy distribution with parameters estimated by mle or eise. *Annals of the Institute of Statistical Mathematics*, v. 57, p. 183–199, 2005.

[197] YU, J. Empirical characteristic function estimation and its applications. *Econometric reviews*, v. 23, p. 93–123, 2004.

[198] PRESS, W. H. et al. *Numerical Recipes in C: The Art of Scientific Computing*. 3. ed. : Cambridge University Press, 2007.

[199] THALER, H. *Nonparametric Density Estimation and the Empirical Characteristic Function*. Dissertação (PhD dissertation) — State University of New York at Buffalo, Dept. of Statistics, State University of New York at Buffalo, Dept. of Statistics, 1974.

[200] GROPARU-COJOCARU, I. *Quadratic distance methods applied to generalized normal Laplace distribution*. Tese (Doutorado) — Université de Montréal, 2007.

[201] LUONG, A.; DORAY, L. G. General quadratic distance methods for discrete distributions definable recursively. *Insurance: Mathematics and Economics*, n. 30, p. 255–267, 2002.

[202] LUONG, A.; DORAY, L. G. Inference for the positive stable laws based on a special quadratic distance. *Statistical Methodology*, n. 6, p. 147–156, 2009.

[203] BARINGHAUS, L.; HENZE, N. A consistent test for multivariate normality based on the empirical characteristic function. *Metrika*, v. 35, p. 339–348, 1988.

[204] JIMÉNEZ-GAMERO, M. D. et al. Goodness-of-fit tests based on empirical characteristic functions. *Computational Statistics and Data Analysis*, v. 53, p. 3957–3971, 2009.

[205] LEE, J.-S.; POTTIER, E. *Polarimetric Radar Imaging: From Basics to Applications*. : CRC press, 2009. 438 p.

[206] OLIVER, C.; QUEGAN, S. *Understanding Synthetic Aperture Radar Images*. : SciTech Publishing, 1998. (The SciTech Radar and Defense series).

[207] GAO, G. Statistical modeling of SAR images: A survey. *Sensors*, v. 10, p. 775–795, 2010.

[208] LI, H.-C. et al. On the empirical-statistical modeling of SAR images with generalized gamma distribution. *IEEE Journal of Selected Topics in Signal Processing*, v. 5, p. 386–397, 2011.

[209] JAKEMAN, E.; PUSEY, P. N. A model for non-Rayleigh sea echo. *IEEE Transactions on Antennas and Propagation*, v. 24, p. 806–814, 1976.

[210] BLACKNELL, D. Comparison of parameter estimators for K-distribution. *IEE Proceedings-Radar, Sonar and Navigation*, v. 141, p. 45–52, 1994. ISSN 1350-2395.

[211] ABRAHAM, D. A.; LYONS, A. P. Reliable methods for estimating the K-distribution shape parameter. *IEEE Journal of Oceanic Engineering*, v. 35, p. 288–302, 2010.

[212] BLACKNELL, D.; TOUGH, R. J. A. Parameter estimation for the K-distribution based on [z log(z)]. *IEE Proceedings - Radar, Sonar and Navigation*, v. 148, p. 309–312, 2001. ISSN 1350-2395.

[213] CINTRA, R. J.; FRERY, A. C.; NASCIMENTO, A. D. C. Parametric and nonparametric tests for speckled imagery. *Pattern Analysis and Applications*, p. 1–21, 2011.

[214] FRERY, A. C.; CRIBARI-NETO, F.; SOUZA, M. O. Analysis of minute features in speckled imagery with maximum likelihood estimation. *EURASIP Journal on Advances in Signal Processing*, v. 2004, p. 2476–2491, 2004.

[215] MEJAIL, M. E. et al. Parametric roughness estimation in amplitude SAR images under the multiplicative model. *Revista de Teledetección*, v. 13, p. 37–49, 2000.

[216] FRERY, A. C. et al. Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images. *Applied Signal Processing*, v. 4, p. 61–76, 1997.

[217] BUSTOS, O. H.; LUCINI, M. M.; FRERY, A. C. M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, v. 2002, p. 105–114, 2002.

[218] ALLENDE, H. et al. M-estimators with asymmetric influence functions: The GA0 distribution case. *The Journal of Statistical Computation and Simulation.*, v. 76, p. 941–956, 2006.

[219] Cribari-Neto, F.; FRERY, A. C.; SILVA, M. F. Improved estimation of clutter properties in speckled imagery. *Computational Statistics & Data Analysis*, v. 40, p. 801–824, 2002.

[220] VASCONCELLOS, K. L. P.; FRERY, A. C.; SILVA, L. B. Improving estimation in speckled imagery. *Computational Statistics*, v. 20, p. 503–519, 2005.

[221] PIANTO, D. M.; CRIBARI-NETO, F. Dealing with monotone likelihood in a model for speckled data. *Computational Statistics & Data Analysis*, v. 55, p. 1394 –1409, 2011.

[222] SILVA, M.; CRIBARI-NETO, F.; FRERY, A. C. Improved likelihood inference for the roughness parameter of the GA0 distribution. *Environmetrics*, v. 19, p. 347–368, 2008.

[223] NASCIMENTO, A. D. C.; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, p. 373–385, 2010.

[224] DONOHUE, K. D. et al. Parametric and nonparametric edge detection for speckle degraded images. *Optical Engineering*, SPIE, v. 32, p. 1935–1946, 1993.

[225] BEAUCHEMIN, M.; THOMSON, K.; EDWARDS, G. On nonparametric edge detection in multilook SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 36, p. 1826–1829, 1998. ISSN 0196-2892.

[226] LI, H.-C. et al. On the empirical-statistical modeling of SAR images with generalized gamma distribution. *IEEE Journal of Selected Topics in Signal Processing*, v. 5, p. 386–397, 2011. ISSN 1932-4553.

[227] ARSENAULT, H. H.; APRIL, G. Properties of speckle integrated with a finite aperture and logarithmically transformed. *Journal of the Optical Society of America A*, OSA, v. 66, p. 1160–1163, 1976.

[228] WARD, K. D. Compound representation of high resolution sea clutter. *Electronics Letters*, v. 17, p. 561–563, 1981. ISSN 0013-5194.

[229] ANFINSEN, S. N.; DOULGERIS, A. P.; ELTOFT, T. Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, v. 47, p. 3795–3809, 2009. ISSN 0196-2892.

[230] SILVA R. S. LOPES, H. F.; MIGON, H. S. The extended generalized inverse Gaussian distribution for log-linear and stochastic volatility models. *Brazilian Journal of Probability and Statistics*, v. 20, p. 67–91, 2006.

[231] KHODABIN, M.; AHMADABADI, A. Some properties of generalized gamma distribution. *Mathematical Sciences*, v. 4, p. 9–28, 2010.

[232] VASCONCELOS, J. M.; CINTRA, R. J.; NASCIMENTO, A. D. C. Goodness-of-fit Mellin-based measures for lifetime data described by the beta-G family. *(Submitted)*, 2018.

[233] VASCONCELOS, J. M.; CINTRA, R. J.; NASCIMENTO, A. D. C. New Mellin-based estimation and GoF criteria for the beta-Weibull model. *(Submitted)*, 2018.

[234] MCQUARRIE, A. D.; TSAI, C.-L. *Regression and time series model selection*. : World Scientific, 1998.

[235] EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. 1. ed. : Springer, 1994. (Chapman & Hall/CRC Monographs on Statistics & Applied Probability).

[236] CHAN, D.; CASSETTI, J.; FRERY, A. C. Texture parameter estimatio in monopolarized SAR imagery, for the single look case, using extreme value theory. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, p. 631–634, 2016.

[237] YOUNG, D. H.; AL-ABOOD, A. M.; BAKIR, S. T. Bias corrected Pearson estimators for the shape parameter in gamma regression. *Journal of Statistical COmputation and SImulation*, v. 27, p. 233–241, 2007.

[238] FERRARI, S.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, v. 31, p. 799–815, 2004. Disponível em: <http://dx.doi.org/10.1080/0266476042000214501>.

[239] MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. : London: Chapman and Hall, 1989.

[240] ATKINSON, A. C. *Plots, transformations and regression: an introduction to graphical methods of diagnostic regression analysis*. : New York: Oxford University Press, 1985.

[241] WANG, H.; OUCHI, K. Accuracy of the K-distribution regression model for forest biomass estimation by high-resolution polarimetric SAR: Comparison of model estimation and field data. *IEEE transactions on geoscience and remote sensing*, v. 46, p. 1058–1064, 2008.

[242] HUZURBAZAR, V. S. Probability distributions and orthogonal parameters. *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, v. 46, n. 2, p. 281–284, 1950.

[243] HARDIN, J.; HILBE, J. *Generalized Linear Models and Extensions*. 3. ed. 2012.

[244] AMIN, M.; AMANULLAH, M.; CORDEIRO, G. M. Influence diagnostics in the gamma regression model with adjusted deviance residuals. *Communications in Statistics - Simulation and Computation*, p. 1–31, 2016.

[245] PAULA, G. A. *Modelos de regressão com apoio computacional*. : IME-USP, 2004.

[246] NETER, J.; KUTNER, M.; WASSERMAN, W. *Applied Linear Statistical Models*. 1996.

[247] WEI, B. C.; HU, Y. Q.; FUNG, W. K. Generalized leverage and its applications. *Scandinavian Journal of Statistics*, v. 25, p. 25–37, 1998.

[248] PREGIBON, D. Logistic regression diagnostics. *The annals of statistics*, v. 9, p. 705–724, 1981.

[249] LINDSEY, J. K. *Applying Generalized Linear Models*. : Springer, 1997. 265 p. (II).

[250] COOK, R. D. Detection of influential observations in linear regression. *Technometrics*, v. 19, p. 15–18, 1977.

[251] COOK, R. D. Assessment of local influence (with discussion). *Journal of the Royal Statistical Society B*, v. 48, p. 133–169, 1986.

[252] WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, v. 48, p. 817–838, 1980.

[253] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2015. Disponível em: <https://www.R-project.org/>.

[254] CORDEIRO, G. M. et al. General properties for the beta extended Half-normal model. *Journal of Statistical Computation and Simulation*, v. 84, p. 881–901, 2014.

[255] BARRETO-SOUZA, W.; SANTOS, A. H. S.; CORDEIRO, G. M. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, v. 80, p. 159–172, 2010.

[256] PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I. *Integrals and Series: Special functions*. : Taylor & Francis, 1986. 802 p.

[257] SEARLE, S. R. *Linear Models*. : John Wiley & Sons, Inc., 1997.

[258] RAO, C. R. *Linear Statistical Inference and Its Applications*. 2. ed. : Wiley, 1973.

[259] RENCHER, A. C.; SCHAALJE, G. B. *Linear Models in Statistics*. : John Wiley & Sons, 2007.

Appendix A - Proof of PWMs for Baseline Distributions

In this appendix, we derive PWM expressions for Fréchet and Kumaraswamy distributions.

Fréchet PWM

Applying the Fréchet quantile function $Q[F(x)] = \lambda[-\log(F(x))]^{-1/\alpha}$ to (2.4), and considering the substitution $u = -\log(F(x))$, we obtain:

$$\mathcal{M}_{l,j,k} = \lambda^l \int_0^\infty u^{-l/\alpha} \mathrm{e}^{-u(j+1)} \left(1-\mathrm{e}^{-u}\right)^k \mathrm{d}u. \tag{1}$$

Taking into account the following series expansion described in [254, 255]

$$(1-z)^j = \sum_{r=0}^{\infty} \binom{j}{r} (-1)^r z^r, \tag{2}$$

for $|z| < 1$ and $j > 0$, (1) becomes:

$$\mathcal{M}_{l,j,k} = \lambda^l \sum_{r=0}^{\infty} \binom{k}{r} (-1)^r \int_0^\infty u^{-l/\alpha} \mathrm{e}^{-u(j+r+1)} \mathrm{d}u.$$

From $\Gamma(\delta) = v^\delta \int_0^\infty x^{\delta-1} \mathrm{e}^{-vx} \mathrm{d}x$, the following result holds:

$$\mathcal{M}_{l,j,k} = \lambda^l \Gamma\left(1-\frac{l}{\alpha}\right) \sum_{r=0}^{\infty} \binom{k}{r} (-1)^r \frac{1}{(j+r+1)^{1-l/\alpha}},$$

For integer, non-negative values of k, we have [256, p. 612]:

$$\mathcal{M}_{l,j,k} = \lambda^l \Gamma\left(1-\frac{l}{\alpha}\right) \sum_{r=0}^{k} \binom{k}{r} (-1)^r \frac{1}{(j+r+1)^{1-l/\alpha}}.$$

The last results are valid for $(1-l/\alpha) \notin \mathbb{Z}_-$.

■

Kumaraswamy PWM

Applying the Kumaraswamy model quantile function $Q[F(x)] = \left[1-(1-F(x))^{1/\lambda}\right]^{1/\alpha}$ in (2.4), and considering the substitution $u = 1-(1-F(x))^{1/\lambda}$, we have:

$$\mathcal{M}_{l,j,k} = \lambda \int_0^1 u^{l/\alpha}(1-u)^{\lambda(k+1)-1}[1-(1-u)^{\lambda}]^j \,\mathrm{d}u.$$

Again, take $v = (1-u)^{\lambda}$ and use [256, p. 271]

$$\int_a^b f(x)\,\mathrm{d}(x) = -\int_b^a f(x)\,\mathrm{d}(x),$$

to obatin

$$\mathcal{M}_{l,j,k} = \int_0^1 (1-v^{1/\lambda})^{l/\alpha}(1-v)^j v^k \,\mathrm{d}v.$$

Invoking (2), we have

$$\mathcal{M}_{l,j,k} = \sum_{r=0}^{\infty}\binom{j}{r}(-1)^r \int_0^1 (1-v^{1/\lambda})^{l/\alpha} v^{r+k}\,\mathrm{d}v.$$

Given $t = 1-v^{1/\lambda}$, we obtain

$$\mathcal{M}_{l,j,k} = \lambda\sum_{r=0}^{\infty}\binom{j}{r}(-1)^r \int_0^1 t^{l/\alpha}(1-t)^{\lambda(k+r+1)-1}\,\mathrm{d}t.$$

As $B(\delta,\tau) = \int_0^1 x^{\delta-1}(1-x)^{\tau-1}\,\mathrm{d}x$, the following holds:

$$\mathcal{M}_{l,j,k} = \lambda\sum_{r=0}^{\infty}\binom{j}{r}(-1)^r B\left[1+l/\alpha,\, \lambda(k+r+1)\right].$$

For integer, non-negative values of j, we have that [256, p. 612]:

$$\mathcal{M}_{l,j,k} = \lambda\sum_{r=0}^{j}\binom{j}{r}(-1)^r B\left[1+l/\alpha,\, \lambda(k+r+1)\right].$$

■

Appendix B - Proof of GoF Criteria

In this appendix, we furnish proofs for Propositions 3, 4, 5, and 6.

General Derivation

According to (2.25), in order to derive the sought statistics for the distribution, we need to obtain the following quantities: (i) the estimates $\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$ and (ii) $\widehat{\boldsymbol{K}}^{-1}$.

Algorithm 2: Computation of $\widehat{\widetilde{\kappa}}_2$ and $\widehat{\widetilde{\kappa}}_3$

Step 1. If BW or BKw models are considered, then let $\boldsymbol{\theta} = \begin{bmatrix} \lambda & \alpha & b \end{bmatrix}^\top$; otherwise $\boldsymbol{\theta} = \begin{bmatrix} \lambda & \alpha & a \end{bmatrix}^\top$;

Step 2. Compute log-likelihood function $\ell(\boldsymbol{\theta})$;

Step 3. Derive the ML estimates $\widehat{\boldsymbol{\theta}}$ by solving the score vector at zero, which can be performed by means of iterative methods, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm;

Step 4. Derive estimates for $\widetilde{\kappa}_2$ and $\widetilde{\kappa}_3$ LCs based on $\widehat{\widetilde{\kappa}}_2 = g_2(\widehat{\boldsymbol{\theta}})$ and $\widehat{\widetilde{\kappa}}_3 = g_3(\widehat{\boldsymbol{\theta}})$.

Algorithm 3: Computation of $\widehat{\boldsymbol{K}}^{-1}$

Step 1. Compute the LC matrix $\widehat{\boldsymbol{J}}$ according to (2.23), considering the particular functions $g_2(\bullet)$ and $g_3(\bullet)$;

Step 2. Derive the matrix $\boldsymbol{\Sigma}$ according to (2.24);

Step 3. Compute: $\widehat{\boldsymbol{K}} = \widehat{\boldsymbol{J}}^\top \cdot \widehat{\boldsymbol{\Sigma}} \cdot \widehat{\boldsymbol{J}}$;

Step 4. If $\widehat{\boldsymbol{K}}$ is nonsingular, compute $\widehat{\boldsymbol{K}}^{-1}$ by usual inversion [137, 138]; otherwise compute the generalized inverse of $\widehat{\boldsymbol{K}}$ [139, 140].

With the outputs from Algorithm 2 and 3, the sought statistics can be obtained according to the algebraic manipulation implied by (2.25).

In the next subsections, for each considered model, we state the necessary inputs for the above algorithms and derive the statistics.

Beta-Weibull Distribution

Log-likelihood Function

$$\ell(\boldsymbol{\theta}) = n\log(\alpha) - n\,\alpha\log(\lambda) - n\log B(1,b) + (\alpha-1)\sum_{r=1}^{n}\log(x_r) - b\sum_{r=1}^{n}\left(\frac{x_r}{\lambda}\right)^{\alpha}.$$

Score vector components

$$\begin{aligned}
\frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha} &= \frac{n}{\alpha} + \sum_{r=1}^{n}\log\left(\frac{x_r}{\lambda}\right) - b\sum_{r=1}^{n}\left(\frac{x_r}{\lambda}\right)^{\alpha}\log\left(\frac{x_r}{\lambda}\right),\\
\frac{\partial\ell(\boldsymbol{\theta})}{\partial\lambda} &= -n\frac{\alpha}{\lambda} + b\frac{\alpha}{\lambda^{\alpha+1}}\sum_{r=1}^{n}x_r^{\alpha},\\
\frac{\partial\ell(\boldsymbol{\theta})}{\partial b} &= -n\left\{\psi(b) - \psi(1+b)\right\} - \frac{1}{\lambda^{\alpha}}\sum_{r=1}^{n}x_r^{\alpha}.
\end{aligned}$$

Functions $g_2(\boldsymbol{\theta})$ and $g_3(\boldsymbol{\theta})$

$$g_2(\boldsymbol{\theta}) = \frac{1}{\alpha^2}\psi^{(1)}(1), \quad \text{and} \quad g_3(\boldsymbol{\theta}) = \frac{1}{\alpha^3}\psi^{(2)}(1).$$

Information Matrix and Its Inverse

$$\boldsymbol{H}_{\text{BW}} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\top} \partial \boldsymbol{\theta}} = \begin{bmatrix} \mathrm{U}_{\alpha\alpha} & \mathrm{U}_{\alpha\lambda} & \mathrm{U}_{\alpha b} \\ \mathrm{U}_{\lambda\alpha} & \mathrm{U}_{\lambda\lambda} & \mathrm{U}_{\lambda b} \\ \mathrm{U}_{b\alpha} & \mathrm{U}_{b\lambda} & \mathrm{U}_{bb} \end{bmatrix},$$

where $\mathrm{U}_{\alpha\alpha} = \frac{n}{\alpha^2} + b\xi_2^{\text{BW}}$, $\mathrm{U}_{\alpha\lambda} = \mathrm{U}_{\lambda\alpha} = \frac{1}{\lambda}(n - b\xi_3^{\text{BW}})$, $\mathrm{U}_{\alpha b} = \mathrm{U}_{b\alpha} = \xi_1^{\text{BW}}$, $\mathrm{U}_{\lambda\lambda} = [b\alpha(\alpha+1)\xi_0^{\text{BW}} - n\alpha]/\lambda^2$, $\mathrm{U}_{\lambda b} = \mathrm{U}_{b\lambda} = -\frac{\alpha}{\lambda}\xi_0^{\text{BW}}$, $\mathrm{U}_{bb} = n[\psi^{(1)}(b) - \psi^{(1)}(1+b)]$, $\xi_s^{\text{BW}} = \sum_{r=1}^{n} \left(\frac{x_r}{\widehat{\lambda}}\right)^{\widehat{\alpha}} \log^s\left(\frac{x_r}{\widehat{\lambda}}\right)$, for $s = 0,1,2$; and $\xi_3^{\text{BW}} = \sum_{r=1}^{n} \left(\frac{x_r}{\lambda}\right)^{\alpha} \left[\log\left(\frac{x_r}{\lambda}\right)^{\alpha} + 1\right]$.

If the determinant $|\boldsymbol{H}_{\text{BW}}| \neq 0$, the asymptotic covariance matrix is given by $\boldsymbol{\Sigma}_{\text{BW}} \approx \boldsymbol{H}_{\text{BW}}^{-1}$, where the usual matrix inversion is applied [137,138]; otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Log-cumulant Matrix

$$\boldsymbol{J}_{\text{BW}} = -\begin{bmatrix} \frac{2}{\alpha^3}\psi^{(1)}(1) & \frac{3}{\alpha^4}\psi^{(2)}(1) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Asymptotic Covariance Matrix and Its Inverse

$$\boldsymbol{K}_{\text{BW}} = \left(\frac{\mathrm{U}_{\alpha\alpha}\mathrm{U}_{bb} - \mathrm{U}_{\lambda b}^2}{\alpha^8 |\boldsymbol{H}_{\text{BW}}|}\right) \cdot \begin{bmatrix} 4\alpha^2\psi^{(1)}(1)^2 & 6\alpha\psi^{(1)}(1)\psi^{(2)}(1) \\ 6\alpha\psi^{(1)}(1)\psi^{(2)}(1) & 9\psi^{(2)}(1)^2 \end{bmatrix}.$$

where

$$|\boldsymbol{H}_{\text{BW}}| = \mathrm{U}_{\alpha\alpha}(\mathrm{U}_{\lambda\lambda}\mathrm{U}_{bb} - \mathrm{U}_{\lambda b}^2) + \mathrm{U}_{\alpha\lambda}(\mathrm{U}_{\alpha b}\mathrm{U}_{\lambda b} - \mathrm{U}_{\alpha\lambda}\mathrm{U}_{bb}) + \mathrm{U}_{\alpha b}(\mathrm{U}_{\alpha\lambda}\mathrm{U}_{\lambda b} - \mathrm{U}_{\alpha b}\mathrm{U}_{\lambda\lambda}).$$

Because $\boldsymbol{K}_{\text{BW}}$ is singular, the generalized Moore-Penrose inverse was computed [139, p. 508]:

$$\boldsymbol{K}_{\text{BW}}^{-1} = \left(\frac{\alpha^6 |\boldsymbol{H}_{\text{BW}}|}{\text{U}_{\alpha\alpha}\text{U}_{bb} - \text{U}_{\lambda b}^2} \right) \begin{bmatrix} (2\psi^{(1)}(1))^{-2} & 0 \\ 0 & 0 \end{bmatrix}.$$

Hotelling's T^2 statistic Derivation

Therefore, we obtain:

$$T_{\text{BW}}^2 = \frac{n\widehat{\alpha}^6}{4} \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right)^2 \left(\frac{|\widehat{\boldsymbol{H}}_{\text{BW}}|}{\text{U}_{\widehat{\alpha}\widehat{\alpha}}\text{U}_{\widehat{b}\widehat{b}} - \text{U}_{\widehat{\lambda}\widehat{b}}^2} \right).$$

■

Beta-Fréchet Distribution:

Log-likelihood Function

$$\ell(\boldsymbol{\theta}) = n\log(\alpha) + n\alpha\log(\lambda) - n\log B(a,1) - (1+\alpha)\sum_{r=1}^{n}\log(x_r) - a\sum_{r=1}^{n}\left(\frac{\lambda}{x_r}\right)^{\alpha}.$$

Score vector components

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{n}{\alpha} + n\log(\lambda) - \sum_{r=1}^{n}\log(x_r) - a\sum_{r=1}^{n}\left(\frac{\lambda}{x_r}\right)^{\alpha}\log\left(\frac{\lambda}{x_r}\right), \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} &= n\frac{\alpha}{\lambda} - a\alpha\lambda^{\alpha-1}\sum_{r=1}^{n}x_r^{-\alpha}, \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial a} &= -n\{\psi(a) - \psi(1+a)\} - \lambda^{\alpha}\sum_{r=1}^{n}x_r^{-\alpha}.
\end{aligned}$$

Functions $g_2(\boldsymbol{\theta})$ and $g_3(\boldsymbol{\theta})$

$$g_2(\boldsymbol{\theta}) = \frac{1}{\alpha^2}\psi^{(1)}(1), \quad \text{and} \quad g_3(\boldsymbol{\theta}) = -\frac{1}{\alpha^3}\psi^{(2)}(1).$$

Information Matrix and Its Inverse

$$\boldsymbol{H}_{\text{BF}} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top \partial \boldsymbol{\theta}} = \begin{bmatrix} \mathrm{U}_{\alpha\alpha} & \mathrm{U}_{\alpha\lambda} & \mathrm{U}_{\alpha a} \\ \mathrm{U}_{\lambda\alpha} & \mathrm{U}_{\lambda\lambda} & \mathrm{U}_{\lambda a} \\ \mathrm{U}_{a\alpha} & \mathrm{U}_{a\lambda} & \mathrm{U}_{aa} \end{bmatrix},$$

where $\mathrm{U}_{\alpha\alpha} = \frac{n}{\alpha^2} + a\xi_2^{\text{BF}}$, $\mathrm{U}_{\alpha\lambda} = \mathrm{U}_{\lambda\alpha} = \frac{1}{\lambda}(a\xi_3^{\text{BF}} - n)$, $\mathrm{U}_{\alpha a} = \mathrm{U}_{a\alpha} = \xi_1^{\text{BF}}$, $\mathrm{U}_{\lambda\lambda} = \frac{1}{\lambda^2}\left[n\alpha - a\alpha(1-\alpha)\xi_0^{\text{BF}}\right]$, $\mathrm{U}_{\lambda a} = \mathrm{U}_{a\lambda} = \frac{\alpha}{\lambda}\xi_0^{\text{BF}}$, $\mathrm{U}_{aa} = n[\psi^{(1)}(a) - \psi^{(1)}(1+a)]$, $\xi_s^{\text{BF}} = \sum_{r=1}^{n}\left(\frac{\widehat{\lambda}}{x_r}\right)^{\widehat{\alpha}} \log^s\left(\frac{\widehat{\lambda}}{x_r}\right)$ with $s = 0, 1, 2$, and $\xi_3^{\text{BF}} = \sum_{r=1}^{n}\left(\frac{\lambda}{x_r}\right)^{\alpha}\left[\log(\frac{\lambda}{x_r})^{\alpha} + 1\right]$.

If the determinant $|\boldsymbol{H}_{\text{BF}}| \neq 0$, the asymptotic covariance matrix is given by $\boldsymbol{\Sigma}_{\text{BF}} \approx \boldsymbol{H}_{\text{BF}}^{-1}$ [137, 138]; otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Log-cumulant Matrix

$$\boldsymbol{J}_{\text{BF}} = \begin{bmatrix} -\frac{2}{\alpha^3}\psi^{(1)}(1) & \frac{3}{\alpha^4}\psi^{(2)}(1) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Asymptotic Covariance Matrix and Its Inverse

$$\boldsymbol{K}_{\text{BF}} = \left(\frac{\mathrm{U}_{\alpha\alpha}\mathrm{U}_{aa} - \mathrm{U}_{\lambda a}^2}{\alpha^8 |\boldsymbol{H}_{\text{BF}}|} \right) \cdot \begin{bmatrix} 4\widehat{\alpha}^2 \psi^{(1)}(1)^2 & -6\widehat{\alpha}\psi^{(1)}(1)\psi^{(2)}(1) \\ -6\widehat{\alpha}\psi^{(1)}(1)\psi^{(2)}(1) & 9\psi^{(2)}(1)^2 \end{bmatrix}.$$

Because $\boldsymbol{K}_{\text{BF}}$ is singular, the generalized Moore-Penrose inverse was computed [139, p. 508]:

$$\boldsymbol{K}_{\text{BF}}^{-1} = \left(\frac{\alpha^6 |\boldsymbol{H}_{\text{BF}}|}{\mathrm{U}_{\alpha\alpha}\mathrm{U}_{aa} - \mathrm{U}_{\lambda a}^2} \right) \begin{bmatrix} \left(2\psi^{(1)}(1)\right)^{-2} & 0 \\ 0 & 0 \end{bmatrix}.$$

Hotelling's T^2 statistic Derivation

Therefore, we obtain:

$$T_{\text{BF}}^2 = \frac{n\widehat{\alpha}^6}{4} \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right)^2 \left(\frac{|\widehat{\boldsymbol{H}}_{\text{BF}}|}{\mathrm{U}_{\widehat{\alpha}\widehat{\alpha}}\mathrm{U}_{\widehat{a}\widehat{a}} - \mathrm{U}_{\widehat{\lambda}\widehat{a}}^2} \right).$$

■

Beta-Kumaraswamy Distribution

Log-likelihood Function

$$\ell(\boldsymbol{\theta}) = n\log(\alpha\lambda) - n\log B(1,b) + (\alpha-1)\sum_{r=1}^{n}\log(x_r) + (\lambda b - 1)\sum_{r=1}^{n}\log\left(1 - x_r^{\alpha}\right),$$

Score vector components

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^{n} \log(x_r) - (\lambda b - 1) \sum_{r=1}^{n} \frac{x_r^{\alpha} \log(x_r)}{1 - x_r^{\alpha}}, \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} &= \frac{n}{\lambda} + b \sum_{r=1}^{n} \log(1 - x_r^{\alpha}), \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial b} &= -n\{\psi(b) - \psi(1+b)\} + \lambda \sum_{r=1}^{n} \log(1 - x_r^{\alpha}).
\end{aligned}$$

Functions $g_2(\boldsymbol{\theta})$ and $g_3(\boldsymbol{\theta})$

$$g_2(\boldsymbol{\theta}) = \frac{\psi^{(1)}(1) - \psi^{(1)}(\lambda b + 1)}{\alpha^2}, \quad \text{and} \quad g_3(\boldsymbol{\theta}) = \frac{\psi^{(2)}(1) - \psi^{(2)}(\lambda b + 1)}{\alpha^3}.$$

Information Matrix and Its inverse

$$\boldsymbol{H}_{\text{BKw}} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\top} \partial \boldsymbol{\theta}} = \begin{bmatrix} \mathrm{U}_{\alpha\alpha} & \mathrm{U}_{\alpha\lambda} & \mathrm{U}_{\alpha b} \\ \mathrm{U}_{\lambda\alpha} & \mathrm{U}_{\lambda\lambda} & \mathrm{U}_{\lambda b} \\ \mathrm{U}_{b\alpha} & \mathrm{U}_{b\lambda} & \mathrm{U}_{bb} \end{bmatrix},$$

where $\mathrm{U}_{\alpha\alpha} = \frac{n}{\alpha^2} + (\lambda b - 1) \sum_{r=1}^{n} \frac{x_r^{\alpha} \log^2(x_r)}{(1 - x_r^{\alpha})^2}$, $\mathrm{U}_{\alpha\lambda} = \mathrm{U}_{\lambda\alpha} = b \sum_{r=1}^{n} \frac{x_r^{\alpha} \log(x_r)}{(1 - x_r^{\alpha})}$, $\mathrm{U}_{\alpha b} = \mathrm{U}_{b\alpha} = \lambda \sum_{r=1}^{n} \frac{x_r^{\alpha} \log(x_r)}{(1 - x_r^{\alpha})}$, $\mathrm{U}_{\lambda\lambda} = \frac{n}{\lambda^2}$, $\mathrm{U}_{\lambda b} = \mathrm{U}_{b\lambda} = -\sum_{r=1}^{n} \log(1 - x_r^{\alpha})$, and $\mathrm{U}_{bb} = n[\psi^{(1)}(b) - \psi^{(1)}(1+b)]$.

If the determinant $|\boldsymbol{H}_{\text{BKw}}| \neq 0$, the asymptotic covariance matrix is given by

$$\boldsymbol{\Sigma}_{\text{BKw}} \approx \frac{1}{|\boldsymbol{H}_{\text{BKw}}|} \begin{bmatrix} \mathrm{U}^{b}_{\alpha\alpha} & \mathrm{U}^{b}_{\alpha\lambda} & \mathrm{U}^{b}_{\alpha b} \\ \mathrm{U}^{b}_{\lambda\alpha} & \mathrm{U}^{b}_{\lambda\lambda} & \mathrm{U}^{b}_{\lambda b} \\ \mathrm{U}^{b}_{b\alpha} & \mathrm{U}^{b}_{b\lambda} & \mathrm{U}^{b}_{bb} \end{bmatrix},$$

with $\mathrm{U}^{b}_{\alpha\alpha} = \mathrm{U}_{\lambda\lambda}\mathrm{U}_{bb} - \mathrm{U}^{2}_{\lambda b}$, $\mathrm{U}^{b}_{\alpha\lambda} = \mathrm{U}^{b}_{\lambda\alpha} = \mathrm{U}_{\alpha b}\mathrm{U}_{\lambda b} - \mathrm{U}_{\alpha b}\mathrm{U}_{\lambda\lambda}$, $\mathrm{U}^{b}_{\alpha b} = \mathrm{U}^{b}_{b\alpha} = \mathrm{U}_{\alpha\lambda}\mathrm{U}_{\lambda b} - \mathrm{U}_{\alpha b}\mathrm{U}_{\lambda\lambda}$, $\mathrm{U}^{b}_{\lambda\lambda} = \mathrm{U}_{\alpha\alpha}\mathrm{U}_{bb} - \mathrm{U}^{2}_{\alpha b}$, $\mathrm{U}^{b}_{\lambda b} = \mathrm{U}^{b}_{b\lambda} = \mathrm{U}_{\alpha\lambda}\mathrm{U}_{\alpha b} - \mathrm{U}_{\alpha\alpha}\mathrm{U}_{\lambda b}$, $\mathrm{U}^{b}_{bb} = \mathrm{U}_{\alpha\alpha}\mathrm{U}_{\lambda\lambda} - \mathrm{U}^{2}_{\alpha\lambda}$; otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Log-cumulant Matrix

$$\boldsymbol{J}_{\mathrm{BKw}} = \begin{bmatrix} J_{12} & J_{13} \\ J_{22} & J_{23} \\ J_{32} & J_{33} \end{bmatrix},$$

where

$$\begin{aligned}
J_{12} &= \frac{2}{\alpha^3}\{\psi^{(1)}(\lambda b+1) - \psi^{(1)}(1)\}, \\
J_{13} &= \frac{3}{\alpha^4}\{\psi^{(2)}(\lambda b+1) - \psi^{(2)}(1)\}, \\
J_{22} &= \frac{b}{\alpha^2}\psi^{(2)}(\lambda b+1), \quad J_{23} = \frac{b}{\alpha^3}\psi^{(3)}(\lambda b+1), \\
J_{32} &= \frac{\lambda}{\alpha^2}\psi^{(2)}(\lambda b+1), \quad J_{33} = \frac{\lambda}{\alpha^3}\psi^{(3)}(\lambda b+1).
\end{aligned}$$

Asymptotic Covariance Matrix and Its Inverse

$$\boldsymbol{K}_{\mathrm{BKw}} = \frac{1}{|\boldsymbol{H}_{\mathrm{BKw}}|}\begin{bmatrix} \delta_{22} & \delta_{32} \\ \delta_{23} & \delta_{33} \end{bmatrix},$$

where

$$\begin{aligned}\delta_{22} &= J_{12}(J_{12}\mathrm{U}^b_{\alpha\alpha} + J_{22}\mathrm{U}^b_{\lambda\alpha} + J_{32}\mathrm{U}^b_{b\alpha}) + J_{22}(J_{12}\mathrm{U}^b_{\alpha\lambda} + J_{22}\mathrm{U}^b_{\lambda\lambda} + J_{32}\mathrm{U}^b_{b\lambda}) \\ &\quad + J_{32}(J_{12}\mathrm{U}^b_{\alpha b} + J_{22}\mathrm{U}^b_{\lambda b} + J_{32}\mathrm{U}^b_{bb}),\end{aligned}$$

$$\begin{aligned}\delta_{23} = \delta_{32} &= J_{13}(J_{12}\mathrm{U}^b_{\alpha\alpha} + J_{22}\mathrm{U}^b_{\lambda\alpha} + J_{32}\mathrm{U}^b_{b\alpha}) + J_{23}(J_{12}\mathrm{U}^b_{\alpha\lambda} + J_{22}\mathrm{U}^b_{\lambda\lambda} + J_{32}\mathrm{U}^b_{b\lambda}) \\ &\quad + J_{33}(J_{12}\mathrm{U}^b_{\alpha b} + J_{22}\mathrm{U}^b_{\lambda b} + J_{32}\mathrm{U}^b_{bb}),\end{aligned}$$

$$\begin{aligned}\delta_{33} &= J_{13}(J_{13}\mathrm{U}^b_{\alpha\alpha} + J_{23}\mathrm{U}^b_{\lambda\alpha} + J_{33}\mathrm{U}^b_{b\alpha}) + J_{23}(J_{13}\mathrm{U}^b_{\alpha\lambda} + J_{23}\mathrm{U}^b_{\lambda\lambda} + J_{33}\mathrm{U}^b_{b\lambda}) \\ &\quad + J_{33}(J_{13}\mathrm{U}^b_{\alpha b} + J_{23}\mathrm{U}^b_{\lambda b} + J_{33}\mathrm{U}^b_{bb}).\end{aligned}$$

If $\delta_{22}\delta_{33} > \delta_{23}^2$ and $|\boldsymbol{H}_{\mathrm{BKw}}| \neq 0$, then the inverse is given by

$$\boldsymbol{K}^{-1}_{\mathrm{BKw}} = \frac{|\boldsymbol{H}_{\mathrm{BKw}}|}{\delta_{22}\delta_{33} - \delta_{23}^2}\begin{bmatrix} \delta_{33} & -\delta_{32} \\ -\delta_{23} & \delta_{22} \end{bmatrix},$$

otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Hotelling's T^2 statistic Derivation

The sought T^2_{BKw} statistic is:

$$T^2_{\mathrm{BKw}} = \frac{n|\widehat{\boldsymbol{H}}_{\mathrm{BKw}}|}{\widehat{\delta}_{22}\widehat{\delta}_{33} - \widehat{\delta}_{23}^2}\left[\widehat{\delta}_{33}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)^2 + \widehat{\delta}_{22}\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)^2 - 2\widehat{\delta}_{23}\left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2\right)\left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3\right)\right].$$

■

Beta-log-logistic Distribution:

Log-likelihood Function

$$\ell(\boldsymbol{\theta}) = n\log\left(\frac{\lambda}{\alpha}\right) - n\log B(a,1) + (a\lambda - 1)\sum_{r=1}^{n}\log\left(\frac{x_r}{\alpha}\right) - (a+1)\sum_{r=1}^{n}\log\left[1 + (x_r/\alpha)^{\lambda}\right].$$

Score vector components

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = -\frac{n\lambda a}{\alpha} + \frac{\lambda(a+1)}{\alpha}\sum_{r=1}^{n}\frac{(x_r/\alpha)^{\lambda}}{\left[1+(x_r/\alpha)^{\lambda}\right]},$$
$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} = \frac{n}{\lambda} + a\sum_{r=1}^{n}\log\left(\frac{x_r}{\alpha}\right) - (a+1)\sum_{r=1}^{n}\frac{(x_r/\alpha)^{\lambda}\log(x_r/\alpha)}{\left[1+(x_r/\alpha)^{\lambda}\right]},$$
$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial a} = -n\{\psi(a) - \psi(1+a)\} + \lambda\sum_{r=1}^{n}\log\left(\frac{x_r}{\alpha}\right) - \sum_{r=1}^{n}\log\left[1+(x_r/\alpha)^{\lambda}\right].$$

Functions $g_2(\boldsymbol{\theta})$ and $g_3(\boldsymbol{\theta})$

$$g_2(\boldsymbol{\theta}) = \frac{\psi^{(1)}(1)+\psi^{(1)}(1)}{\lambda^2}, \quad \text{and} \quad g_3(\boldsymbol{\theta}) = \frac{\psi^{(2)}(1)-\psi^{(2)}(1)}{\lambda^3}.$$

Information Matrix and Its inverse

$$\boldsymbol{H}_{\text{BLL}} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\top}\partial \boldsymbol{\theta}} = \begin{bmatrix} \mathrm{U}_{\alpha\alpha} & \mathrm{U}_{\alpha\lambda} & \mathrm{U}_{\alpha a} \\ \mathrm{U}_{\lambda\alpha} & \mathrm{U}_{\lambda\lambda} & \mathrm{U}_{\lambda a} \\ \mathrm{U}_{a\alpha} & \mathrm{U}_{a\lambda} & \mathrm{U}_{aa} \end{bmatrix},$$

where

$$
\begin{aligned}
\mathrm{U}_{\alpha\alpha} &= -\frac{\lambda}{\alpha^2}\left\{na-(a+1)\sum_{r=1}^{n} z_r\left[1+\frac{\lambda}{1+y_r^{\lambda}}\right]\right\},\\
\mathrm{U}_{\alpha\lambda} &= \frac{1}{\alpha}\left[na-(a+1)\sum_{r=1}^{n}\left(z_r+\lambda y_r^{\lambda}\log(y_r)\right)\right],\\
\mathrm{U}_{\alpha a} &= -\frac{\lambda}{\alpha}\sum_{r=1}^{n}(z_r-1),\\
\mathrm{U}_{\lambda\lambda} &= \frac{n}{\lambda^2}+(a+1)\sum_{r=1}^{n}\frac{z_r\log y_r}{1+y_r^{\lambda}}\left[(1+y_r^{\lambda})\log y_r+\frac{\lambda}{\alpha}y_r^{\lambda}\right],\\
\mathrm{U}_{\lambda a} &= -\sum_{r=1}^{n}(1-z_r)\log y_r, \quad \mathrm{U}_{aa} = n[\psi^{(1)}(a)-\psi^{(1)}(1+a)].
\end{aligned}
$$

with $z_r = y_r^{\lambda}/(1+y_r^{\lambda})$, and $y_r = x_r/\alpha$.

If the determinant $|\boldsymbol{H}_{\mathrm{BLL}}| \neq 0$, the asymptotic covariance matrix is given by

$$
\boldsymbol{\Sigma}_{\mathrm{BLL}} \approx \frac{1}{|\boldsymbol{H}_{\mathrm{BLL}}|}\begin{bmatrix}
\mathrm{U}_{\alpha\alpha}^{a} & \mathrm{U}_{\alpha\lambda}^{a} & \mathrm{U}_{\alpha a}^{a}\\
\mathrm{U}_{\lambda\alpha}^{a} & \mathrm{U}_{\lambda\lambda}^{a} & \mathrm{U}_{\lambda a}^{a}\\
\mathrm{U}_{a\alpha}^{a} & \mathrm{U}_{a\lambda}^{a} & \mathrm{U}_{aa}^{a}
\end{bmatrix},
$$

with $\mathrm{U}_{\alpha\alpha}^{a} = \mathrm{U}_{\lambda\lambda}\mathrm{U}_{aa} - \mathrm{U}_{\lambda a}^{2}$, $\mathrm{U}_{\alpha\lambda}^{a} = \mathrm{U}_{\lambda\alpha}^{a} = \mathrm{U}_{\alpha a}\mathrm{U}_{\lambda a} - \mathrm{U}_{\alpha a}\mathrm{U}_{\lambda\lambda}$, $\mathrm{U}_{\alpha a}^{a} = \mathrm{U}_{a\alpha}^{a} = \mathrm{U}_{\alpha\lambda}\mathrm{U}_{\lambda a} - \mathrm{U}_{\alpha a}\mathrm{U}_{\lambda\lambda}$, $\mathrm{U}_{\lambda\lambda}^{a} = \mathrm{U}_{\alpha\alpha}\mathrm{U}_{aa} - \mathrm{U}_{\alpha a}^{2}$, $\mathrm{U}_{\lambda a}^{a} = \mathrm{U}_{a\lambda}^{a} = \mathrm{U}_{\alpha\lambda}\mathrm{U}_{\alpha a} - \mathrm{U}_{\alpha\alpha}\mathrm{U}_{\lambda a}$, $\mathrm{U}_{aa}^{a} = \mathrm{U}_{\alpha\alpha}\mathrm{U}_{\lambda\lambda} - \mathrm{U}_{\alpha\lambda}^{2}$; otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Log-cumulant Matrix

$$
\boldsymbol{J}_{\mathrm{BLL}} = \begin{bmatrix}
J_{12} & J_{13}\\
J_{22} & J_{23}\\
J_{32} & J_{33}
\end{bmatrix},
$$

where

$$J_{12} = J_{13} = 0,$$
$$J_{22} = -\frac{2}{\lambda^3}\{\psi^{(1)}(a) + \psi^{(1)}(1)\},$$

$$J_{23} = -\frac{3}{\lambda^4}\{\psi^{(2)}(a) - \psi^{(2)}(1)\},$$

$$J_{32} = \frac{1}{\lambda^2}\psi^{(2)}(a), \quad J_{33} = \frac{1}{\lambda^3}\psi^{(3)}(a).$$

Asymptotic Covariance Matrix and Its Inverse

$$\boldsymbol{K}_{\text{BLL}} = \frac{1}{|\boldsymbol{H}_{\text{BLL}}|}\begin{bmatrix} \delta_{22} & \delta_{32} \\ \delta_{23} & \delta_{33} \end{bmatrix},$$

where

$$\delta_{22} = J_{22}(J_{22}\mathrm{U}^a_{\lambda\lambda} + J_{32}\mathrm{U}^a_{a\lambda}) + J_{32}(J_{22}\mathrm{U}^a_{\lambda a} + J_{32}\mathrm{U}^a_{aa}),$$

$$\delta_{23} = \delta_{32} = J_{23}(J_{22}\mathrm{U}^a_{\lambda\lambda} + J_{32}\mathrm{U}^a_{a\lambda}) + J_{33}(J_{22}\mathrm{U}^a_{\lambda a} + J_{32}\mathrm{U}^a_{aa}),$$

$$\delta_{33} = J_{23}(J_{23}\mathrm{U}^a_{\lambda\lambda} + J_{33}\mathrm{U}^a_{a\lambda}) + J_{33}(J_{23}\mathrm{U}^a_{\lambda a} + J_{33}\mathrm{U}^a_{aa}).$$

If $\delta_{22}\delta_{33} > \delta_{23}^2$ and $|\boldsymbol{H}_{\text{BLL}}| \neq 0$, then the inverse is given by

$$\boldsymbol{K}^{-1}_{\text{BLL}} = \frac{|\boldsymbol{H}_{\text{BLL}}|}{\delta_{22}\delta_{33} - \delta_{23}^2}\begin{bmatrix} \delta_{33} & -\delta_{32} \\ -\delta_{23} & \delta_{22} \end{bmatrix},$$

otherwise we apply the generalized Moore-Penrose inverse [139, 140].

Hotelling's T^2 statistic Derivation

The sought T^2_{BLL} statistic is:

$$T^2_{\text{BLL}} = \frac{n|\widehat{\boldsymbol{H}}_{\text{BLL}}|}{\widehat{\delta}_{22}\widehat{\delta}_{33} - \widehat{\delta}^2_{23}} \left[\widehat{\delta}_{33} \left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2 \right)^2 + \widehat{\delta}_{22} \left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3 \right)^2 - 2\widehat{\delta}_{23} \left(\widehat{\widetilde{\kappa}}_2 - \widetilde{\kappa}_2 \right) \left(\widehat{\widetilde{\kappa}}_3 - \widetilde{\kappa}_3 \right) \right].$$

■

Appendix C - Proof of Hotelling Statistics Derivation

In this appendix we provide explicit expressions for the elements of asymptotic matrix of LCEs, which are employed to obtain the new GoF measure.

Consider the relations between moment-to-cumulant transformations given in (2.19), and $\boldsymbol{M}_3$ and $\boldsymbol{K}_3$ defined in (3.4) and (3.5), then, we have

$$\boldsymbol{K}_3 = \boldsymbol{J}_3^{\top} \cdot \boldsymbol{M}_3 \cdot \boldsymbol{J}_3,$$

with the Jacobian matrix of the moment-to-cumulant transformations given in [26, 60]

$$\boldsymbol{J}_3 = \begin{bmatrix} 1 & 0 & 0 \\ -2m_1 & 1 & 0 \\ -3(m_2 - 2m_1^2) & -3m_1 & 1 \end{bmatrix},$$

and

$$\boldsymbol{M}_3 = \begin{bmatrix} \widetilde{\kappa}_2 & \widetilde{\kappa}_3 + 2\widetilde{\kappa}_1\widetilde{\kappa}_2 & M_{13} \\ \widetilde{\kappa}_3 + 2\widetilde{\kappa}_1\widetilde{\kappa}_2 & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix},$$

where

$$\begin{aligned}
M_{13} &= \widetilde{\kappa}_4 + 3\widetilde{\kappa}_1\widetilde{\kappa}_3 + 3\widetilde{\kappa}_2^2 + 3\widetilde{\kappa}_1^2\widetilde{\kappa}_2, \\
M_{22} &= \widetilde{\kappa}_4 + 4\widetilde{\kappa}_1\widetilde{\kappa}_3 + 2\widetilde{\kappa}_2^2 + 4\widetilde{\kappa}_1^2\widetilde{\kappa}_2, \\
M_{23} &= \widetilde{\kappa}_5 + 5\widetilde{\kappa}_1\widetilde{\kappa}_4 + 9\widetilde{\kappa}_2\widetilde{\kappa}_3 + 9\widetilde{\kappa}_1^2\widetilde{\kappa}_3 + 12\widetilde{\kappa}_1\widetilde{\kappa}_2^2 + 6\widetilde{\kappa}_1^3\widetilde{\kappa}_2, \\
M_{33} &= \widetilde{\kappa}_6 + 6\widetilde{\kappa}_1\widetilde{\kappa}_5 + 15\widetilde{\kappa}_2\widetilde{\kappa}_4 + 15\widetilde{\kappa}_1^2\widetilde{\kappa}_4 + 9\widetilde{\kappa}_3^2 \\
&\quad + 54\widetilde{\kappa}_1\widetilde{\kappa}_2\widetilde{\kappa}_3 + 18\widetilde{\kappa}_1^3\widetilde{\kappa}_3 + 15\widetilde{\kappa}_2^3 + 36\widetilde{\kappa}_1^2\widetilde{\kappa}_2^2 + 9\widetilde{\kappa}_1^4\widetilde{\kappa}_2.
\end{aligned}$$

Therefore [26, 60],

$$\boldsymbol{K}_3 = \begin{bmatrix} \widetilde{\kappa}_2 & \widetilde{\kappa}_3 & \widetilde{\kappa}_4 \\ \widetilde{\kappa}_3 & \widetilde{\kappa}_4 + 2\widetilde{\kappa}_2^2 & \widetilde{\kappa}_5 + 6\widetilde{\kappa}_2\widetilde{\kappa}_3 \\ \widetilde{\kappa}_4 & \widetilde{\kappa}_5 + 6\widetilde{\kappa}_2\widetilde{\kappa}_3 & \widetilde{\kappa}_6 + 9\widetilde{\kappa}_2\widetilde{\kappa}_4 + 9\widetilde{\kappa}_3^2 + 6\widetilde{\kappa}_2^3 \end{bmatrix}. \tag{3}$$

For use the proposed of the GoF criteria we developed as input the ACM given by two types: (i) with based in matrix (3), we derived the ACM for the $\mathrm{BW}(1, b, \alpha, \lambda)$; (ii) the matrix M_3 must be replaced by observed information matrix of the $\mathrm{BW}(1, b, \alpha, \lambda)$. This way, the ACM for the first case is:

$$\boldsymbol{K}_{\mathrm{LC}} = \begin{bmatrix} \widetilde{\kappa}_4 + 2\widetilde{\kappa}_2^2 & \widetilde{\kappa}_5 + 6\widetilde{\kappa}_2\widetilde{\kappa}_3 \\ \widetilde{\kappa}_5 + 6\widetilde{\kappa}_2\widetilde{\kappa}_3 & \widetilde{\kappa}_6 + 9\widetilde{\kappa}_2\widetilde{\kappa}_4 + 9\widetilde{\kappa}_3^2 + 6\widetilde{\kappa}_2^3 \end{bmatrix} = \frac{1}{\alpha^6} \begin{bmatrix} \widetilde{\kappa}_{22} & \widetilde{\kappa}_{23} \\ \widetilde{\kappa}_{32} & \widetilde{\kappa}_{33} \end{bmatrix},$$

where

$$\begin{aligned} \widetilde{\kappa}_{22} &= \alpha^2[\Psi(3,1) + 2\Psi^2(1,1)], \\ \widetilde{\kappa}_{23} &= \alpha[\Psi(4,1) + 6\Psi(1,1)\Psi(2,1)], \\ \widetilde{\kappa}_{33} &= \Psi(5,1) + 9\Psi(1,1)\Psi(3,1) + 9\Psi^2(2,1) + 6\Psi^3(1,1). \end{aligned}$$

If $|\boldsymbol{K}_{\mathrm{LC}}| \neq 0$, then the matrix inverse is given by

$$\boldsymbol{K}_{\mathrm{LC}}^{-1} = \frac{\alpha^6}{\widetilde{\kappa}_{33}\widetilde{\kappa}_{22} - \widetilde{\kappa}_{23}^2} \begin{bmatrix} \widetilde{\kappa}_{33} & -\widetilde{\kappa}_{32} \\ -\widetilde{\kappa}_{23} & \widetilde{\kappa}_{22} \end{bmatrix},$$

otherwise we apply the generalized Moore-Penrose inverse [139, 140].

On the other hand, the ACM for the BW model using the observed information matrix instead of matrix $\boldsymbol{M}$, is given by

$$\boldsymbol{K}_{\mathrm{ML}} = \begin{bmatrix} \frac{2}{\alpha^3}\Psi(1,1) & 0 & 0 \\ \frac{3}{\alpha^4}\Psi(2,1) & 0 & 0 \end{bmatrix} \boldsymbol{\Sigma}_{\mathrm{ML}} \begin{bmatrix} \frac{2}{\alpha^3}\Psi(1,1) & \frac{3}{\alpha^4}\Psi(2,1) \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

where $\boldsymbol{\Sigma}_{\text{ML}}$ is given in (3.2). Thence,

$$\boldsymbol{K}_{\text{ML}} = \left(\frac{\mathrm{U}_{\alpha\alpha}\mathrm{U}_{bb} - \mathrm{U}_{\lambda b}^2}{\alpha^8 |\boldsymbol{H}|} \right) \cdot \begin{bmatrix} 4\alpha^2\Psi^2(1,1) & 6\alpha\Psi(1,1)\Psi(2,1) \\ 6\alpha\Psi(1,1)\Psi(2,1) & 9\Psi^2(2,1) \end{bmatrix}.$$

Due to the fact of $\boldsymbol{K}_{\text{ML}}$ be singular, the generalized Moore-Penrose inverse was computed [139, p. 508]:

$$\boldsymbol{K}_{\text{ML}}^{-1} = \left(\frac{\alpha^6 |\boldsymbol{H}|}{\mathrm{U}_{\alpha\alpha}\mathrm{U}_{bb} - \mathrm{U}_{\lambda b}^2} \right) \begin{bmatrix} (4\Psi^2(1,1))^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, the T_{LC}^2 statistics is given by:

$$T_{\text{LC}}^2 = \frac{n\alpha^6}{\widetilde{\kappa}_{33}\widetilde{\kappa}_{22} - \widetilde{\kappa}_{23}^2} \left\{ \widehat{\widetilde{\kappa}}^{22} \left[\Psi(1,1) \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right) \right]^2 + \widehat{\widetilde{\kappa}}^{33} \left[\Psi(2,1) \left(\frac{1}{\widehat{\alpha}^3} - \frac{1}{\alpha^3} \right) \right]^2 \right.$$
$$\left. - 2\widehat{\widetilde{\kappa}}^{23} \Psi(1,1)\Psi(2,1) \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right) \left(\frac{1}{\widehat{\alpha}^3} - \frac{1}{\alpha^3} \right) \right\},$$

where $\widehat{\widetilde{\kappa}}^{22}$, $\widehat{\widetilde{\kappa}}^{33}$, $\widehat{\widetilde{\kappa}}^{23}$, and $\widehat{\alpha}$, are the estimates of $\widetilde{\kappa}^{22}$, $\widetilde{\kappa}^{33}$, $\widetilde{\kappa}^{23}$; and α, respectively.

For the T_{ML}^2 statistics we have:

$$T_{\text{ML}}^2 = \frac{n\widehat{\alpha}^6}{4} \left(\frac{1}{\widehat{\alpha}^2} - \frac{1}{\alpha^2} \right)^2 \left(\frac{|\widehat{\boldsymbol{H}}|}{\mathrm{U}_{\widehat{\alpha}\widehat{\alpha}}\mathrm{U}_{\widehat{bb}} - \mathrm{U}_{\widehat{\lambda b}}^2} \right).$$

■

Appendix D - Proof of GoF measure CTPC model

We want to derive the distance

$$\begin{aligned} D_{\text{CTPC}}(\boldsymbol{\theta};\boldsymbol{x}) &= \int_{\mathbb{R}} |\Phi_{X_n}(t)-\Phi_S(t)|^2\, \mathrm{e}^{-\kappa|t|}\, \mathrm{d}t \\ &= \int_{\mathbb{R}} \nu_1(\boldsymbol{t})\, \mathrm{d}t + \int_{\mathbb{R}} \nu_2(\boldsymbol{t})\, \mathrm{d}t + \int_{\mathbb{R}} \nu_3(\boldsymbol{t})\, \mathrm{d}t, \end{aligned}$$

where $\Phi_S(t)$ and $\Phi_{X_n}(t)$ are given in (4.8) and (4.10), respectively. Also, $|\boldsymbol{x}|^2 = \boldsymbol{x}\cdot\overline{\boldsymbol{x}}$ for $\boldsymbol{x}\in\mathbb{C}^p$. Using the Euler′s identity, one has that

$$\int_{\mathbb{R}} \nu_1(\boldsymbol{t})\, \mathrm{d}t = \int_{\mathbb{R}} \nu_{11}(\boldsymbol{t})\, \mathrm{d}t + int_{\mathbb{R}}\, \nu_{12}(\boldsymbol{t})\, \mathrm{d}t,$$

from which

$$\int_{\mathbb{R}} \nu_{11}(\boldsymbol{t})\, \mathrm{d}t = \frac{1}{n^2}\sum_{j,l=1}^{n}\int_{\mathbb{R}} \cos(t(x_j-x_l))\,\mathrm{e}^{-\kappa|t|}\, \mathrm{d}t$$

and

$$\int_{\mathbb{R}} \nu_{12}(\boldsymbol{t})\, \mathrm{d}t = \frac{1}{n^2}\sum_{j,l=1}^{n}\int_{\mathbb{R}} \mathrm{i}\sin(t(x_j-x_l))\,\mathrm{e}^{-\kappa|t|}\, \mathrm{d}t. \tag{4}$$

Note that $\nu_{12}(\boldsymbol{t})$ is an odd function nd that the integral in (4) is convergent. Therefore, the result (4) is equal to zero. Moreover, using the software `wxMaxima`, one obtains the following identity:

$$\int_{\mathbb{R}} \nu_1(\boldsymbol{t})\, \mathrm{d}t = \int_{\mathbb{R}} \nu_{11}(\boldsymbol{t})\, \mathrm{d}t = \frac{2}{n^2}\sum_{j,l=1}^{n}\frac{\kappa}{\kappa^2+(x_j-x_l)^2}.$$

For the second term, we have:

$$\int_{\mathbb{R}} \nu_2(\boldsymbol{\theta},\boldsymbol{t})\, \mathrm{d}t = -\frac{2}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n}\int_{\mathbb{R}} \mathrm{Re}(\vartheta_j)\, \mathrm{d}t, \tag{5}$$

where $\vartheta_j = [\Lambda_t\,(\cos(\delta_t)-\mathrm{i}\sin(\delta_t))-1]\,\mathrm{e}^{\mathrm{i}tX_j}$.

Developing the algebra and using the identity of the cosines in expression (5), we have:

$$\int_{\mathbb{R}} \nu_2(\boldsymbol{\theta},\boldsymbol{t})\, \mathrm{d}t = -\frac{2}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n} g_3(\boldsymbol{\theta};\kappa,x_j) + \frac{4}{n(\mathrm{e}^{\lambda}-1)}\sum_{j=1}^{n}\frac{\kappa}{\kappa^2+x_j^2},$$

where

$$g_3(\boldsymbol{\theta};\kappa,x_j)=\int_{\mathbb{R}}\cos(\delta_t-t\,x_j)\,\Lambda_t\,\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t.$$

Finally for the third case we have:

$$\int_{\mathbb{R}}\nu_3(\boldsymbol{\theta},\boldsymbol{t})\;\mathrm{d}t=\int_{\mathbb{R}}\frac{1}{(\mathrm{e}^{\lambda}-1)^2}\left[\Lambda_t\,(\cos(\delta_t)+\mathrm{i}\sin(\delta_t))-1\right]\left[\Lambda_t\,(\cos(\delta_t)-\mathrm{i}\sin(\delta_t))-1\right]\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t.$$

Again developing the algebra of this expression we have:

$$\begin{aligned}\int_{\mathbb{R}}\nu_3(\boldsymbol{\theta},\boldsymbol{t})\;\mathrm{d}t&=\frac{1}{(\mathrm{e}^{\lambda}-1)^2}\left[\int_{\mathbb{R}}\Lambda_t^2\,\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t+\int_{\mathbb{R}}\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t\right]-\frac{2}{(\mathrm{e}^{\lambda}-1)^2}\int_{\mathbb{R}}\Lambda_t\cos(\delta_t)\,\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t\\&=\frac{1}{(\mathrm{e}^{\lambda}-1)^2}g_1(\boldsymbol{\theta};\kappa)-\frac{2}{(\mathrm{e}^{\lambda}-1)^2}g_2(\boldsymbol{\theta};\kappa)+\frac{2}{\kappa(\mathrm{e}^{\lambda}-1)^2},\end{aligned}$$

where

$$g_1(\boldsymbol{\theta};\kappa)=\int_{\mathbb{R}}\mathrm{e}^{2\lambda\cos(\mu t)\,\mathrm{e}^{-\gamma|t|}}\,\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t$$

and

$$g_2(\boldsymbol{\theta};\kappa)=\int_{\mathbb{R}}\Lambda_t\,\cos(\delta_t)\,\mathrm{e}^{-\kappa|t|}\;\mathrm{d}t,$$

as desired.

Now suppose that $\mu=0$ and $\gamma=1$, then we holds

$$g_1(\lambda,1,0;\kappa)=\int_{\mathbb{R}}\mathrm{e}^{-\kappa|t|+2\lambda\mathrm{e}^{-|t|}}\;\mathrm{d}t=g_1^{-}(\lambda,1,0;\kappa)+g_1^{+}(\lambda,1,0;\kappa),$$

where

$$g_1^{-}(\lambda,1,0;\kappa)=\int_{-\infty}^{0}\mathrm{e}^{\kappa t+2\lambda\mathrm{e}^{t}}\;\mathrm{d}t.$$

With the change of variable $u=\mathrm{e}^{t}$ and using the software `wxMaxima`

$$g_1^{-}(\lambda,1,0;\kappa)=\int_0^1\mathrm{e}^{2\lambda u}\,u^{\kappa-1}\,\mathrm{du}=\frac{\Gamma(\kappa)-\Gamma(\kappa,-2\lambda)}{(-2\lambda)^{\kappa}},$$

where $\Gamma(\kappa,-2\lambda)=\int_{-2\lambda}^{\infty}t^{\kappa-1}\,\mathrm{e}^{-t}\;\mathrm{d}t$. For $g_1^{+}(\lambda,1,0;\kappa)$ the resolution is identical to the previous one. Hence,

$$g_1(\lambda,1,0;\kappa)=\frac{2}{(-2\lambda)^{\kappa}}\left[\Gamma(\kappa)-\Gamma(\kappa,-2\lambda)\right].$$

Using a similar procedure for $g_2(\lambda,1,0;\kappa)$, we get

$$g_2(\lambda,1,0;\kappa)=\frac{2}{(-\lambda)^{\kappa}}\left[\Gamma(\kappa)-\Gamma(\kappa,-\lambda)\right].$$

Finally for $g_3(\lambda,1,0;\kappa,x_j)$ we have:

$$\begin{aligned} g_3(\lambda,1,0;\kappa,x_j) &= \int_{\mathbb{R}} \cos(tx_j)\,\mathrm{e}^{\lambda \mathrm{e}^{-|t|}}\,\mathrm{e}^{-\kappa|t|}\,\mathrm{d}t \\ &= g_3^{-}(\lambda,1,0;\kappa,x_j)+g_3^{+}(\lambda,1,0;\kappa,x_j). \end{aligned}$$

For $g_3^{-}(\lambda,1,0;\kappa,x_j)$, let $u=\mathrm{e}^{t}$, then

$$g_3^{-}(\lambda,1,0;\kappa,x_j)=\int_0^1 \cos(x_j\log(u))\,\mathrm{e}^{\lambda u}\,u^{\kappa-1}\,\mathrm{du}.$$

As $\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}=\mathrm{e}^{\lambda}$, we have:

$$g_3^{-}(\lambda,1,0;\kappa,x_j)=\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}\int_0^1 \cos(x_j\log(u))u^{\kappa+r-1}\,\mathrm{du}.$$

Using the software `wxMaxima`, we get:

$$g_3^{-}(\lambda,1,0;\kappa,x_j)=\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}\frac{(k+r)}{x_j^2+(k+r)^2}.$$

For $g_3^{+}(\lambda,1,0;\kappa,x_j)$ the process is analogous to the previous one. Thus,

$$g_3(\lambda,1,0;\kappa,x_j)=2\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}\frac{(k+r)}{x_j^2+(k+r)^2}.$$

■

Appendix E - The $\Sigma(\theta)$ matrix of Theorem 5

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \mathrm{Cov}(a_1,a_1) & \mathrm{Cov}(a_1,a_2) & \dots & \mathrm{Cov}(a_1,a_k) & \mathrm{Cov}(a_1,b_1) & \dots & \mathrm{Cov}(a_1,b_k) \\ \cdot & \mathrm{Cov}(a_2,a_2) & \dots & \mathrm{Cov}(a_2,a_k) & \mathrm{Cov}(a_2,b_1) & \dots & \mathrm{Cov}(a_2,b_k) \\ \cdot & \cdot & \dots & \mathrm{Cov}(a_3,a_k) & \mathrm{Cov}(a_3,b_1) & \dots & \mathrm{Cov}(a_3,b_k) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \dots & \mathrm{Cov}(a_k,a_k) & \mathrm{Cov}(a_k,b_1) & \dots & \mathrm{Cov}(a_k,b_k) \\ \cdot & \cdot & \dots & \cdot & \mathrm{Cov}(b_1,b_1) & \dots & \mathrm{Cov}(b_1,b_k) \\ \cdot & \cdot & \dots & \cdot & \mathrm{Cov}(b_2,b_1) & \dots & \mathrm{Cov}(b_2,b_k) \\ \cdot & \cdot & \dots & \cdot & \mathrm{Cov}(b_3,b_1) & \dots & \mathrm{Cov}(b_3,b_k) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & \cdot & \mathrm{Cov}(b_k,b_1) & \dots & \mathrm{Cov}(b_k,b_k) \end{bmatrix}, \tag{6}$$

where, $\forall\, i = j \in \mathbb{N}$ we have:

$$\begin{aligned} \mathrm{Cov}(a_i,a_j) &= \frac{1}{2n}\left\{\mathrm{Re}[\Phi_S(2t_i)] + 1 - 2(\mathrm{Re}[\Phi_S(t_i)])^2\right\}, \\ \mathrm{Cov}(a_i,b_j) &= \frac{1}{2n}\left\{\mathrm{Im}[\Phi_S(2t_i)] - 2\mathrm{Re}[\Phi_S(t_i)]\mathrm{Im}[\Phi_S(t_i)]\right\}, \\ \mathrm{Cov}(b_i,b_j) &= \frac{1}{2n}\left\{1 - \mathrm{Re}[\Phi_S(2t_i)] - 2(\mathrm{Im}[\Phi_S(t_i)])^2\right\}, \end{aligned}$$

where, $\forall\, i \neq j \in \mathbb{N}$ we have:

$$\begin{aligned} \mathrm{Cov}(a_i,a_j) &= \frac{1}{2n}\left\{\mathrm{Re}[\Phi_S(t_i+t_j)] + \mathrm{Re}[\Phi_S(t_i-t_j)] - 2\mathrm{Re}[\Phi_S(t_i)]\mathrm{Re}[\Phi_S(t_j)]\right\}, \\ \mathrm{Cov}(a_i,b_j) &= \frac{1}{2n}\left\{\mathrm{Im}[\Phi_S(t_i+t_j)] - \mathrm{Im}[\Phi_S(t_i-t_j)] - 2\mathrm{Re}[\Phi_S(t_i)]\mathrm{Im}[\Phi_S(t_j)]\right\}, \\ \mathrm{Cov}(b_i,b_j) &= \frac{1}{2n}\left\{\mathrm{Re}[\Phi_S(t_i-t_j)] - \mathrm{Re}[\Phi_S(t_i-t_j)] - 2\mathrm{Im}[\Phi_S(t_i)]\mathrm{Im}[\Phi_S(t_j)]\right\}. \end{aligned}$$

Appendix F - Proof of Corollary 7

We proof in this appendix the Corollary 6, being similar to the previous proof. If $Z \sim \mathcal{G}(\alpha, \gamma, \mathcal{L})$ and $T = (\gamma + \mathcal{L}z)$, then by definition of cdf, we have:

$$F_T(t) = P(T \leq t) = P\left(z \leq \frac{t-\gamma}{\mathcal{L}}\right) = F_Z\left(\frac{t-\gamma}{\mathcal{L}}\right), \tag{7}$$

Deriving both sides the expression (7) with respect to variable T, we have:

$$f_T(t) = \frac{1}{\mathcal{L}} f_Y\left(\frac{t-\gamma}{\mathcal{L}}\right) = \frac{\Gamma(\mathcal{L}-\alpha)}{\Gamma(-\alpha)\Gamma(\mathcal{L})} \frac{t^{\alpha-1}}{\gamma^{\alpha}} \left[1 - \frac{\gamma}{t}\right]^{\mathcal{L}-1},$$

where $t \in (\gamma, \infty)$. It is now we are ready to find the nth expected value.

$$\mathrm{E}\left[\frac{1}{T^n}\right] = \int_{\gamma}^{\infty} t^{-n} f_T(t)\, \mathrm{d}t = \int_{\gamma}^{\infty} \frac{\Gamma(\mathcal{L}-\alpha)}{\Gamma(-\alpha)\Gamma(\mathcal{L})} \frac{t^{\alpha-n-1}}{\gamma^{\alpha}} \left[1 - \frac{\gamma}{t}\right]^{\mathcal{L}-1} \mathrm{d}t.$$

With the change of variable $s = \gamma/t$ we have:

$$\mathrm{E}\left[\frac{1}{T^k}\right] = \frac{\Gamma(\mathcal{L}-\alpha)}{\Gamma(-\alpha)\Gamma(\mathcal{L})} \frac{1}{\gamma^k} \int_0^1 s^{-\alpha+k-1} (1-s)^{\mathcal{L}-1}\, \mathrm{ds}.$$

By result of previous appendix

$$\mathrm{E}\left[\frac{1}{T^n}\right] = \frac{1}{\gamma^n} \frac{B(-\alpha+n, \mathcal{L})}{B(-\alpha, \mathcal{L})} = \frac{1}{\gamma^n} \prod_{k=0}^{n-1} \left(\frac{-\alpha+k}{-\alpha+\mathcal{L}+k}\right).$$

■

Appendix G - Fisher's information matrix and its inverse

Now we gain the score function and the Fisher information matrix for the parameter vector $\boldsymbol{\theta}$. From expression (6.5), for $i = 1, 2, \ldots, p$, and from Searle [257], the following identity holds:

$$U_{\boldsymbol{\beta}} = \sum_{k=1}^{n} \frac{\partial \ell_k(\boldsymbol{\theta})}{\partial \beta_j} = \sum_{k=1}^{n} \frac{\partial \ell_k(\boldsymbol{\theta})}{\partial \mu_k} \frac{\partial \mu_k}{\eta_k} \frac{\eta_k}{\partial \beta_j} = \alpha \sum_{k=1}^{n} \left[\frac{(\alpha - \mathcal{L})(-\alpha - 1)}{T_k} - \frac{1}{\mu_k} \right] \frac{1}{g'(\mu_k)} \boldsymbol{x}_{kj}. \tag{8}$$

We next obtain the matrix expression for the score function for $\boldsymbol{\beta}$ that is given in equation (6.6). For the α parameter, we hold:

$$U_{\alpha} = \sum_{k=1}^{n} \frac{\partial \ell_k(\boldsymbol{\theta})}{\partial \alpha} = nU_1(\alpha, \mathcal{L}) + \sum_{k=1}^{n} \log\left(\frac{T_k}{\mu_k}\right) - (\alpha - \mathcal{L}) \sum_{k=1}^{n} \frac{\mu_k}{T_k}. \tag{9}$$

Similarly, it can be shown that the score function for $\mathcal{L}$ can be written as:

$$U_{\mathcal{L}} = nU_2(\alpha, \mathcal{L}) + \sum_{k=1}^{n} \log\left(\frac{z_k}{T_k}\right) + (\alpha - \mathcal{L}) \sum_{k=1}^{n} \frac{z_k}{T_k}. \tag{10}$$

From regularity conditions, it is known that the expected value of the derivative in equation (6.5) equals zero.

From expression (8), the Hessian function at the terms β_j and β_l in $\boldsymbol{\beta}$ is

$$\begin{aligned} U_{\boldsymbol{\beta}\boldsymbol{\beta}} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} = \sum_{k=1}^{n} \frac{\partial}{\partial \mu_k} \left[\frac{\partial \ell_k(\boldsymbol{\theta})}{\partial \mu_k} \frac{\partial \mu_k}{\partial \eta_k} \right] \frac{\partial \mu_k}{\partial \eta_k} \frac{\partial \eta_k}{\partial \beta_j} x_{kl} \\ &= \sum_{k=1}^{n} \left[\frac{\partial^2 \ell_k(\boldsymbol{\theta})}{\partial \mu_k^2} \frac{\partial \mu_k}{\partial \eta_k} + \frac{\partial \ell_k(\boldsymbol{\theta})}{\partial \mu_k} \frac{\partial}{\partial \mu_k} \frac{\partial \mu_k}{\partial \eta_k} \right] \frac{\partial \mu_k}{\partial \eta_k} x_{kj} x_{kl}. \end{aligned} \tag{11}$$

Since $\mathrm{E}\left(\partial \ell_k(\boldsymbol{\theta}) / \partial \mu_k\right) = 0$, we keep

$$\mathrm{E}\left(U_{\boldsymbol{\beta}\boldsymbol{\beta}}\right) = \sum_{k=1}^{n} \mathrm{E}\left[\frac{\partial^2 \ell_k(\boldsymbol{\theta})}{\partial \mu_k^2}\right] \left(\frac{\partial \mu_k}{\partial \eta_k}\right)^2 x_{kj} x_{kl}.$$

Using the expression (6.13), we have

$$\mathrm{E}\left(U_{\boldsymbol{\beta}\boldsymbol{\beta}}\right) = \sum_{k=1}^{n} \left[\frac{\alpha}{\mu_k^2} + c_1(-\alpha - 1)\,\mathrm{E}\left(\frac{1}{T_k^2}\right) \right] \frac{x_{kj} x_{kl}}{\{g'(\mu_k)\}^2},$$

and from Corollary 7 we have

$$\mathrm{E}\left(\frac{1}{T_k^2}\right) = \frac{1}{\mu_k^2(-\alpha-1)}\frac{\alpha(\alpha-1)}{(\mathcal{L}-\alpha+1)c_1},$$

we retain:

$$\mathrm{E}\left(U_{\boldsymbol{\beta}\boldsymbol{\beta}}\right) = \alpha\left(\frac{\mathcal{L}}{\mathcal{L}-\alpha+1}\right)\sum_{k=1}^{n}\frac{1}{\mu_k^2}\frac{1}{\{g'(\mu_k)\}^2}x_{kj}x_{kl},$$

see in equation (6.9) the matrix form. From expression (8), the Hessian function at the $\boldsymbol{\beta}$ and α can be written as

$$U_{\boldsymbol{\beta}^\top\alpha} = \frac{\partial[U_{\boldsymbol{\beta}^\top}]}{\partial\alpha} = \sum_{k=1}^{n}\left[\frac{(\alpha-\mathcal{L})(-\alpha-1)}{T_k^2} - \frac{2\alpha+1-\mathcal{L}}{T_k\mu_k} - \frac{1}{\mu_k^2}\right]\frac{\mu_k x_{kj}}{g'(\mu_k)}.$$

Hence, applying the expected value, we own:

$$\mathrm{E}(U_{\boldsymbol{\beta}^\top\alpha}) = \sum_{k=1}^{n}\left[(\alpha-\mathcal{L})(-\alpha-1)\,\mathrm{E}\left(\frac{1}{T_k^2}\right) - \frac{2\alpha+1-\mathcal{L}}{\mu_k}\mathrm{E}\left(\frac{1}{T_k}\right) - \frac{1}{\mu_k^2}\right]\frac{\mu_k x_{kj}}{g'(\mu_k)}.$$

Again using the Corollary 7, we have:

$$\mathrm{E}(U_{\boldsymbol{\beta}^\top\alpha}) = c_2\sum_{k=1}^{n}\frac{1}{g'(\mu_k)}\frac{1}{\mu_k}x_{kj},$$

where

$$c_2 = -\left[1 + \frac{(2\alpha+1-\mathcal{L})\alpha}{(-\alpha-1)(\alpha-\mathcal{L})} + \frac{\alpha(\alpha-1)}{(-\alpha-1)(\mathcal{L}-\alpha+1)}\right],$$

The matricial expression is in equation (6.9). From expression (8), the Hessian function at the $\boldsymbol{\beta}$ and $\mathcal{L}$ can be written as

$$U_{\boldsymbol{\beta}^\top\mathcal{L}} = \frac{\partial[U_{\boldsymbol{\beta}^\top}]}{\partial\mathcal{L}} = -(-\alpha-1)\sum_{k=1}^{n}\left[\frac{T_k + (\alpha-\mathcal{L})z_k}{T_k^2}\right]\frac{1}{g'(\mu_k)}x_{kj}.$$

Applying the expected value, we get:

$$\mathrm{E}\left(U_{\boldsymbol{\beta}^\top\mathcal{L}}\right) = -(-\alpha-1)\sum_{k=1}^{n}\left[\mathrm{E}\left(\frac{1}{T_k}\right) + (\alpha-\mathcal{L})\,\mathrm{E}\left(\frac{Z_k}{T_k^2}\right)\right]\frac{1}{g'(\mu_k)}x_{kj}.$$

We derived that

$$\mathrm{E}\left[\frac{1}{T_k}\right] = -\frac{\alpha}{\mu_k(-\alpha-1)(\mathcal{L}-\alpha)}. \tag{12}$$

Then derivative both sides the expression (12) with respect to the $\mathcal{L}$, we have

$$\mathrm{E}\left[\frac{Z_n}{T_k^2}\right] = -\frac{\alpha}{\mu_k(-\alpha-1)(\mathcal{L}-\alpha)^2}. \tag{13}$$

With this

$$\mathrm{E}\left(U_{\boldsymbol{\beta}^\top \mathcal{L}}\right) = \mathbf{0}.$$

For obtain $U_{\alpha\alpha}$ we using the expression (9), this is,

$$U_{\alpha\alpha} = \frac{\partial[U_\alpha]}{\partial\alpha} = nU_1^{(1)}(\alpha,\mathcal{L}) - 2\sum_{k=1}^{n}\frac{\mu_k}{T_k} - (\alpha-\mathcal{L})\sum_{k=1}^{n}\frac{\mu_k^2}{T_k^2},$$

where $U_1^{(1)}(\alpha,\mathcal{L})$ is the first derivate of $U_1(\alpha,\mathcal{L})$ with respect to the parameter α. Applying the expected value in expression above

$$\mathrm{E}\left(U_{\alpha\alpha}\right) = nU_1^{(1)}(\alpha,\mathcal{L}) - 2\sum_{k=1}^{n}\mu_k\,\mathrm{E}\left(\frac{1}{T_k}\right) - (\alpha-\mathcal{L})\sum_{k=1}^{n}\mu_k^2\,\mathrm{E}\left(\frac{1}{T_k^2}\right).$$

Using the Corollary 7 in this expression, we hold:

$$\mathrm{E}\left(U_{\alpha\alpha}\right) = nc_3,$$

where

$$c_3 = U_1^{(1)}(\alpha,\mathcal{L}) + \frac{2\alpha}{(\mathcal{L}-\alpha)(-\alpha-1)} + \frac{\alpha(\alpha-1)}{(-\alpha-1)^2(\mathcal{L}-\alpha+1)}.$$

From expression (9), the second derivative of $\ell(\boldsymbol{\theta})$ with respect to the $\mathcal{L}$ can be written as

$$U_{\alpha\mathcal{L}} = \frac{\partial U_\alpha}{\partial\mathcal{L}} = -n\Psi^{(1)}(\mathcal{L}-\alpha) + \sum_{k=1}^{n}\frac{z_k}{T_k} + \sum_{k=1}^{n}\frac{\mu_k}{T_k} + (\alpha-\mathcal{L})\sum_{k=1}^{n}\mu_k\frac{z_k}{T_k^2},$$

where $\Psi^{(k)}(x) = {}^{\partial^{k+1}\log\Gamma(x)}\!/_{\partial x^{k+1}}$ for $x>0$. Applying the expected value, we own:

$$\mathrm{E}\left(U_{\alpha\mathcal{L}}\right) = -n\Psi^{(1)}(\mathcal{L}-\alpha) + \sum_{k=1}^{n}\mathrm{E}\left(\frac{Z_k}{T_k}\right) + \sum_{k=1}^{n}\mu_k\,\mathrm{E}\left(\frac{1}{T_k}\right) + (\alpha-\mathcal{L})\sum_{k=1}^{n}\mu_k\,\mathrm{E}\left(\frac{Z_k}{T_k^2}\right). \tag{14}$$

Then, for expression $\mathrm{E}\left({}^{Z_k}\!/_{T_k}\right)$ we will wake the fact that $\mathrm{E}\left(U_\alpha\right)=0$, that is,

$$\sum_{k=1}^{n}\mathrm{E}\left[\log\left(T_k\right)\right] = -nU_1(\alpha,\mathcal{L}) + \sum_{k=1}^{n}\log(\mu_k) + (\alpha-\mathcal{L})\sum_{k=1}^{n}\mu_k\,\mathrm{E}\left(\frac{1}{T_k}\right).$$

Derivative both sides the expression above with respect the $\mathcal{L}$, we retain

$$\sum_{k=1}^{n} \mathrm{E}\left(\frac{Z_k}{T_k}\right) = n\Psi^{(1)}(\mathcal{L}-\alpha) - \sum_{k=1}^{n} \mu_k \mathrm{E}\left(\frac{1}{T_k}\right) - (\alpha-\mathcal{L}) \sum_{k=1}^{n} \mu_k \mathrm{E}\left(\frac{Z_k}{T_k^2}\right). \tag{15}$$

Hence, using the expression (15) in (14), we have:

$$\mathrm{E}(U_{\alpha\mathcal{L}}) = \mathbf{0}.$$

Ultimately, we have for $U_{\mathcal{L}\mathcal{L}}$ using the expression (10), this is,

$$U_{\mathcal{L}\mathcal{L}} = \frac{\partial U_{\mathcal{L}}}{\partial \mathcal{L}} = n\left[\Psi^{(1)}(\mathcal{L}-\alpha) - \Psi^{(1)}(\mathcal{L}) + \frac{1}{\mathcal{L}}\right] - 2\sum_{k=1}^{n} \frac{z_k}{T_k} - (\alpha-\mathcal{L}) \sum_{k=1}^{n} \left(\frac{z_k}{T_k}\right)^2.$$

Applying the expected value, we hold:

$$\mathrm{E}(U_{\mathcal{L}\mathcal{L}}) = n\left[\Psi^{(1)}(\mathcal{L}-\alpha) - \Psi^{(1)}(\mathcal{L}) + \frac{1}{\mathcal{L}}\right] - 2\sum_{k=1}^{n} \mathrm{E}\left(\frac{Z_k}{T_k}\right) - (\alpha-\mathcal{L}) \sum_{k=1}^{n} \mathrm{E}\left[\left(\frac{Z_k}{T_k}\right)^2\right].$$

Through of (15) and the expressions (12) and (13), we obtain:

$$\sum_{k=1}^{n} \mathrm{E}\left(\frac{Z_k}{T_k}\right) = n\Psi^{(1)}(\mathcal{L}-\alpha).$$

Derivative both sides of the expression above with respect the $\mathcal{L}$, we get:

$$\sum_{k=1}^{n} \mathrm{E}\left[\left(\frac{Z_k}{T_k}\right)^2\right] = -n\Psi^{(2)}(\mathcal{L}-\alpha).$$

Soon,

$$\mathrm{E}(U_{\mathcal{L}\mathcal{L}}) = nc_4,$$

where

$$c_4 = \left[(\alpha-\mathcal{L})\Psi^{(2)}(\mathcal{L}-\alpha) - \Psi^{(1)}(\mathcal{L}-\alpha) - \Psi^{(1)}(\mathcal{L}) + \frac{1}{\mathcal{L}}\right].$$

Therefore follows the Fisher information matrix for $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top, \mathcal{L})$ that is given in equation (6.9).

For obtain the inverse of $K(\boldsymbol{\theta})$, we partition this matrix of following form:

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \\ \boldsymbol{B}^\top & \boldsymbol{D} \end{bmatrix},$$

where

$$\boldsymbol{A} = \begin{bmatrix} \alpha \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X} & c_2 \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{\mu}^* \\ \bullet & nc_3 \end{bmatrix},$$

$$\boldsymbol{B}^\top = \boldsymbol{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \end{bmatrix}, \text{ and } \boldsymbol{D} = nc_4.$$

It is thus useful the matrices can be inverted block-wise (e.g. Rao [258, p. 33], and for more informations see Rencher and Schaalje [259]), this is:

$$K^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{A}^{-1} + \boldsymbol{\upsilon} \boldsymbol{\Phi}^{-1} \boldsymbol{\upsilon}^\top & -\boldsymbol{\upsilon} \boldsymbol{\Phi}^{-1} \\ -\boldsymbol{\Phi}^{-1} \boldsymbol{\upsilon}^\top & \boldsymbol{\Phi}^{-1} \end{pmatrix},$$

where $\boldsymbol{\Phi} = \boldsymbol{D} - \boldsymbol{B}^\top \boldsymbol{A}^{-1} \boldsymbol{B}$, $\boldsymbol{\upsilon} = \boldsymbol{A}^{-1} \boldsymbol{B}$. As $\boldsymbol{B}^\top = \boldsymbol{B} = [\mathbf{0} \ \mathbf{0}]$, we own:

$$K^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (nc_4)^{-1} \end{pmatrix},$$

with,

$$\boldsymbol{A}^{-1} = \begin{pmatrix} (\alpha \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X})^{-1} + \boldsymbol{\zeta} \boldsymbol{\vartheta}^{-1} \boldsymbol{\zeta}^\top & -\boldsymbol{\zeta} \boldsymbol{\vartheta}^{-1} \\ -\boldsymbol{\vartheta}^{-1} \boldsymbol{\zeta}^\top & \boldsymbol{\vartheta}^{-1} \end{pmatrix},$$

where

$$\boldsymbol{\vartheta} = nc_3 - \frac{c_2^2}{\alpha} (\boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{\mu}^*)^\top (\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X})^{-1} (\boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{\mu}^*),$$

and

$$\boldsymbol{\zeta} = \frac{c_2}{\alpha} (\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X})^{-1} (\boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{\mu}^*).$$

Appendix H - Diagnostic measures

In this appendix we obtain in details generalized leverage for $(\boldsymbol{\beta}^\top, \alpha, \mathcal{L})$. The notation used here is defined in the Diagnostic measures section. Thus, for $i = 1, 2, \ldots, p$, and as given in (8), we have

$$D_{\boldsymbol{\beta}} = \sum_{k=1}^{n} \frac{\partial \mu_k}{\partial \beta_j} = \sum_{k=1}^{n} \frac{1}{g'(\mu_k)} x_{kj} = \boldsymbol{EX}.$$

From equation (11) it follows that

$$\begin{aligned} -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} &= \sum_{k=1}^{n} \left\{ \left[\frac{\alpha}{\mu_k^2} + \frac{c_1(-\alpha - 1)}{T_k^2} \right] + \left[\frac{\alpha}{\mu_k} + \frac{c_1}{T_k} \right] \frac{g''(\mu_k)}{g'(\mu_k)} \right\} \frac{1}{\{g'(\mu_k)\}^2} x_{kj} x_{kl} \\ &= \alpha \boldsymbol{X}^\top \boldsymbol{Q} \boldsymbol{X}, \end{aligned}$$

Also, it can be shown that

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{z}^\top} &= \sum_{k=1}^{n} \frac{\partial \mu_k}{\partial g(\mu_k)} \frac{\partial g(\mu_k)}{\partial \beta_j} \left[\frac{\partial}{\partial \mu_k} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{z}^\top} \right] \\ &= \alpha \sum_{k=1}^{n} \frac{x_{kj}}{g'(\mu_k)} \frac{\mathcal{L}}{T_k} \frac{c_1}{\alpha} \\ &= \alpha \boldsymbol{X}^\top \boldsymbol{E} \boldsymbol{T}^*, \end{aligned}$$

Thus, its follows the result.

Appendix I - Numerical results of the pilot simulation

Following are the table and graphs of the pilot test to choose the numerical iterative method to estimate the parameters of the regression model of the $\mathcal{G}$ distribution. After, its have the tables of the first and third scenario of Monte Carlo simulation.

Tabela 1: Parameter estimates using the model addressed $\mu_k = e^{\beta_0+\beta_1 x_{1,k}+\beta_2 x_{2,k}}$, where $X_k \sim \mathcal{G}(\alpha, (-\alpha-1), \mathcal{L})$

Parameter	n	BFGS		CG		NM		SANN	
		Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
	20	−0.0370	0.4812	−0.0135	0.5571	−0.0187	0.6002	−0.0333	0.4742
β_0	50	−0.0221	0.2783	−0.0193	0.2761	−0.0221	0.2784	−0.0189	0.2784
	100	−0.0252	0.2015	−0.0227	0.1996	−0.0252	0.2015	−0.0289	0.2037
	500	0.0149	0.0792	0.0149	0.0792	0.0149	0.0792	0.0170	0.0802
	20	0.0297	0.3350	0.0298	0.3324	0.0300	0.3353	0.0296	0.3338
β_1	50	0.0282	0.1545	0.0288	0.1550	0.0282	0.1545	0.0309	0.1584
	100	0.0158	0.1039	0.0157	0.1039	0.0157	0.1039	0.0166	0.1035
	500	0.0123	0.0360	0.0123	0.0360	0.0123	0.0360	0.0119	0.0378
	20	−0.0095	0.2961	−0.0086	0.2926	−0.0092	0.2964	−0.0122	0.2934
β_2	50	0.0171	0.1574	0.0170	0.1566	0.0171	0.1574	0.0153	0.1565
	100	−0.0060	0.0875	−0.0058	0.0877	−0.0060	0.0875	−0.0036	0.0864
	500	0.0067	0.0371	0.0067	0.0371	0.0067	0.0371	0.0068	0.0374
	20	−16.8667	27.2267	−5.3176	3.8466	−40.3654	76.4936	−6.9478	6.5823
$\alpha = -3$	50	−6.1848	12.5684	−4.0622	2.2177	−7.0349	18.5315	−4.3065	2.9810
	100	−4.3723	4.3521	−3.7270	1.7593	−4.3793	4.3947	−3.9237	2.3201
	500	−3.0705	0.3938	−3.0696	0.3920	−3.0705	0.3940	−3.0572	0.3949
	20	−0.0992	0.4648	−0.0870	0.4569	−0.1005	0.4661	−0.0922	0.4572
β_0	50	−0.0278	0.2556	−0.0203	0.2523	−0.0279	0.2556	−0.0209	0.2578
	100	−0.0004	0.1670	0.0034	0.1652	−0.0005	0.1670	0.0008	0.1667
	500	0.0120	0.0607	0.0124	0.0606	0.0120	0.0606	0.0102	0.0620
	20	0.0216	0.3150	0.0267	0.3142	0.0208	0.3154	0.0270	0.3161
β_1	50	0.0107	0.1523	0.0123	0.1521	0.0107	0.1523	0.0060	0.1543
	100	0.0143	0.0997	0.0144	0.0990	0.0143	0.0997	0.0146	0.0992
	500	0.0033	0.0387	0.0033	0.0386	0.0033	0.0387	0.0029	0.0398
	20	0.0673	0.3103	0.0656	0.3072	0.0684	0.3115	0.0615	0.3049
β_2	50	0.0231	0.1451	0.0238	0.1447	0.0230	0.1450	0.0250	0.1477
	100	0.0003	0.0948	0.0008	0.0945	0.0003	0.0948	−0.0004	0.0961
	500	0.0066	0.0411	0.0066	0.0411	0.0066	0.0411	0.0072	0.0423
	20	−34.8726	46.2230	−8.2580	5.8834	−75.1931	104.7882	−10.0604	7.9690
$\alpha = -5$	50	−18.2678	32.1539	−6.5404	2.9226	−29.0835	57.8295	−7.9320	5.1603

	100	−9.1406	11.1039	−6.0007	2.3826	−11.3666	21.9571	−6.9239	3.9901
	500	−5.1647	0.9517	−5.1223	0.8737	− 5.1641	0.9515	−5.2155	0.9740
	20	−0.0216	0.4107	−0.0105	0.4066	−0.0220	0.4112	−0.0081	0.4189
β_0	50	−0.0362	0.2344	−0.0304	0.2285	−0.0366	0.2351	−0.0363	0.2331
	100	−0.0263	0.1828	−0.0211	0.1794	−0.0263	0.1832	−0.0190	0.1844
	500	0.0183	0.0674	0.0217	0.0682	0.0184	0.0676	0.0230	0.0726
	20	−0.0098	0.2492	−0.0073	0.2484	−0.0106	0.2494	−0.0120	0.2526
β_1	50	0.0257	0.1301	0.0283	0.1290	0.0254	0.1303	0.0286	0.1335
	100	0.0162	0.1141	0.0166	0.1137	0.0162	0.1140	0.0158	0.1147
	500	0.0047	0.0450	0.0047	0.0448	0.0048	0.0450	0.0038	0.0450
	20	0.0108	0.3035	0.0136	0.2996	0.0104	0.3050	0.0125	0.3090
β_2	50	0.0101	0.1591	0.0108	0.1573	0.0102	0.1597	0.0119	0.1594
	100	0.0218	0.0931	0.0220	0.0931	0.0215	0.0932	0.0198	0.0944
	500	0.0024	0.0452	0.0026	0.0448	0.0024	0.0453	0.0031	0.0453
	20	−47.0713	50.9270	−10.4933	3.4479	−105.9656	123.4028	−13.0618	6.7257
$\alpha = -10$	50	−36.7495	43.1158	−10.0741	2.3237	−71.8729	96.7460	−12.4509	5.3565
	100	−27.3761	35.7574	−9.6007	2.2419	−45.6554	72.1477	−10.8695	3.9029
	500	−13.3979	9.7916	−9.4169	1.3951	−13.6746	10.9915	−9.6884	2.1194
	20	0.0334	0.3895	0.0309	0.3870	0.0287	0.3924	0.0306	0.3854
β_0	50	0.0195	0.2604	0.0186	0.2612	0.0212	0.2587	0.0166	0.2610
	100	0.0178	0.1801	0.0165	0.1794	0.0171	0.1802	0.0171	0.1816
	500	0.0054	0.0645	0.0052	0.0644	0.0053	0.0644	0.0048	0.0644
	20	−0.0530	0.2916	−0.0532	0.2883	−0.0515	0.2919	−0.0508	0.2855
β_1	50	−0.0129	0.1557	−0.0114	0.1529	−0.0135	0.1554	−0.0110	0.1528
	100	0.0042	0.1022	0.0051	0.1016	0.0038	0.1022	0.0038	0.1029
	500	0.0124	0.0491	0.0127	0.0491	0.0123	0.0491	0.0128	0.0487
	20	−0.0052	0.2724	−0.0027	0.2709	−0.0044	0.2756	−0.0019	0.2695
β_2	50	−0.0049	0.1642	−0.0060	0.1641	−0.0068	0.1635	−0.0039	0.1666
	100	−0.0075	0.1008	−0.0075	0.1013	−0.0068	0.1014	−0.0070	0.1031
	500	0.0138	0.0440	0.0138	0.0440	0.0138	0.0441	0.0136	0.0448
	20	−66.9966	39.5089	−49.6229	2.6470	−128.8739	103.8445	−51.9579	7.2810
$\alpha = -50$	50	−56.1634	30.7463	−48.9755	1.1506	−105.1543	92.2535	−49.7222	5.8457
	100	−63.0099	42.8523	−49.1635	2.1797	−101.3539	86.8564	−50.5785	5.3853
	500	−58.6298	34.1430	−49.0169	1.0980	−92.7803	78.5264	−47.9205	3.3888

Tabela 2: Parameter estimates using the model $\mu_k = e^{\beta_0+\beta_1 x_{1,k}+\beta_2 x_{2,k}}$, where $X_k \sim G(\alpha, (-\alpha-1), \mathcal{L})$, with $\boldsymbol{\beta} = (0.01, 0.01, 0.01)$

	Parameter	n	$\alpha=-3$			$\alpha=-5$			$\alpha=-10$			$\alpha=-15$		
			Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE
$\mathcal{L}=1$	β_0	20	0.0022	0.0078	0.5559	0.0011	0.0089	0.4891	0.0042	0.0058	0.4654	0.0128	0.0028	0.4592
		50	0.0261	0.0161	0.3118	0.0252	0.0152	0.2817	0.0360	0.0260	0.2698	0.0164	0.0064	0.2737
		100	0.0172	0.0072	0.2174	0.0156	0.0056	0.1935	0.0142	0.0042	0.1838	0.0197	0.0097	0.1807
		500	0.0124	0.0024	0.0934	0.0182	0.0082	0.0843	0.0171	0.0071	0.0802	0.0127	0.0027	0.0792
	β_1	20	−0.0260	0.0360	0.3559	−0.0266	0.0366	0.3360	−0.0251	0.0351	0.3225	−0.0205	0.0305	0.3178
		50	−0.0173	0.0273	0.1662	−0.0110	0.0210	0.1576	−0.0117	0.0217	0.1651	−0.0085	0.0185	0.1668
		100	0.0027	0.0073	0.1006	0.0008	0.0092	0.1088	0.0095	0.0005	0.1047	−0.0001	0.0101	0.1069
		500	0.0071	0.0029	0.0402	0.0053	0.0047	0.0437	0.0059	0.0041	0.0466	0.0077	0.0023	0.0462
	β_2	20	−0.0145	0.0245	0.3478	−0.0140	0.0240	0.3176	−0.0258	0.0358	0.3160	−0.0478	0.0578	0.3035
		50	−0.0073	0.0173	0.1644	−0.0052	0.0152	0.1610	−0.0153	0.0253	0.1647	−0.0075	0.0175	0.1676
		100	−0.0003	0.0103	0.1067	0.0025	0.0075	0.1122	−0.0034	0.0134	0.1096	−0.0022	0.0122	0.1107
		500	0.0072	0.0028	0.0398	0.0078	0.0022	0.0434	0.0079	0.0021	0.0439	0.0088	0.0012	0.0454
	α	20	−6.2574	3.2574	4.9443	−8.2877	3.2877	5.3225	−11.0602	1.0602	3.9170	−15.0471	0.0471	3.2909
		50	−4.5023	1.5023	2.8531	−6.2957	1.2957	2.9071	−9.6979	0.3021	2.6429	−14.3824	0.6176	1.9040
		100	−4.0021	1.0021	2.1548	−6.0878	1.0878	2.6569	−9.6738	0.3262	2.6916	−14.2610	0.7390	2.1281
		500	−3.1708	0.1708	0.6657	−5.2222	0.2222	1.2198	−9.3387	0.6613	1.4733	−14.0496	0.9504	1.5497
$\mathcal{L}=4$	β_0	20	−0.0260	0.0360	0.4205	−0.0175	0.0275	0.3945	0.0059	0.0041	0.3686	0.0223	0.0123	0.3879
		50	−0.0034	0.0134	0.2330	−0.0121	0.0221	0.2240	0.0130	0.0030	0.2255	0.0145	0.0045	0.2321
		100	−0.0041	0.0141	0.1642	0.0049	0.0051	0.1513	0.0137	0.0037	0.1569	0.0061	0.0039	0.1585
		500	0.0112	0.0012	0.0708	0.0078	0.0022	0.0650	0.0144	0.0044	0.0664	0.0150	0.0050	0.0639
	β_1	20	0.0028	0.0072	0.2853	0.0119	0.0019	0.2769	0.0187	0.0087	0.2668	0.0061	0.0039	0.2716
		50	0.0123	0.0023	0.1414	0.0135	0.0035	0.1434	0.0085	0.0015	0.1488	0.0041	0.0059	0.1491
		100	0.0129	0.0029	0.0860	0.0059	0.0041	0.0957	0.0090	0.0010	0.1063	0.0113	0.0013	0.1043
		500	0.0107	0.0007	0.0342	0.0096	0.0004	0.0394	0.0108	0.0008	0.0430	0.0090	0.0010	0.0438
	β_2	20	0.0041	0.0059	0.2883	0.0089	0.0011	0.2631	−0.0044	0.0144	0.2499	−0.0072	0.0172	0.2703
		50	0.0085	0.0015	0.1383	0.0218	0.0118	0.1402	0.0063	0.0037	0.1533	0.0092	0.0008	0.1526
		100	0.0130	0.0030	0.0845	0.0113	0.0013	0.0943	0.0071	0.0029	0.1052	0.0107	0.0007	0.1023
		500	0.0093	0.0007	0.0350	0.0120	0.0020	0.0397	0.0077	0.0023	0.0426	0.0062	0.0038	0.0442
	α	20	−4.9645	1.9645	3.3439	−7.3870	2.3870	4.0181	−10.7511	0.7511	3.4781	−14.8518	0.1482	3.1501
		50	−3.6681	0.6681	1.5700	−6.1732	1.1732	2.5879	−10.3547	0.3547	2.8542	−14.5246	0.4754	2.4713
		100	−3.3181	0.3181	0.8613	−5.5264	0.5264	1.5472	−9.5812	0.4188	1.5693	−14.1028	0.8972	1.2310
		500	−3.0427	0.0427	0.3141	−5.1035	0.1035	0.6612	−9.3989	0.6011	0.9508	−14.0035	0.9965	1.0914
$\mathcal{L}=8$	β_0	20	−0.0642	0.0742	0.3827	−0.0021	0.0121	0.3756	0.0210	0.0110	0.3825	0.0347	0.0247	0.3698
		50	−0.0100	0.0200	0.2170	0.0040	0.0060	0.1997	0.0095	0.0005	0.2145	0.0265	0.0165	0.2219
		100	0.0059	0.0041	0.1509	0.0062	0.0038	0.1460	0.0019	0.0081	0.1500	0.0221	0.0121	0.1541
		500	0.0075	0.0025	0.0647	0.0077	0.0023	0.0611	0.0101	0.0001	0.0644	0.0148	0.0048	0.0671
	β_1	20	0.0217	0.0117	0.2583	0.0069	0.0031	0.2622	−0.0021	0.0121	0.2669	−0.0020	0.0120	0.2593
		50	0.0129	0.0029	0.1295	0.0098	0.0002	0.1287	0.0070	0.0030	0.1470	0.0007	0.0093	0.1535
		100	0.0119	0.0019	0.0856	0.0114	0.0014	0.0926	0.0151	0.0051	0.1007	−0.0012	0.0112	0.1051
		500	0.0110	0.0010	0.0320	0.0118	0.0018	0.0375	0.0106	0.0006	0.0434	0.0079	0.0021	0.0466
	β_2	20	0.0276	0.0176	0.2586	0.0059	0.0041	0.2557	0.0044	0.0056	0.2566	0.0100	0.0200	0.2599
		50	0.0108	0.0008	0.1296	0.0080	0.0020	0.1359	0.0142	0.0042	0.1421	0.0040	0.0060	0.1513
		100	0.0063	0.0037	0.0813	0.0105	0.0005	0.0916	0.0122	0.0022	0.0984	0.0106	0.0006	0.1051
		500	0.0116	0.0016	0.0324	0.0109	0.0009	0.0378	0.0102	0.0002	0.0418	0.0084	0.0016	0.0431
	α	20	−4.6200	1.6200	2.7629	−7.0144	2.0144	3.4686	−11.1767	1.1767	3.6284	−14.9764	0.0236	2.8486
		50	−3.5154	0.5154	1.1697	−5.7896	0.7896	1.7891	−10.9839	0.9839	3.4258	−15.3677	0.3677	4.3034
		100	−3.2240	0.2240	0.6446	−5.3767	0.3767	1.1740	−9.9584	0.0416	1.6680	−14.3366	0.6634	1.5725
		500	−3.0425	0.0425	0.2524	−5.0636	0.0636	0.4599	−9.6433	0.3567	0.8117	−14.1706	0.8294	1.0660

Tabela 3: Parameter estimates using the model $\mu_k = e^{\beta_0+\beta_1 x_{1,k}+\beta_2 x_{2,k}}$, where $X_k \sim \mathcal{G}(\alpha, (-\alpha-1), \mathcal{L})$ with $\boldsymbol{\beta} = (2,2,2)$

	Parameter	n	$\alpha=-3$			$\alpha=-5$			$\alpha=-10$			$\alpha=-15$		
			Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE	Mean	Abias	RMSE
$\mathcal{L}=1$	β_0	20	1.9919	0.0081	0.5548	1.9923	0.0077	0.4912	1.9935	0.0065	0.4648	2.0003	0.0003	0.4574
		50	2.0157	0.0157	0.3090	2.0152	0.0152	0.2818	2.0264	0.0264	0.2696	2.0065	0.0065	0.2737
		100	2.0072	0.0072	0.2173	2.0056	0.0056	0.1936	2.0039	0.0039	0.1838	2.0096	0.0096	0.1808
		500	2.0007	0.0007	0.0930	2.0080	0.0080	0.0843	2.0069	0.0069	0.0801	2.0027	0.0027	0.0792
	β_1	20	1.9640	0.0360	0.3558	1.9632	0.0368	0.3359	1.9654	0.0346	0.3229	1.9703	0.0297	0.3181
		50	1.9728	0.0272	0.1662	1.9789	0.0211	0.1576	1.9782	0.0218	0.1650	1.9815	0.0185	0.1668
		100	1.9927	0.0073	0.1006	1.9908	0.0092	0.1088	1.9995	0.0005	0.1047	1.9899	0.0101	0.1068
		500	1.9979	0.0021	0.0393	1.9953	0.0047	0.0437	1.9960	0.0040	0.0466	1.9976	0.0024	0.0462
	β_2	20	1.9756	0.0244	0.3475	1.9760	0.0240	0.3174	1.9642	0.0358	0.3160	1.9429	0.0571	0.3020
		50	1.9826	0.0174	0.1643	1.9847	0.0153	0.1611	1.9747	0.0253	0.1651	1.9825	0.0175	0.1676
		100	1.9898	0.0102	0.1067	1.9926	0.0074	0.1122	1.9867	0.0133	0.1096	1.9877	0.0123	0.1107
		500	1.9971	0.0029	0.0406	1.9978	0.0022	0.0434	1.9980	0.0020	0.0439	1.9988	0.0012	0.0454
	α	20	−6.2741	3.2741	4.9574	−8.2286	3.2286	5.1440	−11.0468	1.0468	3.9221	−15.0153	0.0153	3.0416
		50	−4.4756	1.4756	2.7953	−6.3484	1.3484	2.9828	−9.7086	0.2914	2.6937	−14.3593	0.6407	1.9766
		100	−3.9868	0.9868	2.0859	−6.0915	1.0915	2.6772	−9.6877	0.3123	2.6652	−14.3964	0.6036	2.9943
		500	−3.1769	0.1769	0.6776	−5.2467	0.2467	1.2644	−9.3986	0.6014	1.5593	−14.0854	0.9146	1.9685
$\mathcal{L}=4$	β_0	20	1.9641	0.0359	0.4204	1.9726	0.0274	0.3942	1.9958	0.0042	0.3691	2.0119	0.0119	0.3880
		50	1.9867	0.0133	0.2330	1.9779	0.0221	0.2240	2.0029	0.0029	0.2257	2.0046	0.0046	0.2322
		100	1.9859	0.0141	0.1642	1.9950	0.0050	0.1515	2.0037	0.0037	0.1569	1.9961	0.0039	0.1585
		500	2.0012	0.0012	0.0708	1.9979	0.0021	0.0650	2.0044	0.0044	0.0664	2.0050	0.0050	0.0640
	β_1	20	1.9928	0.0072	0.2853	2.0017	0.0017	0.2769	2.0087	0.0087	0.2669	1.9961	0.0039	0.2715
		50	2.0023	0.0023	0.1413	2.0036	0.0036	0.1434	1.9986	0.0014	0.1488	1.9940	0.0060	0.1491
		100	2.0029	0.0029	0.0860	1.9959	0.0041	0.0957	1.9990	0.0010	0.1063	2.0014	0.0014	0.1042
		500	2.0007	0.0007	0.0342	1.9996	0.0004	0.0394	2.0008	0.0008	0.0430	1.9991	0.0009	0.0438
	β_2	20	1.9940	0.0060	0.2883	1.9988	0.0012	0.2630	1.9855	0.0145	0.2500	1.9830	0.0170	0.2704
		50	1.9985	0.0015	0.1383	2.0117	0.0117	0.1401	1.9962	0.0038	0.1534	1.9992	0.0008	0.1527
		100	2.0030	0.0030	0.0845	2.0013	0.0013	0.0943	1.9971	0.0029	0.1052	2.0007	0.0007	0.1022
		500	1.9993	0.0007	0.0350	2.0019	0.0019	0.0397	1.9977	0.0023	0.0425	1.9962	0.0038	0.0442
	α	20	−4.9565	1.9565	3.3054	−7.4466	2.4466	4.1412	−10.8153	0.8153	3.6596	−14.8284	0.1716	2.7841
		50	−3.6651	0.6651	1.5758	−6.1602	1.1602	2.5418	−10.3567	0.3567	2.8904	−14.5347	0.4653	2.8065
		100	−3.3187	0.3187	0.8619	−5.5098	0.5098	1.5153	−9.5672	0.4328	1.5599	−14.1085	0.8915	1.2521
		500	−3.0427	0.0427	0.3141	−5.1021	0.1021	0.6613	−9.3944	0.6056	0.9707	−13.9991	1.0009	1.0949
$\mathcal{L}=8$	β_0	20	1.9258	0.0742	0.3826	1.9878	0.0122	0.3755	2.0113	0.0113	0.3825	2.0245	0.0245	0.3697
		50	1.9800	0.0200	0.2170	1.9940	0.0060	0.1996	1.9993	0.0007	0.2144	2.0165	0.0165	0.2217
		100	1.9958	0.0042	0.1509	1.9962	0.0038	0.1460	1.9921	0.0079	0.1500	2.0121	0.0121	0.1540
		500	1.9975	0.0025	0.0647	1.9977	0.0023	0.0611	2.0001	0.0001	0.0644	2.0048	0.0048	0.0671
	β_1	20	2.0117	0.0117	0.2583	1.9969	0.0031	0.2622	1.9877	0.0123	0.2671	1.9879	0.0121	0.2593
		50	2.0029	0.0029	0.1296	1.9997	0.0003	0.1287	1.9970	0.0030	0.1470	1.9907	0.0093	0.1535
		100	2.0019	0.0019	0.0856	2.0014	0.0014	0.0926	2.0050	0.0050	0.1007	1.9888	0.0112	0.1051
		500	2.0010	0.0010	0.0320	2.0018	0.0018	0.0375	2.0006	0.0006	0.0434	1.9979	0.0021	0.0466
	β_2	20	2.0176	0.0176	0.2585	1.9960	0.0040	0.2557	1.9943	0.0057	0.2565	1.9805	0.0195	0.2600
		50	2.0008	0.0008	0.1296	1.9980	0.0020	0.1359	2.0043	0.0043	0.1420	1.9940	0.0060	0.1513
		100	1.9963	0.0037	0.0813	2.0005	0.0005	0.0916	2.0022	0.0022	0.0984	2.0006	0.0006	0.1051
		500	2.0016	0.0016	0.0324	2.0009	0.0009	0.0378	2.0002	0.0002	0.0417	1.9984	0.0016	0.0432
	α	20	−4.6174	1.6174	2.7650	−7.0183	2.0183	3.4802	−11.1248	1.1248	3.3803	−14.9811	0.0189	2.8060
		50	−3.5141	0.5141	1.1662	−5.7939	0.7939	1.8148	−11.0345	1.0345	3.5124	−15.2890	0.2890	4.0108
		100	−3.2242	0.2242	0.6458	−5.3713	0.3713	1.1550	−9.9359	0.0641	1.6897	−14.3302	0.6698	1.4913
		500	−3.0425	0.0425	0.2524	−5.0627	0.0627	0.4578	−9.6485	0.3515	0.8098	−14.1719	0.8281	1.0850