HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 45 Issue 4
2016

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 45 Issue 4
August 2016

A Peer Reviewed Journal
Published Bimonthly by the
Faculty of Science of Hacettepe University

Abstracted/Indexed in

SCI-EXP, Journal Citation Reports, Mathematical Reviews,
Zentralblatt MATH, Current Index to Statistics,
Statistical Theory & Method Abstracts,
SCOPUS, Tübitak-Ulakbim.

ISSN 1303 5010

This Journal is typeset using LaTeX.

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Cilt 45 Sayı 4 (2016)

ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2016 SAYI : 45 - 4 AY : Ağustos
YAYIN SAHİBİNİN ADI : H. Ü. Fen Fakültesi Dekanlığı adına Prof. Dr. A. Nurdan Özer
SORUMLU YAZI İŞL. MD. ADI : Prof. Dr. Yücel Tıraş
YAYIN İDARE MERKEZİ ADRESİ : H. Ü. Fen Fakültesi Dekanlığı
YAYIN İDARE MERKEZİ TEL. : 0 312 297 68 50
YAYININ TÜRÜ : Yaygın
BASIMCININ ADI : Hacettepe Üniversitesi Hastaneleri Basımevi.
BASIMCININ ADRESİ : 06100 Sıhhıye, ANKARA.
BASIMCININ TEL. : 0 312 305 1020
BASIM TARİHİ - YERİ : - ANKARA

Hacettepe Journal of Mathematics and Statistics

A Bimonthly Publication – Volume 45 Issue 4 (2016)

ISSN 1303 – 5010

Published by Hacettepe University
Faculty of Science

CONTENTS

Mathematics

Statistics

MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 999 – 1005

Investigation of spectral analysis of matrix quantum difference equations with spectral singularities

Yelda Aygar*†

Abstract

In this paper, we investigate the Jost solution, the continuous spectrum, the eigenvalues and the spectral singularities of a nonselfadjoint matrix-valued q-difference equation of second order with spectral singularities.

Keywords: Quantum difference equation, Discrete spectrum, Spectral theory, Spectral singularity, Eigenvalue.

2000 AMS Classification: 39A05, 39A70, 47A05, 47A10, 47A55.

Received : 17.02.2015 *Accepted* : 02.07.2015 *Doi* : 10.15672/HJMS.20164513107

1. Introduction

Spectral analysis of nonselfadjoint differential equations including Sturm–Liouville, Schrödinger and Klein–Gordon equations has been treated by various authors since 1960 [23, 9, 11, 22, 12]. Study of spectral theory of nonselfadjoint discrete Schrödinger and Dirac equations were obtained in [1, 20, 8, 10, 7]. Also, spectral analysis of these equations in self-adjoint case is well-known [4, 5]. In addition to differential and discrete equations, spectral theory of q-difference equations has been investigated in recent years [2, 3], and important generalizations and results were given for dynamic equations including q-difference equations as a special case in [14, 13].

Some problems of spectral theory of differential and difference equations with matrix coefficients were studied in [15, 24, 18, 6]. But spectral analysis of the matrix q-difference equations with spectral singularities has not been investigated yet.

In this paper, we let $q > 1$ and use the notation $q^{\mathbb{N}_0} := \{q^n : n \in \mathbb{N}_0\}$, where $\mathbb{N}_0$ denotes the set of nonnegative integers. Let us introduce the Hilbert space $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$ consisting of all vector sequences $y \in \mathbb{C}^m$, $(y = y(t), \quad t \in q^{\mathbb{N}})$, such that $\sum_{t\in q^{\mathbb{N}}} \mu(t)\|y(t)\|^2_{\mathbb{C}^m} < \infty$ with the inner product $\langle y, z\rangle_q := \sum_{t\in q^{\mathbb{N}}} \mu(t)\,(y(t), z(t))_{\mathbb{C}^m}$, where $\mathbb{C}^m$ is m-dimensional $(m < \infty)$ Euclidean space, $\mu(t) = (q-1)t$ for all $t \in q^{\mathbb{N}}$, and $\|\cdot\|_{\mathbb{C}^m}$ and $(\cdot,\cdot)_{\mathbb{C}^m}$ denote

*University of Ankara, Faculty of Science, Department of Mathematics, 06100, Ankara, Turkey, Email: yaygar@science.ankara.edu.tr

†Corresponding Author.

the norm and inner product in $\mathbb{C}^m$, respectively. We denote by L the operator generated in $\ell_2(q^{\mathbb{N}},\mathbb{C}^m)$ by the q-difference expression

$$(ly)(t) := qA(t)y(qt) + B(t)y(t) + A\left(\frac{t}{q}\right)y\left(\frac{t}{q}\right), \quad t \in q^{\mathbb{N}},$$

and the boundary condition $y(1) = 0$, where $A(t)$, $t \in q^{\mathbb{N}_0}$ and $B(t)$, $t \in q^{\mathbb{N}}$ are linear operators (matrices) acting in $\mathbb{C}^m$. Throughout the paper, we will assume that $A(t)$ is invertible and $A(t) \neq A^*(t)$ for all $t \in q^{\mathbb{N}_0}$. Furthermore $B(t) \neq B^*(t)$ for all $t \in q^{\mathbb{N}}$, where $*$ denotes the adjoint operator. It is clear that L is a nonselfadjoint operator in $\ell_2(q^{\mathbb{N}},\mathbb{C}^m)$. Related to the operator L, we will consider the matrix q-difference equation of second order

$$(1.1) \qquad qA(t)y(qt) + B(t)y(t) + A\left(\frac{t}{q}\right)y\left(\frac{t}{q}\right) = \lambda y(t), \quad t \in q^{\mathbb{N}},$$

where λ is a spectral parameter.

The set up of this paper is summarized as follows: Section 2 discusses the Jost solution of (1.1) and contains analytical properties and asymptotic behavior of this solution. In Section 3, we give the continuous spectrum of L, by using the Weyl compact perturbation theorem. In Section 4, we investigate the eigenvalues and the spectral singularities of L. In particular, we prove that L has a finite number of eigenvalues and spectral singularities with a finite multiplicity.

2. Jost solution of L

We assume that the matrix sequences $\{A(t)\}$ and $\{B(t)\}$, $t \in q^{\mathbb{N}}$ satisfy

$$(2.1) \qquad \sum_{t\in q^{\mathbb{N}}} (\|I - A(t)\| + \|B(t)\|) < \infty,$$

where $\|\cdot\|$ denotes the matrix norm in $\mathbb{C}^m$ and I is identity matrix. Let $F(\cdot, z)$, denotes the matrix solution of the q-difference equation

$$(2.2) \qquad qA(t)y(qt) + B(t)y(t) + A\left(\frac{t}{q}\right)y\left(\frac{t}{q}\right) = 2\sqrt{q}\cos z y(t), \quad t \in q^{\mathbb{N}},$$

satisfying the condition

$$(2.3) \qquad \lim_{t\to\infty} F(t,z)e^{i\frac{\ln t}{\ln q}z}\sqrt{\mu(t)} = I, \quad z \in \overline{\mathbb{C}}_+ := \{z \in \mathbb{C} : \operatorname{Im} z \geq 0\}.$$

The solution $F(\cdot, z)$ is called the Jost solution of (2.2).

2.1. Theorem. *Assume* (2.1). *Let the solution* $F(\cdot, z)$ *be the Jost solution of* (2.2). *Then*

$$(2.4) \qquad F(t,z) = \frac{e^{i\frac{\ln t}{\ln q}z}}{\sqrt{\mu(t)}}I + \sum_{s\in[qt,\infty)\cap q^{\mathbb{N}}} \sqrt{\frac{s}{qt}} \, \frac{\sin\left(\frac{\ln s - \ln t}{\ln q}\right)z}{\sin z} H(s),$$

where

$$H(s) := \left[I - A\left(\frac{s}{q}\right)\right]F\left(\frac{s}{q},z\right) - B(s)F(s,z) + q[I - A(s)]F(qs,z).$$

Proof. Using (2.2), we obtain

$$(2.5) \qquad F\left(\frac{t}{q}\right) + qF(qt) - 2\sqrt{q}\cos zF(t) = H(t).$$

Since $\frac{\exp\left(i\frac{\ln t}{\ln q}z\right)}{\sqrt{\mu(t)}}I$ and $\frac{\exp\left(-i\frac{\ln t}{\ln q}z\right)}{\sqrt{\mu(t)}}I$ are linearly independent solutions of the homogeneous equation

$$F\left(\frac{t}{q}\right)+qF(qt)-2\sqrt{q}\cos zF(t)=0,$$

we get the general solution of (2.5) by

$$\begin{aligned}F(t,z)&=\frac{e^{i\frac{\ln t}{\ln q}z}}{\sqrt{\mu(t)}}\alpha+\frac{e^{-i\frac{\ln t}{\ln q}z}}{\sqrt{\mu(t)}}\beta\\&+\sum_{s\in[qt,\infty)\cap q^{\mathbb{N}}}\sqrt{\frac{\mu(s)}{q}}\frac{1}{\sqrt{\mu(t)}}\frac{\sin\left(\frac{\ln s-\ln t}{\ln q}\right)z}{\sin z}H(s),\end{aligned}\tag{2.6}$$

where α and β are constants in $\mathbb{C}^m$. Using (2.1), (2.3), and (2.6), we find $\alpha=I$ and $\beta=0$. This completes the proof, i.e., $F(t,z)$ satisfies (2.4). □

2.2. Theorem. *Assume* (2.1). *Then the Jost solution $F(\cdot,z)$ has a representation*

$$F(t,z)=T(t)\frac{e^{i\frac{\ln t}{\ln q}z}}{\sqrt{\mu(t)}}\left(I+\sum_{r\in q^{\mathbb{N}}}K(t,r)e^{i\frac{\ln r}{\ln q}z}\right),\quad t\in q^{\mathbb{N}_0}\tag{2.7}$$

where $z\in\overline{\mathbb{C}}_+$, $T(t)$ and $K(t,r)$ are expressed in terms of $\{A(t)\}$ and $\{B(t)\}$.

Proof. If we put $F(\cdot,z)$ defined by (2.7) into (2.2), then we have the relations

$$\begin{aligned}&A(t)T(t)=T(qt),\quad K(t,q)-K(\frac{t}{q},q)=\frac{1}{\sqrt{q}}T^{-1}(t)B(t)T(t),\\&K(\frac{t}{q},q^2)-K(t,q^2)=T^{-1}(t)\left(T(t)-A^2(t)T(t)-\frac{1}{\sqrt{q}}B(t)T(t)K(t,q)\right),\\&K(t,rq^2)-K(\frac{t}{q},rq^2)=T^{-1}(t)\left(A^2(t)T(t)K(qt,r)+\frac{1}{\sqrt{q}}B(t)T(t)K(t,qr)\right)-K(t,r),\end{aligned}$$

and using these relations, we obtain

$$\begin{aligned}&T(t)=\prod_{p\in[t,\infty)\cap q^{\mathbb{N}}}[A(p)]^{-1},\quad K(t,q)=-\frac{1}{\sqrt{q}}\sum_{p\in[qt,\infty)\cap q^{\mathbb{N}}}T^{-1}(p)B(p)T(p),\\&K(t,q^2)=\sum_{p\in[qt,\infty)\cap q^{\mathbb{N}}}T^{-1}(p)\left[-\frac{1}{\sqrt{q}}B(p)T(p)K(p,q)+(I-A^2(p))T(p)\right],\\&K(t,rq^2)=K(qt,r)+\sum_{p\in[qt,\infty)\cap q^{\mathbb{N}}}T^{-1}(p)\left[I-A^2(p)\right]T(p)K(qp,r)\\&-\frac{1}{\sqrt{q}}\sum_{p\in[qt,\infty)\cap q^{\mathbb{N}}}T^{-1}(p)B(p)T(p)K(p,qr),\end{aligned}$$

for $r\in q^{\mathbb{N}}$ and $t\in q^{\mathbb{N}_0}$. Due to the condition (2.1), the infinite product and the series in the definition of $T(t)$ and $K(t,r)$ are absolutely convergent. □

Note that, in analogy to the Sturm–Liouville equation the function $F(1,z):=\frac{T(1)}{\sqrt{\mu(1)}}\left(I+\sum_{r\in q^{\mathbb{N}}}K(1,r)e^{i\frac{\ln r}{\ln q}z}\right)$ is called the Jost function.

2.3. Theorem. *Assume*

$$\sum_{t\in q^{\mathbb{N}}}\frac{\ln t}{\ln q}\left(\|I-A(t)\|+\|B(t)\|\right)<\infty.\tag{2.8}$$

Then the Jost solution $F(\cdot, z)$ is continuous in $\overline{\mathbb{C}}_+$ and analytic with respect to z in $\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$.

Proof. Using the equalities for $K(t,r)$ given in Theorem 2.2 and mathematical induction, we get

$$\|K(t,r)\| \leq C \sum_{p \in \left[tq^{\lfloor \frac{\ln r}{2\ln q} \rfloor}, \infty\right) \cap q^{\mathbb{N}}} \left(\|I - A(p)\| + \|B(p)\|\right), \tag{2.9}$$

where $C > 0$ is a constant and $\lfloor \frac{\ln r}{2\ln q} \rfloor$ is the integer part of $\frac{\ln r}{2\ln q}$. From (2.8) and (2.9), we get that the series

$$\sum_{r \in q^{\mathbb{N}}} K(t,r) e^{i\frac{\ln r}{\ln q} z}, \quad \sum_{r \in q^{\mathbb{N}}} \frac{\ln r}{\ln q} K(t,r) e^{i\frac{\ln r}{\ln q} z}$$

are absolutely convergent in $\overline{\mathbb{C}}_+$ and in $\mathbb{C}_+$, respectively. This completes the proof. □

2.4. Theorem. *Under the condition* (2.8)*, the Jost solution satisfies*

$$F(t,z) = \frac{e^{i\frac{\ln t}{\ln q} z}}{\sqrt{\mu(t)}} \left(I + o(1)\right), \ z \in \overline{\mathbb{C}}_+, \ t \to \infty, \tag{2.10}$$

$$F(t,z) = T(t) \frac{e^{i\frac{\ln t}{\ln q} z}}{\sqrt{\mu(t)}} \left(I + o(1)\right), \ t \in q^{\mathbb{N}_0}, \ \operatorname{Im} z \to \infty. \tag{2.11}$$

Proof. It follows from the definition of $T(t)$, (2.8), and (2.9) that

$$\lim_{t \to \infty} T(t) = I, \tag{2.12}$$

and

$$\sum_{r \in q^{\mathbb{N}}} K(t,r) e^{i\frac{\ln r}{\ln q} z} = o(1), \ z \in \overline{\mathbb{C}}_+, \ t \to \infty. \tag{2.13}$$

From (2.7), (2.12), and (2.13), we get (2.10). Using (2.8) and (2.9), we have

$$\sum_{r \in q^{\mathbb{N}}} K(t,r) e^{i\frac{\ln r}{\ln q} z} = o(1), \ z \in \overline{\mathbb{C}}_+, \ \operatorname{Im} z \to \infty. \tag{2.14}$$

From (2.7) and (2.14), we get (2.11). □

3. Continuous spectrum of L

Let L_1 and L_2 denote the q-difference operators generated in $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$ by the q-difference expressions

$$(l_1 y)(t) = q y(qt) + y\left(\frac{t}{q}\right)$$

and

$$(l_2 y)(t) = q\left[A(t) - I\right] y(qt) + B(t) y(t) + \left[A\left(\frac{t}{q}\right) - I\right] y\left(\frac{t}{q}\right)$$

with the boundary condition $y(1) = 0$, respectively. It is clear that $L = L_1 + L_2$.

3.1. Lemma. *The operator L_1 is self-adjoint in $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$.*

Proof. Since

$$\|L_1 y\|_q \leq 2\sqrt{q}\|y\|_q$$

for all $y \in \ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$, L_1 is bounded in the Hilbert space $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$, and since

$$\begin{aligned} \langle l_1 y, z\rangle_q &= \sum_{t\in q^{\mathbb{N}}} \mu(t)(z(t))^* \left(qy(qt) + y\left(\frac{t}{q}\right)\right) \\ &= \sum_{t\in q^{\mathbb{N}}} \mu(t)\left(qz(qt) + z\left(\frac{t}{q}\right)\right)^* y(t) = \langle y, l_1 z\rangle_q, \end{aligned}$$

the operator L_1 is self-adjoint in $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$. □

3.2. Theorem. *Assume* (2.8). *Then* $\sigma_c(L) = [-2\sqrt{q}, 2\sqrt{q}]$, *where* $\sigma_c(L)$ *denotes the continuous spectrum of* L.

Proof. It is easy to see that L_1 has no eigenvalues, so the spectrum of the operator L_1 consists only its continuous spectrum and

$$\sigma(L_1) = \sigma_c(L_1) = [-2\sqrt{q}, 2\sqrt{q}],$$

where $\sigma(L_1)$ denotes the spectrum of the operator L_1. Using (2.8), we find that L_2 is compact operator in $\ell_2(q^{\mathbb{N}}, \mathbb{C}^m)$[21]. Since $L = L_1 + L_2$ and $L_1 = (L_1)^*$, we obtain that

$$\sigma_c(L) = \sigma_c(L_1) = [-2\sqrt{q}, 2\sqrt{q}]$$

by using Weyl's theorem of a compact perturbation [19, p.13]. □

4. Eigenvalues and spectral singularities of L

If we define

$$f(z) := detF(1, z),\ z \in \overline{\mathbb{C}}_+, \tag{4.1}$$

then the function f is analytic in $\mathbb{C}_+$, $f(z) = f(z + 2\pi)$ and is continuous in $\overline{\mathbb{C}}_+$. Let us define the semi-strips $P_0 = \{z \in \mathbb{C}_+ : 0 \leq \operatorname{Re} z \leq 2\pi\}$ and $P = P_0 \cup [0, 2\pi]$. We will denote the set of all eigenvalues and spectral singularities of L by $\sigma_d(L)$ and $\sigma_{ss}(L)$, respectively. From the definitions of eigenvalues and spectral singularities of nonselfadjoint operators[22, 23], we have

$$\sigma_d(L) = \{\lambda \in \mathbb{C} :\ \lambda = 2\sqrt{q}\cos z,\ z \in P_0,\ f(z) = 0\}, \tag{4.2}$$

$$\sigma_{ss}(L) = \{\lambda \in \mathbb{C} :\ \lambda = 2\sqrt{q}\cos z,\ z \in [0, 2\pi],\ f(z) = 0\} \backslash \{0\}. \tag{4.3}$$

4.1. Theorem. *Assume* (2.8). *Then*

i) *the set* $\sigma_d(L)$ *is bounded and countable, and its limit points lie only in the interval* $[-2\sqrt{q}, 2\sqrt{q}]$,
ii) $\sigma_{ss}(L) \subset [-2\sqrt{q}, 2\sqrt{q}]$ *and the linear Lebesgue measure of the set* $\sigma_{ss}(L)$ *is zero.*

Proof. In order to investigate the quantitative properties of eigenvalues and spectral singularities of L, it is necessary to discuss the quantitative properties of zeros of f in P from (4.2) and (4.3). Using (2.11) and (4.1), we get

$$f(z) = \det T(1)\frac{1}{\mu(1)}[I + o(1)],\ \operatorname{Im} z > 0,\ z \in P_0,\ \operatorname{Im} z \to \infty, \tag{4.4}$$

where $detT(1) \neq 0$. From (4.4), we get the boundedness of zeros of f in P_0. Since f is a 2π-periodic function and is analytic in $\mathbb{C}_+$, we obtain that f has at most a countable number of zeros in P_0. By the uniqueness of analytic functions, we find that the the limit points of zeros of f in P_0 can lie only in $[0, 2\pi]$. We get $\sigma_{ss}(L) \subset [-2\sqrt{q}, 2\sqrt{q}]$ using

(4.3). Since $f(z) \neq 0$ for all $z \in \mathbb{C}_+$, we get that the linear Lebesgue measure of the set of zeros of f on real axis is not positive, by using the boundary uniqueness theorem of analytic functions [17], i.e., the linear Lebesgue measure of the $\sigma_{ss}(L)$ is zero. □

4.2. Definition. The multiplicity of a zero of f in P is called the multiplicity of the corresponding eigenvalue or spectral singularity of L.

4.3. Theorem. *If, for some $\varepsilon > 0$,*

$$\sup_{t\in q^{\mathbb{N}}} \left\{ e^{\varepsilon \frac{\ln t}{\ln q}} (\|I - A(t)\| + \|B(t)\|) \right\} < \infty \tag{4.5}$$

then the operator L has a finite number of eigenvalues and spectral singularities, and each of them is of finite multiplicity.

Proof. Since $F(1,z) = \frac{T(1)}{\sqrt{q-1}} \left(I + \sum_{r\in q^{\mathbb{N}}} K(1,r) e^{i \frac{\ln r}{\ln q} z} \right)$, using (2.9) and (4.5), we get that

$$\|K(1,r)\| \leq D e^{-\frac{\varepsilon}{4} \frac{\ln r}{\ln q}}, \; r \in q^{\mathbb{N}}, \tag{4.6}$$

where $D > 0$ is a constant. From (4.1) and (4.6), we obtain that the function f has an analytic continuation to the half-plane $\operatorname{Im} z > -\frac{\varepsilon}{4}$. Because the series

$$\sum_{r\in q^{\mathbb{N}}} iK(1,r) \frac{\ln r}{\ln q} e^{i \frac{\ln r}{\ln q} z}$$

is uniformly convergent in $\operatorname{Im} z > -\frac{\varepsilon}{4}$. Since f is a 2π periodic function, the limit points of its zeros in P cannot lie in $[0, 2\pi]$. Using Theorem 4.1, we find that the bounded sets $\sigma_d(L)$ and $\sigma_{ss}(L)$ have no limit points, i.e., the sets $\sigma_d(L)$ and $\sigma_{ss}(L)$ have a finite number of elements. From the analyticity of f in $\operatorname{Im} z > -\frac{\varepsilon}{4}$, we get that all zeros of f in P have a finite multiplicity. □

References

[1] Adıvar, M. and Bairamov, E. *Spectral properties of non-selfadjoint difference operators*, J. Math. Anal. Appl. **261** (2), 461-478, 2001.

[2] Adıvar, M. and Bohner, M. *Spectral analysis of q-difference equations with spectral singularities*, Math. Comput. Modelling **43** (7), 695–703, 2006.

[3] Adıvar, M. and Bohner, M. *Spectrum and principal vectors of second order q-difference equations*, Indian J. Math. **48** (1), 17–33, 2006.

[4] Agarwal, R.P. *Difference equations and inequalities*, Monographs and Textbooks in Pure and Applied Mathematics (Marcel Dekker Inc., New York, 2000).

[5] Agarwal, R.P. and Wong, P.J.Y. *Advanced topics in difference equations* (Kluwer Academic Publishers Group, Dordrecht, 1997).

[6] Aygar, Y. and Bairamov, B. *Jost solution and the spectral properties of the matrix-valued difference operators*, Appl. Math. Comput. **218** (19), 9676–9681, 2012.

[7] Bairamov, E. Aygar, Y. and Koprubasi, T. *The spectrum of eigenparameter-dependent discrete Sturm-Liouville equations*, J. Comput. Appl. Math. **235**, 4519–4523, 2011.

[8] Bairamov, E. and Coskun, C. *Jost solutions and the spectrum of the system of difference equations*, Appl. Math. Lett. (2) **17** (9), 1039–1045, 2004.

[9] Bairamov, E. Çakar, Ö. and Çelebi, A.O. *Quadratic pencil of Schrödinger operators with spectral singularities: discrete spectrum and principal functions*, J. Math. Anal. Appl. **216** (1), 303–320, 1997.

[10] Bairamov, E. Çakar, Ö. and Krall, A.M. *Non-selfadjoint difference operators and Jacobi matrices with spectral singularities*, Math. Nachr. **229**, 5–14, 2001.

[11] Başcanbaz Tunca, G. *Spectral properties of the Klein-Gordon s-wave equation with spectrAL parameter-dependent boundary condition*, Int. J. Math. Math. Sci. (25-28), 1437–1445, 2004.

[12] Başcanbaz Tunca, G. and Bairamov, E. *Discrete spectrum and principal functions of non-selfadjoint differential operator*, Czechoslovak Math. J. **49** (4), 689–700, 1999.

[13] Bohner, M. Guseinov, G. and Peterson, A. *Introduction to the time scales calculus. In Advances in dynamic equations on time scales* (Birkhäuser Boston, MA, 2003), 1–15.

[14] Bohner, M. and Peterson, A. *Dynamic equations on time scales* (Birkhäuser Boston, MA, 2001).

[15] Clark, S. Gesztesy, F. and Renger, W. *Trace formulas and Borg-type theorems for matrix-valued Jacobi and Dirac finite difference operators*, J. Differential Equations **219** (1), 144–182, 2005.

[16] Coskun, C. and Olgun, M. *Principal functions of non-selfadjoint matrix Sturm-Liouville equations*, J. Comput. Appl. Math. **235** (16), 4834–4838, 2011.

[17] Dolzhenko, E.P., *Boundary-value uniqueness theorems for analytic functions*, Mathematical notes of the Academy of Sciences of the USSR **25**, (6), 437–442, 1979.

[18] Gesztesy, F. Kiselev, A. and Makarov, K.A. *Uniqueness results for matrix-valued Schrödinger, Jacobi, and Dirac-type operators*, Math. Nachr. **239/240**, 103–145, 2002.

[19] Glazman, I.M. *Direct methods of qualitative spectral analysis of singular differential operators* (Jerusalem, 1966).

[20] Krall, A.M. Bairamov, E. and Çakar, Ö. *Spectral analysis of non-selfadjoint discrete Schrödinger operators with spectral singularities*, Math. Nachr. **231**, 89–104, 2001.

[21] L.A. Lusternik and V.J. Sobolev, *Elements of functional analysis* (Hindustan Publishing Corp., Delhi, 1974).

[22] Lyance, V.E. *A differential operator with spectral singularities I-II*, Amer. Math. Soc. Transl. **60** (2) 185–225, 227–283, 1967.

[23] Naimark, M.A. *Linear differential operators. Part II: Linear differential operators in Hilbert space* (Frederick Ungar Publishing Co., New York, 1960).

[24] Olgun, M. and Coskun, C. *Non-selfadjoint matrix Sturm-Liouville operators with spectral singularities*, Appl. Math. Comput. **216**, (8), 2271–2275, 2011.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1007 – 1014

Spectral singularities of the matrix Schrödinger equations

Elgiz Bairamov* and Şerifenur Cebesoy†

Abstract

In this paper, we investigate analytical and asymptotical properties of the Jost function of the matrix Schrödinger equation. Later, using the analytic continuation and the uniqueness theorems of analytic functions we study the eigenvalues and the spectral singularities of that equation.

Keywords: Differential Equations, Jost Function, Eigenvalues, Spectral Singularities

2000 AMS Classification: 34B05, 34B27, 34L05

Received : 19.06.2015 *Accepted* : 05.10.2015 *Doi* : 10.15672/HJMS.20164514275

1. Introduction

Schrödinger differential equations subject to the general point interaction can be found in many real world problems. These equations describe observed evolution phenomena. For instance, many chemical, physical phenomena and pharmacokinetics do exhibit point interaction effects [1]. The spectral analysis of Schrödinger equations with general point interaction have been investigated in detail in [2]-[6]. To be more precise, we should note that these equations have bound states, i.e., eigenvalues with square-integrable eigenfunctions and spectral singularities. It is well known that the bound state of quantum mechanical system correspond to the energy. Also a physical interpretation for the spectral singularities that identifies with the energies of scattering states having infinite reflection and transmission
coefficients. So spectral singularities correspond to the resonance states having a real energy. Consequently, the spectral analysis of Schrödinger equations with
spectral singularities are important to study in quantum mechanics. So in this paper, we investigate the spectral analysis of general matrix Schrödinger equations with spectral singularities.

*Department of Mathematics, Ankara University, 06100 Tandoğan, Ankara, TURKEY
Email : bairamov@science.ankara.edu.tr
†Department of Mathematics, Ankara University, 06100 Tandoğan, Ankara, TURKEY
Email : scebesoy@ankara.edu.tr

The spectral analysis of differential equations with spectral singularities was investigated by Naimark [7]. Schwartz studied the spectral singularities of a certain class of abstract linear operators in a Hilbert space [8]. The following definition of spectral singularities is given by Schwartz.

Let H be a Hilbert space and $A : H \to H$ be a linear operator such that its spectrum $\sigma(A)$ consists of an interval J of the real axis and a finite number of complex numbers outside J. We will denote the resolvent operator of A by $R_\mu(A) := (A - \mu I)^{-1}$, where $\mu \in \mathbb{C}$. Let J_0 be a finite subset of J. Assume that for any finite subinterval Δ of J, whose closure do not contain any point of J_0, the limit operator

$$E_\Delta = \lim_{\epsilon \to 0^+} \frac{1}{2\pi i} \int\limits_{\Delta} \left[R_{\mu+i\epsilon}(A) - R_{\mu-i\epsilon}(A)\right] d\mu$$

exists in the strong limit sense, so that E_Δ is a linear bounded operator on H. Denote by d the distance from the interval Δ to the set J_0. If

$$\lim_{d \to 0} \parallel E_\Delta \parallel = \infty$$

then any point of the set J_0, is called a spectral singularity of the operator A. For the selfadjoint operators $\parallel E_\Delta \parallel \leq 1$, so that selfadjoint operators have no spectral singularities.

The sets of the spectral singularities for closed linear operators on a Banach space was given by Nagy [9]. Nagy shows that the set of spectral singularities defined according to his general definition coincides in the case of differential operators as defined by Naimark and Lyance [7], [10]. Pavlov established the dependence of the structure of the spectral singularities of Schrödinger operators on the behaviour of the potential function at infinity [11].

Note that the principal functions corresponding to the spectral singularities are not the elements of the Hilbert space. Also, the spectral singularities belong to the continuous spectrum and are not the eigenvalues. However, the spectral singularities play a certain critical role in the spectral analysis of operators. Their existence is accompanied by specific phenomenon which are new in the sense that they do not occur either in the spectral analysis of selfadjoint or normal operators.

The spectral singularities of the Sturm-Liouville operators with the general boundary conditions was investigated in detail by Krall [12], [13]. Some problems of spectral theory of differential equations and operators with spectral singularities were also studied in [14]-[19].

Let S be a n-dimensional $(n < \infty)$ Euclidian space and we denote by $L^2(\mathbb{R}, S)$ the Hilbert space of vector-valued functions with values in S and the norm

$$\|f\|^2 := \int\limits_{-\infty}^{\infty} \|f(x)\|_S^2 \, dx.$$

Let L denote the operator generated in $L^2(\mathbb{R}, S)$ by the matrix Schrödinger equation

$$-y'' + Q(x)y = \lambda^2 y \quad , -\infty < x < \infty \tag{1.1}$$

where Q is a non-self adjoint matrix-valued potential function $(Q \neq Q^*)$ and λ is a spectral parameter. It is clear that, the operator L is non-selfadjoint. L is called the matrix Schrödinger operator.

In this paper, we investigate asymptotics and analytical properties of Jost function of (1.1). We also obtain the resolvent of L. Later, we study the eigenvalues and the spectral singularities of L using the analytic continuation and uniqueness theorems of analytic functions. Afterwards we prove that the equation (1.1) (i.e. the operator L) has a finite number of eigenvalues and spectral singularities and each of them is of finite multiplicity under the condition

$$\sup_{-\infty<x<\infty}\left\{\exp\left(\epsilon\sqrt{|x|}\right)\|Q(x)\|\right\}<\infty \quad , \ \epsilon>0.$$

2. Jost function

Suppose the matrix function Q satisfies

$$\int_{-\infty}^{\infty}(1+|x|)\,\|Q(x)\|\,dx<\infty. \tag{2.1}$$

We introduce the notation

$$\eta^{+}(x)=\int_{x}^{\infty}\|Q(t)\|\,dt\ , \qquad \eta_1^{+}(x)=\int_{x}^{\infty}\eta^{+}(t)dt,$$

$$\eta^{-}(x)=\int_{-\infty}^{x}\|Q(t)\|\,dt\ , \qquad \eta_1^{-}(x)=\int_{-\infty}^{x}\eta^{-}(t)dt.$$

Let $E^{+}(x,\lambda)$ and $F^{-}(x,\lambda)$ denote the solutions of (1.1) subject to the conditions

$$\lim_{x\to\infty}E^{+}(x,\lambda)\,e^{-i\lambda x}=I, \qquad \lambda\in\overline{\mathbb{C}}_{+}:=\{\lambda:\lambda\in\mathbb{C},\ \operatorname{Im}\lambda\geq 0\},$$

and

$$\lim_{x\to-\infty}F^{-}(x,\lambda)\,e^{i\lambda x}=I, \qquad \lambda\in\overline{\mathbb{C}}_{+},$$

respectively, where I denotes the identity matrix in S. Under the condition (2.1) the solution $E^{+}(x,\lambda)$ has the following integral representation [20]

$$E^{+}(x,\lambda)=e^{i\lambda x}I+\int_{x}^{\infty}K^{+}(x,t)\,e^{i\lambda t}dt, \qquad \lambda\in\overline{\mathbb{C}}_{+}.$$

We also denote the solution of the equation

$$-z^{''}+zQ(x)=\lambda^2 z \qquad , \ -\infty<x<\infty,$$

subject to the condition

$$\lim_{x\to-\infty}z^{-}(x,\lambda)\,e^{i\lambda x}=I, \qquad \lambda\in\overline{\mathbb{C}}_{+}$$

by $E^{-}(x,\lambda)$.

Under the conditon (2.1), the solution $E^{-}(x,\lambda)$ has the similar integral representation

$$E^{-}(x,\lambda)=e^{-i\lambda x}I+\int_{-\infty}^{x}K^{-}(x,t)\,e^{-i\lambda t}dt, \qquad \lambda\in\overline{\mathbb{C}}_{+}$$

where the matrix-functions $K^{\pm}(x,t)$ are differentiable with respect to x and t and satisfies

$$\left\|K^{\pm}(x,t)\right\|\leq\frac{1}{2}\eta^{\pm}\left(\frac{x+t}{2}\right)\exp\left\{\eta_1^{\pm}(x)-\eta_1^{\pm}\left(\frac{x+t}{2}\right)\right\}, \tag{2.2}$$

$$\left\|K_x^{\pm}(x,t)\mp\frac{1}{4}Q\left(\frac{x+t}{2}\right)\right\|\leq\frac{1}{2}\eta_1^{\pm}(x)\eta^{\pm}\left(\frac{x+t}{2}\right)\exp\eta_1^{\pm}(x), \tag{2.3}$$

$$\left\| K_t^{\pm}(x,t) \mp \frac{1}{4} Q\left(\frac{x+t}{2}\right) \right\| \le \frac{1}{2} \eta_1^{\pm}(t) \eta^{\pm}\left(\frac{x+t}{2}\right) \exp \eta_1^{\pm}(t) \tag{2.4}$$

The matrix-functions $E^+(x,\lambda)$ and $E^-(x,\lambda)$ are analytic with respect to λ in $\mathbb{C}_+ := \{\lambda : \lambda \in \mathbb{C},\ \operatorname{Im}\lambda > 0\}$ and continuous up to the real axis.

Now, let us introduce

$$D(\lambda) := W\left[E^-(x,\lambda), E^+(x,\lambda)\right], \quad \lambda \in \overline{\mathbb{C}}_+,$$

where $W\left[E^-(x,\lambda), E^+(x,\lambda)\right]$ is the Wronskian of the solutions of $E^-(x,\lambda)$ and $E^+(x,\lambda)$. The function D is called Jost function of (1.1). Note that Jost function is analytic in $\mathbb{C}_+$ and continuous on the real axis.

Theorem 2.1. *The function D satisfies*

$$D(\lambda) = 2i\lambda I - 2K^+(0,0) - 2K^-(0,0) + \int_0^\infty f(t) e^{i\lambda t} dt \tag{2.5}$$

where

$$\begin{aligned} f(t) = {} & K_x^+(0,t) - K_x^-(0,-t) - K_t^+(0,t) + K_t^-(0,-t) - K^-(0,0)\, K^+(0,t) \\ & - K^+(0,0)\, K^-(0,-t) + K^-(0,-t) * K_x^+(0,t) + K_x^-(0,-t) * K^+(0,t) \end{aligned}$$

in which $(*)$ is the convolution operation and $f \in L_1(\mathbb{R}, S)$.

Proof. By the definition of the Wronskian of the solutions $E^-(x,\lambda)$ and $E^+(x,\lambda)$ we have

$$D(\lambda) = E^-(0,\lambda) E_x^+(0,\lambda) - E_x^-(0,\lambda) E^+(0,\lambda).$$

Using the integral represantations of $E^-(x,\lambda)$ and $E^+(x,\lambda)$ we obtain (2.5). It follows from (2.2) – (2.4) that $f \in L_1(\mathbb{R}, S)$. □

Theorem 2.2. *The following asymptotics hold*

$$D(\lambda) = 2i\lambda I - 2K^+(0,0) - 2K^-(0,0) + o(1) \quad , \lambda \in \overline{\mathbb{C}}_+,\ |\lambda| \to \infty, \tag{2.6}$$

$$D(\lambda) = 2i\lambda I + O(1) \quad \lambda \in \overline{\mathbb{C}}_+,\ |\lambda| \to \infty. \tag{2.7}$$

Proof. a) Let $\lambda \in \mathbb{R}$. By the Riemann-Lebesgue lemma for the Fourier transforms we get that [21]

$$\int_0^\infty f(t) e^{i\lambda t} dt = o(1) \quad , \quad \lambda \in \mathbb{R},\ |\lambda| \to \pm\infty. \tag{2.8}$$

b) Let $\lambda \in \mathbb{C}_+$. In this case, by the Lebesgue theorem we obtain that [21]

$$\int_0^\infty f(t) e^{i\lambda t} dt = o(1) \quad , \quad \lambda \in \mathbb{C}_+,\ |\lambda| \to \infty. \tag{2.9}$$

It follows from (2.8) and (2.9) that

$$\int_0^\infty f(t) e^{i\lambda t} dt = o(1) \quad , \quad \lambda \in \overline{\mathbb{C}}_+,\ |\lambda| \to \infty. \tag{2.10}$$

From (2.5) and (2.10) we have (2.6). In a similar way we may also prove (2.7). □

3. Eigenvalues and Spectral Singularities of L

Let us suppose that

(3.1) $$G(\lambda) := \det D(\lambda)$$

Also, $\sigma_d(L)$ and $\sigma_{ss}(L)$ will denote the eigenvalues and spectral singularities of L, respectively. By the definition of eigenvalues and spectral singularities of differential operators we can write [7], [10]

(3.2) $$\sigma_d(L) = \left\{z : z = \lambda^2, \lambda \in \mathbb{C}_+, G(\lambda) = 0\right\}$$

(3.3) $$\sigma_{ss}(L) = \left\{z : z = \lambda^2, \lambda \in \mathbb{R}\backslash \{0\}, G(\lambda) = 0\right\}.$$

Definition 3.1. *The multiplicity of a zero of G in $\overline{\mathbb{C}}_+$ is defined as the multiplicity of the corresponding eigenvalue and spectral singularity of L.*

In order to investigate the quantitative properties of the eigenvalues and the spectral singularities of L, we need to discuss the quantitative properties of the zeros of G in $\overline{\mathbb{C}}_+$.

Let M_1 denotes the zeros of the function G in $\mathbb{C}_+$ and M_2 denotes the zeros of the function G on the real axis. Therefore, using (3.2) and (3.3) we obtain

(3.4) $$\sigma_d(L) = \left\{z : z = \lambda^2, \lambda \in M_1\right\}$$

(3.5) $$\sigma_{ss}(L) = \left\{z : z = \lambda^2, \lambda \in M_2\right\} \setminus \{0\}$$

Lemma 3.2. *(i) The set M_1 is bounded and has at most countable number of elements and its limit points can lie only in a bounded subinterval of the real axis.*

(ii) The set M_2 is compact and its Lebesque measure is zero.

Proof. From (2.7), we can obtain

(3.6) $$\| K^{\pm}(x,t)\| \leq c\eta^{\pm}(\frac{x+t}{2})$$

where $c > 0$ is a constant. Using (3.1) and (3.6), we get that the function G is analytic in $\mathbb{C}_+$, continuous on the real axis and satisfies the following

(3.7) $$G(\lambda) = 2i\lambda + O(1) \quad , \quad \lambda \in \overline{\mathbb{C}}_+, |\lambda| \to \infty$$

Equation (3.7) shows that the sets M_1 and M_2 are bounded. Since $D(\lambda)$ is analytic in $\mathbb{C}_+$, then the set M_1 has at most countable number of elements. By (3.7) and the boundary value uniqueness theorem of analytic functions, we get that the set M_2 is closed and $\mu(M_2) = 0$, where $\mu(M_2)$ denote Lebesgue measure of the set M_2 [22]. $\square$

From (3.4), (3.5) and *Lemma* 3.2 we obtain the following theorem.

Theorem 3.3. *Under the condition* (2.1)

(i) The set of eigenvalues of L is bounded, is no more than countable and its limit points can lie only in a bounded subinterval of the positive semiaxis.

(ii) The set of spectral singularities of L is bounded and $\mu(\sigma_{ss}(L)) = 0$.

Now, let us assume that, for some $\epsilon > 0$,

(3.8) $$\int_{-\infty}^{\infty} \exp(\epsilon\,|x|)\|Q(x)\|dx < \infty.$$

From (2.2) – (2.4), we get the following

(3.9) $$\|K^{\pm}(x,t)\|, \|K_x^{\pm}(x,t)\|, \|K_t^{\pm}(x,t)\| \leq c\exp\left[-\epsilon(\frac{x+t}{2})\right]$$

where $c > 0$ is a constant. Also under the condition (3.8) we have

$$\|f(t)\| \leq ce^{-\frac{\epsilon}{2}|t|} \tag{3.10}$$

by (3.9).

Theorem 3.4. *Under the condition* (3.8), *the operator L has a finite number of eigenvalues and spectral singularities and each of them is of finite multiplicity.*

Proof. By using (3.9) and (3.10) we observe that the function G has analytic continuation to the half plane $\operatorname{Im}\lambda > -\frac{\epsilon}{2}$. So, the limit points of zeros of G can not lie in $\mathbb{R}$. Using Lemma 3.2, we get that the bounded sets M_1 and M_2 have a finite number of elements. Since G is analytic for $\operatorname{Im}\lambda > -\frac{\epsilon}{2}$, we obtain that all zeros of G in $\overline{\mathbb{C}}_+$ have a finite multiplicity. So that the sets $\sigma_d(L)$ and $\sigma_{ss}(L)$ have a finite number of element with a finite multiplicity. □

Now, let us suppose that

$$\int_{-\infty}^{\infty} \exp(\epsilon\sqrt{|\,x\,|})\|Q(x)\|dx < \infty \quad , \quad \epsilon > 0 \tag{3.11}$$

which is weaker than (3.8).

It is evident that under the condition (3.11), the function G is analytic in $\mathbb{C}_+$and infinite differentiable on the real axis.

Let us denote the sets of all limit points of M_1 and M_2 by M_3 and M_4 respectively, and the set of all zeros of G with infinite multiplicity in $\overline{\mathbb{C}}_+$ by M_5. It is obvious from the uniqueness theorem of the analytic functions that

$$M_3 \subset M_2,\ M_4 \subset M_2,\ M_5 \subset M_2,\ M_3 \subset M_5,\ M_4 \subset M_5$$

and $\mu(M_3) = \mu(M_4) = 0$ [22].

Lemma 3.5. *Under the condition* (3.11), *the function G and its derivatives provide the following inequality:*

$$\left|G^{(n)}(\lambda)\right| \leq A_n \quad , \quad n = 1, 2, ..., \quad \operatorname{Im}\lambda > 0 \tag{3.12}$$

where

$$A_1 = 2 + c2^2 \int_0^{\infty} te^{-\frac{\epsilon}{2}\sqrt{t}}dt$$

$$A_n = c2^{n+1} \int_0^{\infty} t^n e^{-\frac{\epsilon}{2}\sqrt{t}}dt, \quad n = 2, 3, ... \tag{3.13}$$

are constants. In addition, for all $n \in \mathbb{N}$

$$A_n \leq Bb^n n! n^n \tag{3.14}$$

holds where B, b are constants.

Proof. We easily get the proof of the Lemma using (2.5) and (3.1). □

Theorem 3.6. *If* (3.11) *holds, then $M_5 = \varnothing$.*

Proof. Using Lemma 3.2, for sufficiently large $T > 0$, we get

$$\left|\int_{-\infty}^{T} \frac{\ln|G(\lambda)|}{1+\lambda^2}d\lambda\right| < \infty, \qquad \left|\int_{T}^{\infty} \frac{\ln|G(\lambda)|}{1+\lambda^2}d\lambda\right| < \infty \tag{3.14}$$

Since $G(\lambda) \neq 0$, we obtain

$$\int_0^h \ln H(s) d\mu(M_5, s) > -\infty \tag{3.15}$$

by (3.12), (3.14) and Pavlov's Theorem, where $H(s) = \inf_n \frac{A_n s^n}{n!}$, $\mu(M_5, s)$ is the Lebesque measure of the s-neighbourhood of M_5 and $h > 0$ is a constant [5]. Substituting (3.14) into the definition $H(s)$, we arrive at

$$H(s) \leq B \inf_n \{b^n s^n n^n\} \leq B \exp\left\{-b^{-1} s^{-1} e^{-1}\right\}$$

and

$$\ln H(s) \leq -b^{-1} s^{-1} e^{-1}$$

Consequently,

$$\int_0^h \frac{1}{s} d\mu(M_5, s) < \infty$$

holds by using (3.15) for arbitrary s, if and only if $\mu(M_5, s) = 0$ or $M_5 = \varnothing$. □

Lemma 3.7. *G has a finite number of zeros with finite multiplicity in $\overline{\mathbb{C}}_+$.*

Proof. Since $M_3 \subset M_5$ and $M_4 \subset M_5$, we get

$$M_3 = M_4 = \varnothing. \tag{3.16}$$

By using Lemma 3.2 and (3.16), we obtain the finiteness of the sets M_1 and M_2. Because of $M_5 = \varnothing$, all of the zeros of the function G have finite multiplicities. □

From Lemma 3.7, we get the following theorem.

Theorem 3.8. *The operator L has a finite number of eigenvalues and spectral singularities and each of them is of finite multiplicity if condition* (3.11) *holds.*

References

[1] V. Lakshmikantham, D. D. Bainov, P. S. Simeonov, Theory of Impulsive Differential Equations, Series in Modern Applied Mathematics, vol. 6 (World Scientific Publishing Co, Inc., Teaneck, NJ, 1989).

[2] A. Mostafazadeh, H. Mehri-Dehnavi, Spectral singularities, biorthonotmal systems and a two-parameter family of complex point interactions, J. Phys. A. Math. Theory, 42 (2009), 125303 (27 pp).

[3] A. Mostafazadeh, Spectral singularities of a general point interaction, J. Phys. A. Math. Theory, 44 (2011), 375302 (9 pp).

[4] A. Mostafazadeh, Optical spectral singularities as threshold resonances, Physical Review A, 83, 045801 (2011).

[5] A. Mostafazadeh, Resonance phenomenon related to spectral singularities, complex barrier potential, and resonating waveguides, Physical Review A, 80, 032711 (2009).

[6] A. Mostafazadeh, Spectral singularities of complex scattering potentials and infinite reflection and transmission coefficients at real energies, Physical Review Letters, 102 , 220402 (2009).

[7] M.A. Naimark, Investigation of the spectrum and the expansion in eigenfunctions of a nonselfadjoit operators of second order on a semi-axis. AMS Transl. (2), 16 (1960), 103-193.

[8] J.T. Schwartz, Some non-self adjoint operators, Comm. Pure Appl. Math., 13 (1960), 609-639.

[9] B. Nagy, Operators with spectral singularities, J. Oper. Theory, 15 (1986), 2, 307-325.

[10] V. E. Lyance, A differential operator with spectral singularities, I, II, AMS Transl. (2), 60 (1967), 185-225, 227-283.

[11] B.S. Pavlov, The non-selfadjoint Schrödinger operator, Topics in Math Phys. 1, (1967), 87-110.

[12] A.M. Krall, The adjoint of a differential operator with integral boundary conditions, Proc. Amer. Math. Soc., 16 (1965), 738-742.

[13] A.M. Krall, A non-homogenous eigenfunction expansion, Trans. Amer. Soc., 117 (1965), 352-361.

[14] A.M. Krall, E. Bairamov and O. Cakar, Spectrum and spectral singularities of a quadratic pencil of a Schrödinger operator with a general boundary condition, J. Differential Equations, 151 (1999), 2, 252–267.

[15] A.M. Krall, E. Bairamov and O. Cakar, Spectral analysis of a non-selfadjoint discrete Schrödinger operators with spectral singularities, Math. Nachr., 231 (2001), 89-104.

[16] E. Bairamov, O. Cakar and A.M. Krall, An eigenfunction expansion for a quadratic pencil of a Schrödinger operator with spectral singularities, J. Differential Equations 151 (1999), 268-289.

[17] E. Bairamov and A.M. Krall, Dissipative operators generated by Sturm-Liouville differential expression in the Weyl limit circle case J. Math. Anal. Appl., 254 (2001), 178-190.

[18] E. Bairamov, O. Cakar and C. Yanık, Spectral singularities Klein-Gordon s-wave equation, Indian J. Pure Appl. Math., 32 (2001), 851-857.

[19] E. Bairamov and O. Karaman, Spectral singularities of Klein-Gordon s-wave equations with an integral boundary condition, Acta Math. Hungar., 97 (2002), 1-2, 121–131.

[20] Z.S. Agranovich, V.A. Marchenko, The Inverse Problem of Scattering Theory, Gordon and Breach, New York, 1965.

[21] A.N. Kolmogorov and S.V. Fomin, Introductory Real Analysis, Dover Publ., New York, 1975.

[22] E.P. Dolzhenko, Boundary value uniqueness theorems for analytic functions, Math. Notes, 25 (1979), 437-442.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1015 – 1022

Coefficient bounds for certain subclasses of analytic functions of complex order

Serap Bulut*

Abstract

In this paper, we introduce and investigate two subclasses of analytic functions of complex order, which are introduced here by means of a certain nonhomogeneous Cauchy–Euler-type differential equation of order m. Several corollaries and consequences of the main results are also considered.

Keywords: Analytic functions, Differential operator, Nonhomogeneous Cauchy-Euler differential equations, Coefficient bounds, Subordination.

2000 AMS Classification: 30C45

Received : 19.12.2014 *Accepted* : 16.09.2015 *Doi* : 10.15672/HJMS.20164514273

1. Introduction, definitions and preliminaries

Let $\mathbb{R} = (-\infty, \infty)$ be the set of real numbers, $\mathbb{C}$ be the set of complex numbers,

$$\mathbb{N} := \{1, 2, 3, \ldots\} = \mathbb{N}_0 \setminus \{0\}$$

be the set of positive integers and

$$\mathbb{N}^* := \mathbb{N} \setminus \{1\} = \{2, 3, 4, \ldots\}.$$

Let $\mathcal{A}$ denote the class of functions of the form

$$f(z) = z + \sum_{i=2}^{\infty} a_i z^i \tag{1.1}$$

which are analytic in the open unit disk

$$\mathbb{U} = \{z : z \in \mathbb{C} \quad \text{and} \quad |z| < 1\}.$$

*Kocaeli University, Faculty of Aviation and Space Sciences, Arslanbey Campus, 41285 Kartepe-Kocaeli, TURKEY
Email : serap.bulut@kocaeli.edu.tr

Recently, Faisal and Darus [8] defined the following differential operator:

$$
\begin{aligned}
D^0 f(z) &= f(z), \\
D^1_\lambda(\alpha,\beta,\mu) f(z) &= \left(\frac{\alpha-\mu+\beta-\lambda}{\alpha+\beta}\right) f(z) + \left(\frac{\mu+\lambda}{\alpha+\beta}\right) z f'(z), \\
D^2_\lambda(\alpha,\beta,\mu) f(z) &= D\left(D^1_\lambda(\alpha,\beta,\mu) f(z)\right) \\
&\ \ \vdots \\
D^n_\lambda(\alpha,\beta,\mu) f(z) &= D\left(D^{n-1}_\lambda(\alpha,\beta,\mu) f(z)\right). \qquad (1.2)
\end{aligned}
$$

If f is given by (1.1), then it is easily seen from (1.2) that

$$
D^n_\lambda(\alpha,\beta,\mu) f(z) = z + \sum_{i=2}^{\infty} \left(\frac{\alpha+(\mu+\lambda)(i-1)+\beta}{\alpha+\beta}\right)^n a_i z^i \qquad (1.3)
$$

$$
(f \in \mathcal{A};\ \alpha,\beta,\mu,\lambda \geq 0;\ \alpha+\beta \neq 0;\ n \in \mathbb{N}_0).
$$

By using the operator $D^n_\lambda(\alpha,\beta,\mu)$, Faisal and Darus [8] defined a function class $\Psi(n,\alpha,\beta,\mu,\lambda,\zeta,\gamma,\xi)$ by

$$
\Re\left\{1+\frac{1}{\xi}\left(\frac{z\left[\zeta D^{n+1}_\lambda(\alpha,\beta,\mu) f(z)+(1-\zeta) D^n_\lambda(\alpha,\beta,\mu) f(z)\right]'}{\zeta D^{n+1}_\lambda(\alpha,\beta,\mu) f(z)+(1-\zeta) D^n_\lambda(\alpha,\beta,\mu) f(z)}-1\right)\right\} > \gamma,
$$

$$
(z \in \mathbb{U};\ 0 \leq \gamma < 1;\ 0 \leq \zeta \leq 1;\ \xi \in \mathbb{C}\backslash\{0\})
$$

and also investigated the subclass $\Phi(n,\alpha,\beta,\mu,\lambda,\zeta,\gamma,\xi,\tau)$ of the analytic function class $\mathcal{A}$, which consists of functions $f \in \mathcal{A}$ satisfying the following nonhomogenous Cauchy-Euler differential equation:

$$
z^2\frac{d^2w}{dz^2} + 2(1+\tau) z\frac{dw}{dz} + \tau(1+\tau) w = (1+\tau)(2+\tau) q(z)
$$

$$
(w = f(z) \in \mathcal{A};\ q \in \Psi(n,\alpha,\beta,\mu,\lambda,\zeta,\gamma,\xi);\ \tau \in (-1,\infty)).
$$

In the same paper [8], coefficient bounds for the subclasses $\Psi(n,\alpha,\beta,\mu,\lambda,\zeta,\gamma,\xi)$ and $\Phi(n,\alpha,\beta,\mu,\lambda,\zeta,\gamma,\xi,\tau)$ of analytic functions of complex order were obtained.

Making use of the differential operator $D^n_\lambda(\alpha,\beta,\mu)$, we now introduce each of the following subclasses of analytic functions.

1. Definition. Let $g : \mathbb{U} \to \mathbb{C}$ be a convex function such that

$$
g(0) = 1 \qquad \text{and} \qquad \Re\{g(z)\} > 0 \quad (z \in \mathbb{U}).
$$

We denote by $\mathcal{M}_g(n,\alpha,\beta,\mu,\lambda,\zeta,\xi)$ the class of functions $f \in \mathcal{A}$ satisfying

$$
1+\frac{1}{\xi}\left(\frac{z\left[\zeta D^{n+1}_\lambda(\alpha,\beta,\mu) f(z)+(1-\zeta) D^n_\lambda(\alpha,\beta,\mu) f(z)\right]'}{\zeta D^{n+1}_\lambda(\alpha,\beta,\mu) f(z)+(1-\zeta) D^n_\lambda(\alpha,\beta,\mu) f(z)}-1\right) \in g(\mathbb{U}),
$$

where $z \in \mathbb{U};\ 0 \leq \zeta \leq 1;\ \xi \in \mathbb{C}\backslash\{0\}$.

2. Definition. A function $f \in \mathcal{A}$ is said to be in the class $\mathcal{M}_g(n,\alpha,\beta,\mu,\lambda,\zeta,\xi;m,\tau)$ if it satisfies the following nonhomogenous Cauchy-Euler differential equation:

$$
z^m\frac{d^mw}{dz^m} + \binom{m}{1}(\tau+m-1) z^{m-1}\frac{d^{m-1}w}{dz^{m-1}} + \cdots + \binom{m}{m} w \prod_{j=0}^{m-1}(\tau+j) = q(z)\prod_{j=0}^{m-1}(\tau+j+1)
$$

$$
(w = f(z) \in \mathcal{A};\ q \in \mathcal{M}_g(n,\alpha,\beta,\mu,\lambda,\zeta,\xi);\ m \in \mathbb{N}^*;\ \tau \in (-1,\infty)). \qquad (1.4)
$$

Remark 1. There are many choices of the function g which would provide interesting subclasses of analytic functions of complex order. In particular,
(**i**) if we choose the function g as

$$g(z) = \frac{1+Az}{1+Bz} \quad (-1 \leq B < A \leq 1;\ z \in \mathbb{U}),$$

it is easy to verify that g is a convex function in $\mathbb{U}$ and satisfies the hypotheses of Definition 1. If $f \in \mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi)$, then we have

$$1 + \frac{1}{\xi}\left(\frac{z\left[\zeta D_\lambda^{n+1}(\alpha,\beta,\mu) f(z) + (1-\zeta) D_\lambda^n(\alpha,\beta,\mu) f(z)\right]'}{\zeta D_\lambda^{n+1}(\alpha,\beta,\mu) f(z) + (1-\zeta) D_\lambda^n(\alpha,\beta,\mu) f(z)} - 1\right) \prec \frac{1+Az}{1+Bz} \quad (z \in \mathbb{U}).$$

We denote this new class by $\mathcal{H}(n, \alpha, \beta, \mu, \lambda, \zeta, \xi, A, B)$. Also we denote by $\mathcal{B}(n, \alpha, \beta, \mu, \lambda, \zeta, \xi, A, B; m, \tau)$ for corresponding class to $\mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi; m, \tau)$;
(**ii**) if we choose the function g as

$$g(z) = \frac{1+(1-2\gamma)z}{1-z} \quad (0 \leq \gamma < 1;\ z \in \mathbb{U}),$$

it is easy to verify that g is a convex function in $\mathbb{U}$ and satisfies the hypotheses of Definition 1. If $f \in \mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi)$, then we have

$$\Re\left\{1 + \frac{1}{\xi}\left(\frac{z\left[\zeta D_\lambda^{n+1}(\alpha,\beta,\mu) f(z) + (1-\zeta) D_\lambda^n(\alpha,\beta,\mu) f(z)\right]'}{\zeta D_\lambda^{n+1}(\alpha,\beta,\mu) f(z) + (1-\zeta) D_\lambda^n(\alpha,\beta,\mu) f(z)} - 1\right)\right\} > \gamma \quad (z \in \mathbb{U}),$$

that is

$$f \in \Psi(n, \alpha, \beta, \mu, \lambda, \zeta, \gamma, \xi).$$

Remark 2. In view of Remark 1(ii), by taking

$$g(z) = \frac{1+(1-2\gamma)z}{1-z} \quad (0 \leq \gamma < 1;\ z \in \mathbb{U})$$

in Definitions 1 and 2, we easily observe that the function classes

$$\mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi) \qquad \text{and} \qquad \mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi; 2, \tau)$$

become the aforementioned function classes

$$\Psi(n, \alpha, \beta, \mu, \lambda, \zeta, \gamma, \xi) \qquad \text{and} \qquad \Phi(n, \alpha, \beta, \mu, \lambda, \zeta, \gamma, \xi, \tau),$$

respectively.

In this work, by using the principle of subordination, we obtain coefficient bounds for functions in the subclasses

$$\mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi) \qquad \text{and} \qquad \mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi; m, \tau)$$

of analytic functions of complex order, which we have introduced here. Our results would unify and extend the corresponding results obtained earlier by Robertson [13], Nasr and Aouf [12], Altıntaş et al. [1], Faisal and Darus [8], Srivastava et al. [16], and others.

In our investigation, we shall make use of the principle of subordination between analytic functions, which is explained in Definition 3 below (see [11]).

3. Definition. For two functions f and g, analytic in $\mathbb{U}$, we say that the function $f(z)$ is subordinate to $g(z)$ in $\mathbb{U}$, and write

$$f(z) \prec g(z) \quad (z \in \mathbb{U}),$$

if there exists a Schwarz function $\mathfrak{w}(z)$, analytic in $\mathbb{U}$, with

$$\mathfrak{w}(0) = 0 \quad \text{and} \quad |\mathfrak{w}(z)| < 1 \quad (z \in \mathbb{U}),$$

such that

$$f(z) = g(\mathfrak{w}(z)) \quad (z \in \mathbb{U}).$$

In particular, if the function g is univalent in $\mathbb{U}$, the above subordination is equivalent to

$$f(0) = g(0) \quad \text{and} \quad f(\mathbb{U}) \subset g(\mathbb{U}).$$

2. Main results and their demonstration

In order to prove our main results (Theorems 1 and 2 below), we first recall the following lemma due to Rogosinski [14].

1. Lemma. Let the function $\mathfrak{g}$ given by

$$\mathfrak{g}(z) = \sum_{k=1}^{\infty} \mathfrak{b}_k z^k \qquad (z \in \mathbb{U})$$

be convex in $\mathbb{U}$. Also let the function $\mathfrak{f}$ given by

$$\mathfrak{f}(z) = \sum_{k=1}^{\infty} \mathfrak{a}_k z^k \qquad (z \in \mathbb{U})$$

be holomorphic in $\mathbb{U}$. If

$$\mathfrak{f}(z) \prec \mathfrak{g}(z) \qquad (z \in \mathbb{U}),$$

then

$$|\mathfrak{a}_k| \leq |\mathfrak{b}_1| \qquad (k \in \mathbb{N}).$$

We now state and prove each of our main results given by Theorems 1 and 2 below.

1. Theorem. Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi)$, then

$$(2.1) \qquad |a_i| \leq \frac{(\alpha+\beta)^{n+1} \prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{(i-1)! \, [\alpha + \zeta(\mu+\lambda)(i-1) + \beta] \, [\alpha + (\mu+\lambda)(i-1) + \beta]^n} \qquad (i \in \mathbb{N}^*).$$

Proof. Let the function $f \in \mathcal{A}$ be given by (1.1). Suppose that the function $\mathcal{F}(z)$ is defined, in terms of the differential operator $D_\lambda^n(\alpha, \beta, \mu)$, by

$$(2.2) \qquad \mathcal{F}(z) = \zeta D_\lambda^{n+1}(\alpha, \beta, \mu) f(z) + (1-\zeta) D_\lambda^n(\alpha, \beta, \mu) f(z) \quad (z \in \mathbb{U}).$$

Then, clearly, $\mathcal{F}$ is an analytic function in $\mathbb{U}$, and a simple computation shows that $\mathcal{F}$ has the following power series expansion:

$$(2.3) \qquad \mathcal{F}(z) = z + \sum_{i=2}^{\infty} A_i z^i \quad (z \in \mathbb{U}),$$

where, for convenience,

$$(2.4) \qquad A_i = \frac{[\alpha + \zeta(\mu+\lambda)(i-1) + \beta] \, [\alpha + (\mu+\lambda)(i-1) + \beta]^n}{(\alpha+\beta)^{n+1}} a_i \quad (i \in \mathbb{N}^*).$$

From Definition 1 and (2.2), we thus have

$$1 + \frac{1}{\xi}\left(\frac{z\mathcal{F}'(z)}{\mathcal{F}(z)} - 1\right) \in g(\mathbb{U}) \quad (z \in \mathbb{U}).$$

Let us define the function $p(z)$ by

$$(2.5) \qquad p(z) = 1 + \frac{1}{\xi}\left(\frac{z\mathcal{F}'(z)}{\mathcal{F}(z)} - 1\right) \quad (z \in \mathbb{U}).$$

Hence we deduce that

$$p(0) = g(0) = 1 \quad \text{and} \quad p(z) \in g(\mathbb{U}) \quad (z \in \mathbb{U}).$$

Therefore, we have

$$p(z) \prec g(z) \quad (z \in \mathbb{U}).$$

Thus, according to the Lemma 1, we obtain

$$\left|\frac{p^{(l)}(0)}{l!}\right| \leq \left|g'(0)\right| \quad (l \in \mathbb{N}). \tag{2.6}$$

Also from (2.5), we find

$$z\mathcal{F}'(z) = [1 + \xi(p(z) - 1)]\,\mathcal{F}(z). \tag{2.7}$$

Next, we suppose that

$$p(z) = 1 + c_1 z + c_2 z^2 + \cdots \quad (z \in \mathbb{U}). \tag{2.8}$$

Since $A_1 = 1$, in view of (2.3), (2.7) and (2.8), we obtain

$$(i-1)\,A_i = \xi\,\{c_{i-1} + c_{i-2}A_2 + \cdots + c_1 A_{i-1}\} \quad (i \in \mathbb{N}^*). \tag{2.9}$$

By combining (2.6) and (2.9), for $i = 2, 3, 4$, we obtain

$$|A_2| \leq |\xi|\left|g'(0)\right|,$$

$$|A_3| \leq \frac{|\xi|\,|g'(0)|\,(1 + |\xi|\,|g'(0)|)}{2!},$$

$$|A_4| \leq \frac{|\xi|\,|g'(0)|\,(1 + |\xi|\,|g'(0)|)\,(2 + |\xi|\,|g'(0)|)}{3!},$$

respectively. Also, by using the principle of mathematical induction, we obtain

$$|A_i| \leq \frac{\prod_{j=0}^{i-2}[j + |\xi|\,|g'(0)|]}{(i-1)!} \quad (i \in \mathbb{N}^*).$$

Now from (2.4), it is clear that

$$|a_i| \leq \frac{(\alpha+\beta)^{n+1}\prod_{j=0}^{i-2}[j + |\xi|\,|g'(0)|]}{(i-1)!\,[\alpha + \zeta(\mu+\lambda)(i-1) + \beta]\,[\alpha + (\mu+\lambda)(i-1) + \beta]^n} \quad (i \in \mathbb{N}^*).$$

This evidently completes the proof of Theorem 1. ∎

2. Theorem. Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi; m, \tau)$, then

$$|a_i| \leq \frac{(\alpha+\beta)^{n+1}\prod_{j=0}^{i-2}[j + |\xi|\,|g'(0)|]\prod_{j=0}^{m-1}(\tau + j + 1)}{(i-1)!\,[\alpha + \zeta(\mu+\lambda)(i-1) + \beta]\,[\alpha + (\mu+\lambda)(i-1) + \beta]^n\prod_{j=0}^{m-1}(\tau + j + i)} \quad (i \in \mathbb{N}^*).$$

Proof. Let the function $f \in \mathcal{A}$ be given by (1.1). Also let

$$h(z) = z + \sum_{i=2}^{\infty} b_i z^i \in \mathcal{M}_g(n, \alpha, \beta, \mu, \lambda, \zeta, \xi).$$

Hence, from (1.4), we deduce that

$$a_i = \frac{\prod_{j=0}^{m-1} (\tau + j + 1)}{\prod_{j=0}^{m-1} (\tau + j + i)} b_i \quad \left(i \in \mathbb{N}^*, \ \tau \in (-1, \infty)\right).$$

Thus, by using Theorem 1, we obtain

$$|a_i| \leq \frac{(\alpha+\beta)^{n+1} \prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{(i-1)! \, [\alpha + \zeta (\mu+\lambda)(i-1) + \beta] \, [\alpha + (\mu+\lambda)(i-1) + \beta]^n} \frac{\prod_{j=0}^{m-1} (\tau + j + 1)}{\prod_{j=0}^{m-1} (\tau + j + i)}.$$

This completes the proof of Theorem 2. ∎

3. Corollaries and consequences

In this section, we apply our main results (Theorems 1 and 2 of Section 2) in order to deduce each of the following corollaries and consequences.

1. Corollary. ([19]) Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(0, \alpha, \beta, \mu, \lambda, \zeta, \xi) \equiv \mathcal{S}_g(\zeta, \xi)$, then

$$|a_i| \leq \frac{\prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{(i-1)! \, (1 + \zeta (i-1))} \quad \left(i \in \mathbb{N}^*\right).$$

2. Corollary. ([19]) Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(0, \alpha, \beta, \mu, \lambda, \zeta, \xi; m, \tau) \equiv \mathcal{K}_g(\zeta, \xi, m; \tau)$, then

$$|a_i| \leq \frac{\prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{(i-1)! \, (1 + \zeta (i-1))} \frac{\prod_{j=0}^{m-1} (\tau + j + 1)}{\prod_{j=0}^{m-1} (\tau + j + i)} \quad \left(i \in \mathbb{N}^*\right).$$

3. Corollary. ([17]) Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(n, 1, 0, 0, 1, \zeta, \xi) \equiv \mathcal{M}_g(n, \zeta, \xi)$, then

$$|a_i| \leq \frac{\prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{i^n \, (1 + \zeta (i-1)) \, (i-1)!} \quad \left(i \in \mathbb{N}^*\right).$$

4. Corollary. ([17]) Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\mathcal{M}_g(n, 1, 0, 0, 1, \zeta, \xi; 2, \tau) \equiv \mathcal{M}_g(n, \zeta, \xi; \tau)$, then

$$|a_i| \leq \frac{(1+\tau)(2+\tau) \prod_{j=0}^{i-2} [j + |\xi| \, |g'(0)|]}{i^n \, (1 + \zeta (i-1)) \, (i-1)! \, (i+\tau)(i+1+\tau)} \quad \left(i \in \mathbb{N}^*\right).$$

Setting

$$m = 2 \quad \text{and} \quad g(z) = \frac{1 + (1-2\gamma) z}{1 - z} \quad (0 \leq \gamma < 1; \ z \in \mathbb{U})$$

in Theorems 1 and 2, we have following corollaries, respectively.

5. Corollary. [8] Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\Psi(n, \alpha, \beta, \mu, \lambda, \zeta, \gamma, \xi)$, then

$$|a_i| \leq \frac{(\alpha+\beta)^{n+1} \prod_{j=0}^{i-2} [j + 2|\xi|(1-\gamma)]}{(i-1)!\ [\alpha + \zeta(\mu+\lambda)(i-1) + \beta]\ [\alpha + (\mu+\lambda)(i-1) + \beta]^n} \quad (i \in \mathbb{N}^*).$$

6. Corollary. [8] Let the function $f \in \mathcal{A}$ be defined by (1.1). If the function f is in the class $\Phi(n, \alpha, \beta, \mu, \lambda, \zeta, \gamma, \xi, \tau)$, then

$$|a_i| \leq \frac{(1+\tau)(2+\tau)(\alpha+\beta)^{n+1} \prod_{j=0}^{i-2} [j + 2|\xi|(1-\gamma)]}{(i+\tau)(i+1+\tau)(i-1)!\ [\alpha + \zeta(\mu+\lambda)(i-1) + \beta]\ [\alpha + (\mu+\lambda)(i-1) + \beta]^n} \quad (i \in \mathbb{N}^*).$$

For several other closely-related investigations, see (for example) the recent works [1-7, 12, 13].

References

[1] Altıntaş, O., Irmak, H., Owa, S. and Srivastava, H.M. *Coefficient bounds for some families of starlike and convex functions of complex order*, Appl. Math. Letters **20**, 1218-1222, 2007.

[2] Altıntaş, O., Irmak, H. and Srivastava, H.M. *Fractional calculus and certain starlike functions with negative coefficients*, Comput. Math. Appl. **30** (2), 9-15, 1995.

[3] Altıntaş, O. and Özkan, Ö. *Starlike, convex and close-to-convex functions of complex order*, Hacet. Bull. Nat. Sci. Eng. Ser. B **28**, 37-46, 1991.

[4] Altıntaş, O. and Özkan, Ö. *On the classes of starlike and convex functions of complex order*, Hacet. Bull. Nat. Sci. Eng. Ser. B **30**, 63-68, 2001.

[5] Altıntaş, O., Özkan, Ö. and Srivastava, H.M. *Neighborhoods of a class of analytic functions with negative coefficients*, Appl. Math. Letters **13** (3), 63-67, 1995.

[6] Altıntaş, O., Özkan, Ö. and Srivastava, H.M. *Majorization by starlike functions of complex order*, Complex Variables Theory Appl. **46**, 207-218, 2001.

[7] Altıntaş, O., Özkan, Ö. and Srivastava, H.M. *Neighborhoods of a certain family of multivalent functions with negative coefficient*, Comput. Math. Appl. **47**, 1667-1672, 2004.

[8] Faisal, I. and Darus, M. *Application of nonhomogenous Cauchy-Euler differential equation for certain class of analytic functions*, Hacet. J. Math. Stat. **43** (3), 375–382, 2014.

[9] Flett, T.M. *The dual of an inequality of Hardy and Littlewood and some related inequalities*, J. Math. Anal. Appl. **38**, 746-765, 1972.

[10] Jurg, I.B., Kim, Y.C. and Srivastava, H.M. *The Hardy space of analytic functions associated with certain one-parameter families of integral operators*, J. Math. Anal. Appl. **176**, 138-147, 1993.

[11] Miller, S.S. and Mocanu, P.T. *Differential subordinations*: Theory and applications. Monographs and Textbooks in Pure and Applied Mathematics, **225**. Marcel Dekker, Inc., New York, 2000.

[12] Nasr, M.A. and Aouf, M.K. *Radius of convexity for the class of starlike functions of complex order*, Bull. Fac. Sci. Assiut Univ. A **12** (1), 153-159, 1983.

[13] Robertson, M.S. *On the theory of univalent functions*, Ann. of Math. (2) **37** (2), 374-408, 1936.

[14] Rogosinski, W. *On the coefficients of subordinate functions*, Proc. London Math. Soc. (Ser. 2) **48**, 48-82, 1943.

[15] Sălăgean, G.Ş. *Subclasses of univalent functions*, Complex Analysis-Fifth Romanian-Finnish seminar, Part 1 (Bucharest, 1981), Lecture Notes in Math., vol. **1013**, Springer, Berlin, 1983, pp. 362-372.

[16] Srivastava, H.M., Altıntaş, O. and Serenbay, S.K., *Coefficient bounds for certain subclasses of starlike functions of complex order*, Appl. Math. Letters **24** (8), 1359-1363, 2011.

[17] Srivastava, H.M., Xu, Q.-H. and Wu, G.-P. *Coefficient estimates for certain subclasses of spiral-like functions of complex order*, Appl. Math. Letters **23**, 763-768, 2010.

[18] Uralegaddi, B.A. and Somanatha, C. *Certain classes of univalent functions*. Current topics in analytic function theory, 371-374, World Sci. Publ., River Edge, NJ, 1992.
[19] Xu, Q.-H., Gui, Y.-C. and Srivastava, H.M. *Coefficient estimates for certain subclasses of analytic functions of complex order*, Taiwanese J. Math. **15** (5), 2377-2386, 2011.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1023 – 1032

Some normal subgroups of extended generalized Hecke groups

Bilal Demir*†, Özden Koruoğlu‡ and Recep Şahin§

In memory of my dear son Can Şahin.

Abstract

Generalized Hecke group $H_{p,\infty}(\lambda)$ is generated by $X(z) = -(z - \lambda_p)^{-1}$ and $Y(z) = -(z + \lambda)^{-1}$ where $\lambda_p = 2\cos\frac{\pi}{p}$, $p \geq 2$ integer and $\lambda \geq 2$. Extended generalized Hecke group $\overline{H}_{p,\infty}(\lambda)$ is obtained by adding the reflection $R(z) = 1/\overline{z}$ to the generators of generalized Hecke group $H_{p,\infty}(\lambda)$. In this paper, we study the commutator subgroups of extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$. Also, we determine the power subgroups of generalized Hecke groups $H_{p,\infty}(\lambda)$ and extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$.

Keywords: Generalized Hecke groups, Extended generalized Hecke groups, Commutator subgroups, Power subgroups.

2000 AMS Classification: 20H10, 11F06.

Received : 13.01.2015 *Accepted* : 24.08.2015 *Doi* : 10.15672/HJMS.20164513108

*Balıkesir University, Necatibey Faculty of Education, Department of Secondary Mathematics Education,10100 Balıkesir, Turkey, Email: bdemir@balikesir.edu.tr

†Corresponding Author.

‡Balıkesir University, Necatibey Faculty of Education, Department of Elementary Mathematics Education,10100 Balıkesir, Turkey, Email: ozdenk@balikesir.edu.tr

§Balıkesir University, Faculty of Arts and Sciences,Department of Mathematics,10145 Çağış Campus, Balıkesir, Turkey, Email: rsahin@balikesir.edu.tr

1. Introduction

In [1], Hecke introduced the groups $H(\lambda)$ generated by two linear fractional transformations

$$T(z) = -\frac{1}{z} \quad \text{and} \quad U(z) = z + \lambda,$$

where λ is a fixed positive real number. Let $S = TU$, i.e.,

$$S(z) = -\frac{1}{z+\lambda}.$$

Hecke showed that $H(\lambda)$ is discrete if and only if either $\lambda = \lambda_q = 2\cos(\frac{\pi}{q})$, $q \geq 3$ integer, or $\lambda \geq 2$. These groups have come to be known as the *Hecke groups* and we will denote them by H_q, or by $H(\lambda)$, respectively. The first few Hecke groups are $H_3 = PSL(2,\mathbb{Z})$ (the modular group), $H_4 = H(\sqrt{2})$, $H_5 = H(\frac{1+\sqrt{5}}{2})$, and $H_6 = H(\sqrt{3})$ for $q = 3, 4, 5$ and 6, respectively.

It is known that when $\lambda = \lambda_q = 2\cos(\frac{\pi}{q})$, $q \geq 3$ integer, Hecke group H_q is isomorphic to the free product of two finite cyclic groups of orders 2 and q,

$$H_q =< T, S \mid T^2 = S^q = I >\cong C_2 * C_q,$$

and when $\lambda \geq 2$, Hecke group $H(\lambda)$ is a free product of a cyclic group of order 2 and infinity, so all such $H(\lambda)$ have the same algebraic structure, i.e.

$$H(\lambda) =< T,\ S \mid T^2 = I >\cong C_2 * \mathbb{Z}.$$

Also Hecke group H_q or $H(\lambda)$ is the Fuchsian group of the first kind when either $\lambda = \lambda_q = 2\cos(\frac{\pi}{q})$, $q \geq 3$ integer or $\lambda = 2$, and $H(\lambda)$ is the Fuchsian group of the second kind when $\lambda > 2$.

On the other hand, Lehner studied in [2] more general class $H_{p,q}$ of Hecke groups H_q, by taking

$$X = \frac{-1}{z-\lambda_p} \quad \text{and} \quad V = z + \lambda_p + \lambda_q,$$

where $2 \leq p \leq q \leq \infty$, $p + q > 4$. Here if we take $Y = XV = -\frac{1}{z+\lambda_q}$, then we have the presentation,

$$H_{p,q} =< X, Y \mid X^p = Y^q = I >\cong C_p * C_q. \tag{1.1}$$

We call these groups as *generalized Hecke groups* $H_{p,q}$. We know from [2] that $H_{2,q} = H_q$, $|H_q : H_{q,q}| = 2$, and there is no group $H_{2,2}$. Also, all Hecke groups H_q are included in generalized Hecke groups $H_{p,q}$. Also, generalized Hecke groups $H_{p,q}$ have been studied extensively for many aspects in the literature (for examples, please see, [3], [4], [5], [6], [7] and [8]).

Extended generalized Hecke groups $\overline{H}_{p,q}$ have been defined in [9] and [10], similar to extended Hecke groups $\overline{H}_q$ (please see, [11] and [12]), by adding the reflection $R(z) = 1/\overline{z}$ to the generators of generalized Hecke group $H_{p,q}$. From [9], extended generalized Hecke groups $\overline{H}_{p,q}$ have a presentation

$$\overline{H}_{p,q} =< X, Y, R \mid X^p = Y^q = R^2 = I,\ RX = X^{-1}R, RY = Y^{-1}R >,$$

or

$$\overline{H}_{p,q} =< X, Y, R \mid X^p = Y^q = R^2 = (XR)^2 = (YR)^2 = I >\cong D_p *_{C_2} D_q.$$

The group $H_{p,q}$ is a subgroup of index 2 in $\overline{H}_{p,q}$.

In (1.1), if $q = \infty$, then we have more general class $H_{p,\infty}$, of Hecke groups $H(\lambda)$.

Now we can give the following definitions;

1.1. Definition. Let $\lambda_p = 2\cos\frac{\pi}{p}$, $p \geq 2$ integer and let $\lambda \geq 2$. Generalized Hecke groups $H_{p,\infty}(\lambda)$ are defined as the groups generated by

$$X = \frac{-1}{z-\lambda_p} \text{ and } Y = -\frac{1}{z+\lambda},$$

and have a presentation

$$H_{p,\infty}(\lambda) = < X, Y \mid X^p = Y^\infty = I > \cong C_p * \mathbb{Z}.$$

1.2. Definition. Extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$, are defined by adding reflection $R(z) = 1/\overline{z}$ to the generators of generalized Hecke groups $H_{p,\infty}(\lambda)$ and have a presentation

$$\overline{H}_{p,\infty}(\lambda) = < X, Y, R \mid X^p = Y^\infty = R^2 = I, RX = X^{p-1}R, RY = Y^{-1}R >,$$

or

$$\begin{aligned} \overline{H}_{p,\infty}(\lambda) &= < X, Y, R \mid X^p = Y^\infty = R^2 = (XR)^2 = (YR)^2 = I >, \\ &\cong D_p *_{C_2} D_\infty. \end{aligned}$$

In this paper, we study the commutator subgroups of extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$. Then, we determine the power subgroups of generalized Hecke groups $H_{p,\infty}(\lambda)$ and extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$. We use the Reidemeister-Schreier method to get the generators of all these subgroups.

Let G be a group and N be a normal subgroup of G with finite index. According to the Reidemeister-Schreier method we get the generators of N as follows: We first choose a Schreier transversal Σ for the quotient group G/N such that all certain words of generators including.Note that this transversal is not unique. Then we get the generators of N as following order:

$$\text{(An element of } \Sigma) \times \text{(A generator of } G) \times$$
$$\text{(coset representative of the preceeding product)}^{-1}.$$

For more details please see [13].

Commutator subgroups and power subgroups of Hecke and extended Hecke groups have been studied in, [14], [15], [17], [20], [23], [24] and [25]. Here, our aim is to generalize the results given in [14] and [15] for Hecke groups $H(\lambda)$ and extended Hecke groups $\overline{H}(\lambda)$ to extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$.

2. Commutator Subgroups of Extended Generalized Hecke Groups $\overline{H}_{p,\infty}(\lambda)$

Since the index of the commutator subgroup $H'_{p,\infty}(\lambda)$ in $H_{p,\infty}(\lambda)$ is infinite, we study only the commutator subgroup $\overline{H}'_{p,\infty}(\lambda)$ of extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$.

Here, we investigate the cases of p, odd or even, seperately.

2.1. Theorem. *Let $p \geq 3$ be an odd integer and let $\lambda \geq 2$. Then*

1) $\left|\overline{H}_{p,\infty}(\lambda) : \overline{H}'_{p,\infty}(\lambda)\right| = 4.$

2) $\overline{H}'_{p,\infty}(\lambda) = < X, YXY^{-1}, Y^2 \mid X^p = (YXY^{-1})^p$
$= (Y^2)^\infty = I > \cong C_p * C_p * \mathbb{Z}.$

Proof. 1) Firstly, we set up the quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda)$ which can be construct by adding the abelianizing relation to the relations of $\overline{H}_{p,\infty}(\lambda)$. Then

$$\begin{aligned} \overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda) = < X, Y, R \mid X^p = Y^\infty = R^2 = I,\ RX = X^{p-1}R, \\ RY = Y^{-1}R,\ XR = RX,\ YR = RY,\ XY = YX > . \end{aligned}$$

Since p is odd and from the relations $RX = X^{p-1}R$ and $RX = XR$, we have $X = I$. Also we get $Y^2 = I$ from the relations $RY = Y^{-1}R$ and $YR = RY$. Thus we have

$$\overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda) =< Y, R \mid Y^2 = R^2 = (YR)^2 = I >\simeq C_2 \times C_2.$$

2) Now we determine the set of generators for $\overline{H}'_{p,\infty}(\lambda)$. We choose a Schreier transversal for $\overline{H'}_{p,\infty}(\lambda)$ as $\Sigma = \{I, Y, R, YR\}$. According to Reidemeister-Schreier method we can form all possible products;

$$\begin{array}{lll} I.X.(I)^{-1} = X, & I.Y.(Y)^{-1} = I, & I.R.(R)^{-1} = I, \\ Y.X.(Y)^{-1} = YXY^{-1}, & Y.Y.(I)^{-1} = Y^2, & Y.R.(YR)^{-1} = I, \\ R.X.(R)^{-1} = X^{p-1}, & R.Y.(YR)^{-1} = Y^{-2}, & R.R.(I)^{-1} = I, \\ YR.X.(YR)^{-1} = YX^{p-1}Y^{-1}, & YR.Y.(R)^{-1} = I, & YR.R.(Y)^{-1} = I. \end{array}$$

Since $X^{-1} = X^{p-1}, (YXY^{-1})^{-1} = YX^{p-1}Y^{-1}$ and $(Y^2)^{-1} = Y^{-2}$, the generators are X, YXY^{-1} and Y^2. Thus $\overline{H'}_{p,\infty}(\lambda)$ has a presentation

$$\begin{aligned} \overline{H}'_{p,\infty}(\lambda) &= < X, YXY^{-1}, Y^2 \mid X^p = (YXY^{-1})^p \\ &= (Y^2)^\infty = I >\cong C_p * C_p * \mathbb{Z}. \end{aligned}$$

□

2.2. Theorem. *Let $p \geq 2$ be an even integer and let $\lambda \geq 2$. Then*

1) $\left|\overline{H}_{p,\infty}(\lambda) : \overline{H}'_{p,\infty}(\lambda)\right| = 8.$

2)

$$\begin{aligned} \overline{H}'_{p,\infty}(\lambda) &= < X^2, YX^2Y^{-1},\ XYXY^{-1}, Y^2, XY^2X^{-1} \mid (X^2)^{p/2} \\ &= (YX^2Y^{-1})^{p/2} = (XYXY^{-1})^\infty = (Y^2)^\infty = (XY^2X^{-1})^\infty = I > \\ &\cong C_{p/2} * C_{p/2} * \mathbb{Z} * \mathbb{Z} * \mathbb{Z}. \end{aligned}$$

Proof. 1) Similar to the previous proof, we have the quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda)$ as

$$\begin{aligned} \overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda) =&< X, Y, R \mid X^p = Y^\infty = R^2 = I,\ RX = X^{p-1}R, \\ &RY = Y^{-1}R,\ XR = RX,\ YR = RY,\ XY = YX > . \end{aligned}$$

Since p is even and from the relations $RX = X^{p-1}R$, $XR = RX$, $RY = Y^{-1}R$ and $YR = RY$, we have $X^2 = I$ and $Y^2 = I$. Thus we get

$$\begin{aligned} \overline{H}_{p,\infty}(\lambda)/\overline{H}'_{p,\infty}(\lambda) &= < X, Y, R : X^2 = Y^2 = R^2 = (XY)^2 = (XR)^2 = (YR)^2 = I >, \\ &\cong C_2 \times C_2 \times C_2. \end{aligned}$$

2) Now we can determine the Schreier transversal as $\Sigma = \{I,\ X,\ Y,\ R,\ XR,\ YR,\ XY,\ XYR\}$. From the Reidemeister-Schreier method all possible products are;

$$\begin{array}{cc}
I.X.(X)^{-1} = I, & I.Y.(Y)^{-1} = I, \\
X.X.(I)^{-1} = X^2, & X.Y.(XY)^{-1} = I, \\
Y.X.(XY)^{-1} = YXY^{-1}X^{p-1}, & Y.Y.(I)^{-1} = Y^2, \\
R.X.(XR)^{-1} = X^{p-2}, & R.Y.(YR)^{-1} = Y^{-2}, \\
XR.X.(R)^{-1} = I, & XR.Y.(XYR)^{-1} = XY^{-2}X^{-1}, \\
YR.X.(XYR)^{-1} = YX^{-1}Y^{-1}X^{-1}, & YR.Y.(R)^{-1} = I, \\
XY.X.(Y)^{-1} = XYXY^{-1}, & XY.Y.(X)^{-1} = XY^2X^{-1}, \\
XYR.X.(YR)^{-1} = XYX^{-1}Y^{-1}, & XYR.Y.(XR)^{-1} = I,
\end{array}$$

$$\begin{array}{c}
I.R.(R)^{-1} = I, \\
X.R.(XR)^{-1} = I, \\
Y.R.(YR)^{-1} = I, \\
R.R.(I)^{-1} = I, \\
XR.R.(X)^{-1} = I, \\
YR.R.(Y)^{-1} = I, \\
XY.R.(XYR)^{-1} = I, \\
XYR.R.(XY)^{-1} = I.
\end{array}$$

Since $(X^2)^{-1} = X^{p-2}$, $(YXY^{-1}X^{p-1})^{-1} = XYX^{-1}Y^{-1}$, $(YX^{-1}Y^{-1}X^{-1})^{-1} = XYXY^{-1}$, $(Y^2)^{-1} = Y^{-2}$, $(XY^2X^{-1})^{-1} = XY^{-2}X^{-1}$, we have the presentation of $\overline{H}'_{p,\infty}(\lambda)$ as

$$\begin{aligned}
\overline{H}'_{p,\infty}(\lambda) &= \ < X^2, YX^2Y^{-1},\ XYXY^{-1}, Y^2, XY^2X^{-1} \mid (X^2)^{p/2} \\
&= (YX^2Y^{-1})^{p/2} = (XYXY^{-1})^{\infty} = (Y^2)^{\infty} = (XY^2X^{-1})^{\infty} = I > \\
&\cong C_{p/2} * C_{p/2} * \mathbb{Z} * \mathbb{Z} * \mathbb{Z}.
\end{aligned}$$

□

3. Power Subgroups of $H_{p,\infty}(\lambda)$ and $\overline{H}_{p,\infty}(\lambda)$

In this section, we consider the power subgroups of generalized Hecke groups $H_{p,\infty}(\lambda)$ and extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$. Here, we note that the power subgroups of Hecke groups H_q, or $H(\lambda)$ and extended Hecke groups $\overline{H}_q$, or $\overline{H}(\lambda)$ have been studied by many authors in [6], [7], [10], [11], [12], [14], [16], [18], [19], [21], [22].

Now we give some information about the power subgroups.

Let m be a positive integer. Let us define G^m to be the subgroup generated by the m^{th} powers of all elements of $G = H_{p,\infty}(\lambda)$ or $\overline{H}_{p,\infty}(\lambda)$. The subgroup G^m is called the m^{th} *– power subgroup* of G. As fully invariant subgroups, they are normal in G.

From the definition, it is easy to see that

$$G^{mk} < G^m$$

and

$$G^{mk} < (G^m)^k.$$

We now discuss the group theoretical structure of these subgroups. We find a presentation for the quotient G/G^m by adding the relation $A^m = I$ to the presentation of G. The order of G/G^m gives us the index. Thus we use the Reidemeister-Schreier process to find the presentation of the power subgroups G^m.

Let us start with $H_{p,\infty}(\lambda)$.

3.1. Theorem. 1) *Let $p > 2$ be an odd integer and $\lambda \geq 2$. Then,*
$H^2_{p,\infty}(\lambda) = < X, YXY^{-1},\ Y^2 \mid X^p = (YXY^{-1})^p = (Y^2)^{\infty} = I > \cong C_p * C_p * \mathbb{Z}$.

2) *Let $p \geq 2$ be an even integer and $\lambda \geq 2$. Then,*

$$H^2_{p,\infty}(\lambda) =< X^2, YX^2Y^{-1},\ XYXY^{-1}, Y^2, XY^2X^{-1} \mid (X^2)^{p/2}$$
$$= (YX^2Y^{-1})^{p/2} = (XYXY^{-1})^{\infty} = (Y^2)^{\infty} = (XY^2X^{-1})^{\infty} = I > .$$

Proof. 1) The quotient group $H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda)$ is

$$H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda) =< X, Y \mid X^p = Y^{\infty} = (XY)^{\infty} = X^2 = Y^2 = (XY)^2 = \cdots = I > .$$

Since $p > 2$ is an odd integer and from the relations $X^2 = X^p = I$ and $Y^2 = Y^{\infty} = I$, we have $X = Y^2 = I$. Thus we get

$$H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda) =< Y \mid Y^2 = I >\cong C_2.$$

If we choose a Schreier transversal as $\{I,\ Y\}$ and use the Reidemeister-Schreier method, we obtain all possible products;

$$\begin{array}{ll} I.X.(I)^{-1} = X, & I.Y.(Y)^{-1} = I, \\ Y.X.(Y)^{-1} = YXY^{-1}, & Y.Y.(I)^{-1} = Y^2. \end{array}$$

So we get the presentation of $H^2_{p,\infty}(\lambda)$ as

$$H^2_{p,\infty}(\lambda) =< X, YXY^{-1},\ Y^2 \mid X^p = (YXY^{-1})^p = (Y^2)^{\infty} = I >\cong C_p * C_p * \mathbb{Z}.$$

2) The quotient group $H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda)$ is

$$\begin{aligned} H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda) &= < X, Y \mid X^p = Y^{\infty} = (XY)^{\infty} \\ &= X^2 = Y^2 = (XY)^2 = \cdots = I > . \end{aligned}$$

Since $p \geq 2$ is an even integer and from the relations $X^2 = X^p = I$ and $Y^2 = Y^{\infty} = I$, we obtain $X^2 = Y^2 = I$. Thus we have

$$H_{p,\infty}(\lambda)/H^2_{p,\infty}(\lambda) =< X, Y \mid X^2 = Y^2 = (XY)^2 = I >\cong D_2.$$

Now we choose a Schreier transversal as $\{I,\ X,\ Y,\ XY\}$ for $H^2_{p,\infty}(\lambda)$. According to the Reidemeister-Schreier method, we can form all possible products;

$$\begin{array}{ll} I.X.(X)^{-1} = I, & I.Y.(Y)^{-1} = I, \\ X.X.(I)^{-1} = X^2, & X.Y.(XY)^{-1} = I, \\ Y.X.(XY)^{-1} = YXY^{-1}X^{-1}, & Y.Y.(I)^{-1} = Y^2, \\ XY.X.(Y)^{-1} = XYXY^{-1}, & XY.Y.(X)^{-1} = XY^2X^{-1}. \end{array}$$

Thus we obtain a presentation of $H^2_{p,\infty}(\lambda)$ as

$$\begin{aligned} H^2_{p,\infty}(\lambda) &= < X^2, YX^2Y^{-1},\ XYXY^{-1}, Y^2, XY^2X^{-1} \mid (X^2)^{p/2} \\ &= (YX^2Y^{-1})^{p/2} = (XYXY^{-1})^{\infty} = (Y^2)^{\infty} = (XY^2X^{-1})^{\infty} = I > \\ &\cong C_{p/2} * C_{p/2} * \mathbb{Z} * \mathbb{Z} * \mathbb{Z}. \end{aligned}$$

□

3.2. Theorem. *Let $\lambda \geq 2$. If m and p are positive integers such that $(m,p) = 1$, then*

$$\begin{aligned} H^m_{p,\infty}(\lambda) &= < X, YXY^{-1},\ Y^2XY^{-2}, \cdots, Y^{m-1}XY^{1-m}, Y^m \mid X^p \\ &= (YXY^{-1})^p = (Y^2XY^{-2})^p = \cdots = (Y^{m-1}XY^{1-m})^p = (Y^m)^{\infty} = I > \\ &\cong \underbrace{C_p * C_p * \cdots * C_p}_{m\ times} * \mathbb{Z}. \end{aligned}$$

Proof. The quotient group $H_{p,\infty}(\lambda)/H^m_{p,\infty}(\lambda)$ is

$$H_{p,\infty}(\lambda)/H^m_{p,\infty}(\lambda) =< X, Y \mid X^p = Y^{\infty} = (XY)^{\infty} = X^m = Y^m = (XY)^m = \cdots = I > .$$

Since $(m,p) = 1$ and from the relations $X^p = X^m = I$, we find $X = I$. Thus we have

$$H_{p,\infty}(\lambda)/H^m_{p,\infty}(\lambda) =< Y : Y^m = I >\cong C_m.$$

Then we choose the Schreier transversal as $\Sigma = \{I,\, Y,\, Y^2, ...,\, Y^{m-1}\}$. According to the Reidemeister-Schreier method, we get the following products;

$$\begin{array}{ll} I.X.(I)^{-1} = X, & I.Y.(Y)^{-1} = I, \\ Y.X.(Y)^{-1} = YXY^{-1}, & Y.Y.(Y^2)^{-1} = I, \\ Y^2.X.(Y^2)^{-1} = Y^2XY^{-2}, & Y^2.Y.(Y^3)^{-1} = I, \\ Y^3.X.(Y^3)^{-1} = Y^3XY^{-3}, & Y^3.Y.(Y^4)^{-1} = I, \\ \vdots & \vdots \\ Y^{m-1}.X.(Y^{m-1})^{-1} = Y^{m-1}XY^{1-m}, & Y^{m-1}.Y.(I)^{-1} = Y^m. \end{array}$$

So we have a presentation of $H^2_{p,\infty}(\lambda)$ as

$$\begin{aligned} H^m_{p,\infty}(\lambda) &= \ < X, YXY^{-1},\ Y^2XY^{-2}, \cdots, Y^{m-1}XY^{1-m}, Y^m \mid X^p \\ &= (YXY^{-1})^p = (Y^2XY^{-2})^p = \cdots = (Y^{m-1}XY^{1-m})^p = (Y^m)^\infty = I > \\ &\cong \underbrace{C_p * C_p * \cdots * C_p}_{m \text{ times}} * \mathbb{Z}. \end{aligned}$$

□

The case $(m,p) = d > 1$, except of $m = 2$ and p even, is more complex, since the index of quotient group $H_{p,\infty}(\lambda)/H^m_{p,\infty}(\lambda)$ is unknown. In this case, we have the relations $X^d = Y^m = (XY)^m = \cdots = I$ and can not say anything about the power subgroups $H^m_{p,\infty}(\lambda)$.

Now we consider the power subgroups $\overline{H}^m_{p,\infty}(\lambda)$ of extended generalized Hecke groups $\overline{H}_{p,\infty}(\lambda)$. Here, we interest with the cases such that the index of the quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}^m_{p,\infty}(\lambda)$ is finite.

3.3. Theorem. 1) *Let $p > 2$ be an odd integer and $\lambda \geq 2$. Then,*
$\overline{H}^2_{p,\infty}(\lambda) =< X, YXY^{-1},\ Y^2 \mid X^p = (YXY^{-1})^p = (Y^2)^\infty = I > \cong C_p * C_p * \mathbb{Z}$.
2) *Let $p \geq 2$ be an even integer and $\lambda \geq 2$. Then,*
$\overline{H}^2_{p,\infty}(\lambda) =< X^2, YX^2Y^{-1},\ XYXY^{-1}, Y^2, XY^2X^{-1} \mid (X^2)^{p/2}$
$= (YX^2Y^{-1})^{p/2} = (XYXY^{-1})^\infty = (Y^2)^\infty = (XY^2X^{-1})^\infty = I >$.

Proof. The quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}^2_{p,\infty}(\lambda)$ is

$$\begin{aligned} \overline{H}_{p,\infty}(\lambda)/\overline{H}^2_{p,\infty}(\lambda) =< X, Y, R \mid X^p = Y^\infty = R^2 = (XR)^2 = (YR)^2 \\ = X^2 = Y^2 = (XY)^2 = \cdots = I > . \end{aligned}$$

The rest of the proof is similar to the proof of the Theorems 1 and 2. □

By using the Theorems 1, 2, 3 and 5, we can give the following.

3.4. Corollary. $\overline{H}^2_{p,\infty}(\lambda) = H^2_{p,\infty}(\lambda) = \overline{H}'_{p,\infty}(\lambda)$.

3.5. Theorem. 1) *Let $\lambda \geq 2$ and let $p \geq 3$ be an odd number. If m is an even positive integer such that $(m,p) = 1$, then*

$$\begin{aligned} \overline{H}^m_{p,\infty}(\lambda) &= \ < X, YXY^{-1},\ Y^2XY^{-2}, \cdots, Y^{m-1}XY^{1-m}, Y^m \mid X^p \\ &= (YXY^{-1})^p = (Y^2XY^{-2})^p = \cdots = (Y^{m-1}XY^{1-m})^p = (Y^m)^\infty = I > \\ &\cong \underbrace{C_p * C_p * \cdots * C_p}_{m\ times} * \mathbb{Z}. \end{aligned}$$

2) *Let $\lambda \geq 2$. If $m > 0$ is odd integer such that $(m,p) = 1$, then $\overline{H}^m_{p,\infty}(\lambda) = \overline{H}_{p,\infty}(\lambda)$.*

Proof. 1) The quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda)$ is

$$\begin{aligned}\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda) = &< X, Y, R \mid X^p = Y^{\infty} = R^2 = (XR)^2 = (YR)^2 \\ &= X^m = Y^m = (XY)^m = \cdots = I > .\end{aligned}$$

Since $(m, p) = 1$ and m is even, we have $X = I$.

$$\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda) =< Y, R : Y^m = R^2 = (YR)^2 = ... = I >\cong D_m.$$

Considering the presentation of quotient group we can choose Schreier transversal as $\Sigma = \{I,\ Y,\ Y^2, ...,\ Y^{m-1},\ R,\ RY,\ RY^2, ...,\ RY^{m-1}\}$. Then the process as following;

$$\begin{array}{ll}
I.X.(I)^{-1} = X, & I.Y.(Y)^{-1} = I, \\
Y.X.(Y)^{-1} = YXY^{-1}, & Y.Y.(Y^2)^{-1} = I, \\
Y^2.X.(Y^2)^{-1} = Y^2XY^{-2}, & Y^2.Y.(Y^3)^{-1} = I, \\
\vdots & \vdots \\
Y^{m-1}.X.(Y^{m-1})^{-1} = Y^{m-1}XY^{1-m}, & Y^{m-1}.Y.(I)^{-1} = Y^m, \\
R.X.(R)^{-1} = X^{p-1}, & R.Y.(RY)^{-1} = I, \\
RY.X.(RY)^{-1} = Y^{-1}X^{p-1}Y, & RY.Y.(RY^2)^{-1} = I, \\
RY^2.X.(RY^2)^{-1} = Y^{-2}X^{p-1}Y^{-2}, & RY^2.Y.(RY^3)^{-1} = I, \\
\vdots & \vdots \\
RY^{m-1}.X.(RY^{m-1})^{-1} = Y^{1-m}X^{p-1}Y^{m-1}, & RY^{m-1}.Y.(R)^{-1} = Y^{-m},
\end{array}$$

$$\begin{array}{l}
I.R.(R)^{-1} = I, \\
Y.R.(RY^{m-1})^{-1} = Y^m, \\
Y^2.R.(RY^{m-2})^{-1} = Y^m, \\
\vdots \\
Y^{m-1}.R.(RY)^{-1} = Y^m, \\
R.R.(I)^{-1} = I, \\
RY.R.(Y^{m-1})^{-1} = Y^{-m}, \\
RY^2.R.(Y^{m-2})^{-1} = Y^{-m}, \\
\vdots \\
RY^{m-1}.R.(Y)^{-1} = Y^{-m},
\end{array}$$

After required calculations, we have a presentation of $\overline{H}^{m}_{p,\infty}(\lambda)$ as

$$\begin{aligned}
\overline{H}^{m}_{p,\infty}(\lambda) &= < X, YXY^{-1},\ Y^2XY^{-2}, \cdots, Y^{m-1}XY^{1-m}, Y^m \mid X^p = (YXY^{-1})^p \\
&= (Y^2XY^{-2})^p = \cdots = (Y^{m-1}XY^{1-m})^p = (Y^m)^{\infty} = I > \\
&\cong \underbrace{C_p * C_p * \cdots * C_p}_{m \text{ times}} * \mathbb{Z}.
\end{aligned}$$

2) The quotient group $\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda)$ is

$$\begin{aligned}
\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda) &= < X, Y, R \mid X^p = Y^{\infty} = R^2 \\
&= (XR)^2 = (YR)^2 = X^m = Y^m = (XY)^m = \cdots = I >
\end{aligned}$$

Since $m > 0$ is an odd integer and from the relations $X^m = X^p = I$, $Y^m = (YR)^2 = I$ and $R^2 = R^m = I$, we have $X = Y = R = I$. Obviously we have $X = I$. As a result, we obtain

$$\overline{H}_{p,\infty}(\lambda)/\overline{H}^{m}_{p,\infty}(\lambda) \cong \{I\},$$

and so $\overline{H}^{m}_{p,\infty}(\lambda) = \overline{H}_{p,\infty}(\lambda)$. $\square$

3.6. Corollary. *Let $p \geq 3$ be an odd integer and let $\lambda \geq 2$. If m is an even positive integer such that $(m,p) = 1$, then $\overline{H}^m_{p,\infty}(\lambda) = H^m_{p,\infty}(\lambda)$.*

The case $(m,p) = d > 1$,except of $m = 2$ and p even, is unknown and so we can not say anything about the power subgroups $\overline{H}^m_{p,\infty}(\lambda)$, similar to $H^m_{p,\infty}(\lambda)$.

3.7. Remark. In this paper, if we take $p = 2$, then our results coincide with the ones given in [14] and [15].

References

[1] Hecke, E. *Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung*, Math. Ann. **112**, 664-699, 1936.

[2] Lehner, J. *Uniqueness of a class of Fuchsian groups,* Illinois J. Math. **19**, 308–315, 1975.

[3] Calta K. and Schmidt, T. A. *Infinitely many lattice surfaces with special pseudo-Anosov maps,* J. Mod. Dyn. **7**, No. 2, 239-254, 2013.

[4] Calta K. and Schmidt, T. A. *Continued fractions for a class of triangle groups,* J. Aust. Math. Soc. **93**, No. 1-2, 21-42, 2012.

[5] Dehornoy, P. *Geodesic flow, left-handedness, and templates,* http://arxiv.org/pdf/1112.6296.

[6] Doran, C. F., Gannon, T., Movasati, H. and Shokri, K. M. *Automorphic forms for triangle groups,* Commun. Number Theory Phys. **7** , no. 4, 689–737, 2013.

[7] Pinsky, T. *Templates for geodesic flows,* Ergodic Theory Dynam. Systems **34** no. 1, 211-235, 2014,.

[8] Tsanov, V. V. *Triangle groups, automorphic forms, and torus knots,* Enseign. Math. **59** (2), no. 1-2, 73–113, 2013.

[9] Demir, B., Koruoglu, O. and Sahin, R. *Conjugacy classes in extended generalized Hecke groups,* Rev. Un. Mat. Argentina **57**, No 1, 49-56, 2016.

[10] Kaymak, S., Demir, B., Koruoglu, O. and Sahin, R. *Commutator subgroups of generalized Hecke and extended generalized Hecke groups,* submitted for publication.

[11] Sahin, R. and Bizim, O. *Some subgroups of the extended Hecke groups $\overline{H}(\lambda_q)$*, Acta Math. Sci., Ser. B, Engl. Ed., **23**, No.4, 497-502, 2003.

[12] Sahin, R., Bizim, O. and Cangül, I. N. *Commutator subgroups of the extended Hecke groups,* Czech. Math. J., **54**, No.1, 253-259, 2004.

[13] Magnus, W., Karras, A. and Solitar, D. *Combinatorial group theory,* Dover Publications, INC. New York, 1975.

[14] Koruoğlu, Ö., Sahin, R. and Ikikardes, S. *The normal subgroup structure of the extended Hecke groups,* Bull. Braz. Math. Soc. (N.S.) **38**, no. 1, 51–65, 2007.

[15] Koruoğlu, Ö., Sahin, R., Ikikardes, S. and Cangül, I. N. *Normal subgroups of Hecke groups $H(\lambda)$*, Algebr. Represent. Theory **13**, no. 2, 219–230, 2010.

[16] Cangül, I. N. and Singerman, D. *Normal subgroups of Hecke groups and regular maps,* Math. Proc. Camb. Phil. Soc. **123**, 59-74, 1998.

[17] Cangül, I. N. and Bizim, O. *Commutator subgroups of Hecke groups,* Bull. Inst. Math. Acad. Sinica **30**, no. 4, 253–259, 2002.

[18] Cangül, I. N., Sahin, R., Ikikardes S. and Koruoğlu, Ö. *Power subgroups of some Hecke groups. II.*, Houston J. Math. **33**, no. 1, 33–42, 2007.

[19] Ikikardes, S., Koruoğlu, Ö. and Sahin, R. *Power subgroups of some Hecke groups,* Rocky Mountain J. Math. **36**, no. 2, 497–508, 2006.

[20] Lang, C. L. and Lang, M. L. *The commutator subgroup of the Hecke group G_5 is not congruence,* http://arxiv.org/abs/1401.1567.

[21] Sahin, R., Ikikardes, S. and Koruoğlu, Ö. *On the power subgroups of the extended modular group $\overline{\Gamma}$*, Tr. J. of Math. **29**, 143-151, 2004.

[22] Sahin, R., Ikikardes, S. and Koruoğlu, Ö. *Some normal subgroups of the extended Hecke groups $\overline{H}(\lambda_p)$*, Rocky Mountain J. Math. **36** , no. 3, 1033-1048, 2006.

[23] Sahin, R. Koruoğlu, Ö. and Ikikardes, S. *On the extended Hecke groups $\overline{H}(\lambda_5)$*, Algebra Colloq. **13**, no. 1, 17-23, 2006.

[24] Sarıgedik, Z., Ikikardes, S. and Sahin, R. *Power subgroups of extended Hecke groups,* Miskolc Mathematical Notes **16**, No. 1, 483-490, 2015.

[25] Yilmaz, N. and Cangül, I. N. *Power subgroups of Hecke groups* $H(\sqrt{n})$, Int. J. Math. Math. Sci. **25,** no. 11, 703-708, 2001.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1033 – 1047

On quasi-contractions in metric spaces with a graph

Kamal Fallahi* and Aris Aghanians†‡

Abstract

In the present work, we introduce G-quasi-contractions using directed graphs in metric spaces with a graph and we show that this contraction generalizes a large number of contractions. We then investigate the existence of fixed points for G-quasi-contractions under two different conditions and discuss the main theorem. Finally, we list some consequences of our theorem where either the contractive condition is replaced with a stronger one or the underlying space is changed to a complete metric space or a complete cone metric space.

Keywords: $(\widetilde{\mathrm{C}})$-Graph, $\widetilde{G}$-Quasi-contraction, Fixed point.

2000 AMS Classification: 47H10, 05C40

Received : 22.05.2015 *Accepted* : 20.10.2015 *Doi* : 10.15672/HJMS.20164514282

1. Introduction and Preliminaries

In 1974, Lj. B. Ćirić [9] introduced (single-valued) quasi-contractions in metric spaces and gave an example to show that this new contraction is a real generalization of some well-known linear contractions. He investigated the existence and uniqueness of fixed points for quasi-contractions in T-orbitally complete metric spaces via a different approach rather than using merely the iterates of a point. He also introduced multi-valued quasi-contractions and showed that a similar result is valid for these contractions in F-orbitally complete metric spaces.

In [21], B. E. Rhoades compared various definitions of contractive mappings in metric spaces and showed that Ćirić's contractive condition is one of the most general contractive definitions in metric spaces and includes a large number of different types of contractions. Thus, many authors became interested in studying quasi-contractions. The existence and uniqueness of fixed points for these contractions as well as some interesting properties

*Department of Mathematics, Payame Noor University, P.O. Box 19395-3697, Tehran, Iran, Email: fallahi1361@gmail.com

†Department of Mathematics, Payame Noor University, P.O. Box 19395-3697, Tehran, Iran, Email: a.aghanians@dena.kntu.ac.ir

‡Corresponding Author.

of them have been investigated not only in metric spaces but in different spaces such as modular spaces (see, e.g., [17]) and cone metric spaces (see, e.g., [13, 15, 16, 20]) so far. Quasi-contractions have also been studied in Banach spaces (see, e.g., [10]).

The most important graph theory approach to metric fixed point theory introduced so far is attributed to J. Jachymski [14]. In this approach, the underlying metric space is equipped with a directed graph and the Banach contraction is formulated in a graph language. Using this simple but very interesting idea, J. Jachymski generalized several well-known versions of Banach contraction principle in metric spaces simultaneously and from various aspects. As an application, he proved the Kelisky-Rivlin theorem on the iterates of the Bernstein operators defined on the Banach space of continuous functions on $[0,1]$. In the recent years, many authors followed J. Jachymski's idea to formulate different types of contractions via directed graphs in metric spaces and generalized the concerned fixed point theorems (see, e.g. [1, 2, 3, 6]).

The main goal of this paper is to formulate single-valued quasi-contractions in metric spaces with a graph and find sufficient conditions which guarantee the existence of a fixed point. A large number of different types of contractive mappings formulated using directed graphs satisfy the presented contractive condition and our main result is a natural generalization of [9, Theorem 1] from metric spaces to metric spaces with a graph.

We start by reviewing a few basic notions in graph and fixed point theory that are frequently used in the paper. For more details on graphs, the reader is refered to [4].

In an arbitrary (not necessarily simple) graph G, a link is an edge of G with distinct ends and a loop is an edge of G with identical ends. Two or more links of G with the same pairs of ends are called parallel edges of G.

Suppose that (X,d) is a metric space and G is a directed graph whose vertex set $V(G)$ coincides with X and edge set $E(G)$ contains all loops (note that in general, G can have uncountably many vertices). Suppose further that G has no parallel edges. In this case, (X,d) is called a metric space with the graph G.

By G^{-1}, it is meant the conversion of G as usual, i.e. a directed graph obtained from G by reversing the directions of the edges of G, and by $\widetilde{G}$, it is always meant the undirected graph obtained from G by ignoring the directions of the edges G. Thus, it is clear that $V(G^{-1}) = V(\widetilde{G}) = V(G) = X$ and we have

$$E(G^{-1}) = \big\{(x,y) \in X \times X : (y,x) \in E(G)\big\} \quad \text{and} \quad E(\widetilde{G}) = E(G) \cup E(G^{-1}).$$

If $(X,\preccurlyeq)$ is a partially ordered set, then by comparable elements of $(X,\preccurlyeq)$, it is meant two elements $x,y \in X$ satisfying either $x \preccurlyeq y$ or $y \preccurlyeq x$, and following A. C. M. Ran and M. C. B. Reurings [19, Theorem 2.1], a mapping $T : X \to X$ is called order-preserving whenever $x \preccurlyeq y$ implies $Tx \preccurlyeq Ty$ for all $x,y \in X$. Furthermore, following the idea of A. Petruşel and I. A. Rus in L-spaces [18, Definitions 3.1 and 3.6] (see also [23]), one can naturally formulate Picard and weakly Picard operators in metric spaces as follows:

1.1. Definition ([14, 18, 23]). Let (X,d) be a metric space and $T : X \to X$ be a mapping.

- **a)** T is called a Picard operator if T has a unique fixed point $x^\star \in X$ and $T^n x \to x^\star$ for all $x \in X$.
- **b)** T is called a weakly Picard operator if $\{T^n x\}$ is a convergent sequence and its limit (which depends on x) is a fixed point of T for all $x \in X$.

Finally, we need a weaker type of continuity defined in metric spaces with a graph which was first introduced by J. Jachymski (see [14, Definition 2.4]). The idea of this definition comes from the definition of orbital continuity defined by Lj. B. Ćirić [8].

1.2. Definition ([14]). Let (X,d) be a metric space with a graph G. A mapping $T : X \to X$ is called orbitally G-continuous on X if $T^{b_n} x \to y$ implies $T(T^{b_n} x) \to Ty$ for all $x, y \in X$ and all sequences $\{b_n\}$ of positive integers such that $(T^{b_n}x, T^{b_n+1}x) \in E(G)$ for all $n \in \mathbb{N}$.

2. Main Results

Let (X, d) be a metric space with a graph G and let $T : X \to X$ be a mapping. In this section, by C_T, we mean the set of all points $x \in X$ such that $(T^m x, T^n x)$ is an edge of $\widetilde{G}$ for all $m, n \in \mathbb{N} \cup \{0\}$, i.e.

$$C_T = \big\{x \in X : (T^m x, T^n x) \in E(\widetilde{G}) \quad m, n = 0, 1, \dots\big\}.$$

Note that C_T may be an empty set. For instance, consider the set $\mathbb{R}$ of all real numbers with the usual Euclidean metric and a graph G given by $V(G) = \mathbb{R}$ and $E(G) = \{(x, x) : x \in \mathbb{R}\}$. If $T : \mathbb{R} \to \mathbb{R}$ is defined by the rule $Tx = x + 1$ for all $x \in \mathbb{R}$, then it is easily seen that $C_T = \emptyset$.

Given $x \in X$ and $n \in \mathbb{N} \cup \{0\}$, the n-th orbit of x under T is denoted by $O(x; n)$, i.e.

$$O(x; n) = \{x, Tx, \dots, T^n x\}.$$

Finally, if A is a subset of X, then by $\operatorname{diam}(A)$, it is meant the diameter of A in X, i.e.

$$\operatorname{diam}(A) = \sup\big\{d(x, y) : x, y \in A\big\}.$$

Following the idea of S. M. A. Aleomraninejad et al. [1], we say that G is a $(\widetilde{\mathrm{C}})$-graph whenever the triple (X, d, G) has the following property:

> If $x \in X$ and $\{x_n\}$ is a sequence in (X, d, G) such that $x_n \to x$ and $(x_n, x_{n+1}) \in E(\widetilde{G})$ for all $n \in \mathbb{N}$, then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $(x_{n_k}, x) \in E(\widetilde{G})$ for all $k \in \mathbb{N}$.

Now, we are ready to give the definition of G-quasi-contractions in metric spaces with a graph which is motivated by [9, Definition 1] and [14, Definition 2.1].

2.1. Definition. Let (X, d) be a metric space with a graph G and $T : X \to X$ be a mapping. We say that T a G-quasi-contraction if

Q1) T preserves the edges of G, i.e. $(x, y) \in E(G)$ implies $(Tx, Ty) \in E(G)$ for all $x, y \in X$;

Q2) there exists a $\lambda \in [0, 1)$ such that

$$d(Tx, Ty) \le \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\}$$

for all $x, y \in X$ with $(x, y) \in E(G)$.

We also call the number λ in (Q2) a quasi-contractive constant of T.

We now give some examples of G-quasi-contractions.

2.2. Example. Suppose that (X, d) is a metric space with a graph G and $x_0 \in X$. It is easy to verify that the constant mapping $x \mapsto x_0$ is a G-quasi-contraction. So the cardinality of the set of all G-quasi-contractions defined on a metric space (X, d) with a graph G is no less than the cardinality of X.

2.3. Example. Suppose that (X, d) is a metric space and $T : X \to X$ is a quasi-contraction in the sense of Lj. B. Ćirić [9, Definition 1], i.e. there exists a $\lambda \in [0, 1)$ such that

$$d(Tx, Ty) \le \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\} \tag{2.1}$$

for all $x, y \in X$. Define a graph G_0 by $V(G_0) = X$ and $E(G_0) = X \times X$, i.e. G_0 is the complete graph whose vertex set coincides with X. Clearly, T preserves the edges of G_0 and (2.1) guarantees that T satisfies (Q2) for the complete graph G_0. Thus, T is a

G_0-quasi-contraction. Hence G_0-quasi-contractions on metric spaces with the graph G_0 are precisely the quasi-contractions on metric spaces, and so G-quasi-contractions are a generalization of quasi-contractions from metric spaces to metric spaces with a graph.

2.4. Example. Suppose that $(X, \preccurlyeq)$ is a partially ordered set and d is a metric on X. Define a graph G_1 by $V(G_1) = X$ and $E(G_1) = \{(x, y) \in X \times X : x \preccurlyeq y\}$. A mapping $T : X \to X$ preserves the edges of G_1 if and only if T is order-preserving, and T satisfies (Q2) for the graph G_1 if and only if there exists a $\lambda \in [0, 1)$ such that

$$d(Tx, Ty) \leq \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\}$$

for all comparable elements $x, y \in X$.

2.5. Example. Suppose that $(X, \preccurlyeq)$ is a partially ordered set and d is a metric on X. Define a graph G_2 by $V(G_2) = X$ and $E(G_2) = \{(x, y) \in X \times X : x \preccurlyeq y \ \vee \ y \preccurlyeq x\}$. A mapping $T : X \to X$ preserves the edges of G_2 if and only if T maps comparable elements of $(X, \preccurlyeq)$ onto comparable elements, and T satisfies (Q2) for the graph G_2 if and only if there exists a $\lambda \in [0, 1)$ such that

$$d(Tx, Ty) \leq \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\} \tag{2.2}$$

for all comparable elements $x, y \in X$. In particular, if T is a G_1-quasi-contraction, then T is a G_2-quasi-contraction. Hence G-quasi-contractions are a generalization of ordered quasi-contractions from metric spaces equipped with a partial order to metric spaces with a graph.

2.6. Example. Suppose that (X, d) is a metric space and $\varepsilon > 0$ is a fixed real number. Recall that two elements $x, y \in X$ are said to be ε-close if $d(x, y) < \varepsilon$. Define a graph G_3 by $V(G_3) = X$ and $E(G_3) = \{(x, y) \in X \times X : d(x, y) < \varepsilon\}$. A mapping $T : X \to X$ preserves the edges of G_3 if and only if T maps ε-close elements of (X, d) onto ε-close elements, and T satisfies (Q2) for the graph G_3 if and only if there exists a $\lambda \in [0, 1)$ such that

$$d(Tx, Ty) \leq \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\} \tag{2.3}$$

for all ε-close elements $x, y \in X$.

Hereafter, we assume that the graphs G_0, G_1, G_2 and G_3 are as defined in Examples 2.3, 2.4, 2.5 and 2.6, respectively.

2.7. Remark. In the definitions of $(\widetilde{\mathrm{C}})$-graph and the set C_T, let's set G the special graphs G_0, G_1, G_2 and G_3. Then we obtain the following special cases:

- The set C_T related to the complete graph G_0 coincides with X and G_0 is a $(\widetilde{\mathrm{C}})$-graph.
- If $\preccurlyeq$ is a partial order on X, then the set C_T related to the graph G_1 (and also G_2) consists of all points $x \in X$ whose every two iterates under T are comparable elements of $(X, \preccurlyeq)$. In addition, G_1 (and also G_2) is a $(\widetilde{\mathrm{C}})$-graph whenever the triple $(X, d, \preccurlyeq)$ has the following property:
  - $(*)$ If $\{x_n\}$ is a sequence in (X, d) converging to an $x \in X$ whose successive terms are pairwise comparable elements of $(X, \preccurlyeq)$, then there exists a subsequence of $\{x_n\}$ whose terms and x are comparable elements of $(X, \preccurlyeq)$.
- If $\varepsilon > 0$, then the set C_T relative to the graph G_3 consists of all points $x \in X$ whose every two iterates under T are ε-close elements of (X, d). In addition, G_3 is a $(\widetilde{\mathrm{C}})$-graph. Indeed, if $\{x_n\}$ is a sequence in (X, d) converging to an $x \in X$, then for sufficiently large indices n, say $n \geq N$, we have $d(x_n, x) < \varepsilon$. Therefore, $\{x_{n+N}\}$ is a subsequence of $\{x_n\}$ whose terms and x are ε-close elements of (X, d).

2.8. Example. Suppose that (X,d) is a metric space with a graph G and $T : X \to X$ is a Banach G-contraction in the sense of J. Jachymski [14, Definition 2.1], i.e. T preserves the edges of G and there exists an $\alpha \in (0,1)$ such that

$$d(Tx,Ty) \le \alpha d(x,y)$$

for all $x,y \in X$ with $(x,y) \in E(G)$. If $(x,y) \in E(G)$, then

$$d(Tx,Ty) \le \alpha d(x,y) \le \alpha \cdot \max\big\{d(x,y), d(x,Tx), d(y,Ty), d(x,Ty), d(y,Tx)\big\}.$$

Therefore, T satisfies (Q2) and so T is a G-quasi-contraction. Hence every G-contraction is a G-quasi-contraction.

2.9. Example. Suppose that (X,d) is a metric space with a graph G and $T : X \to X$ is a G-Kannan mapping in the sense of F. Bojor [2, Definition 4], i.e. T preserves the edges of G and there exists an $\alpha \in [0, \frac{1}{2})$ such that

$$d(Tx,Ty) \le \alpha\big(d(x,Tx) + d(y,Ty)\big)$$

for all $x,y \in X$ with $(x,y) \in E(G)$. If $(x,y) \in E(G)$, then

$$\begin{aligned} d(Tx,Ty) &\le \alpha\big(d(x,Tx) + d(y,Ty)\big) \\ &\le 2\alpha \cdot \max\big\{d(x,Tx), d(y,Ty)\big\} \\ &\le 2\alpha \cdot \max\big\{d(x,y), d(x,Tx), d(y,Ty), d(x,Ty), d(y,Tx)\big\}. \end{aligned}$$

Therefore, T satisfies (Q2) and so T is a G-quasi-contraction. Hence every G-Kannan mapping is a G-quasi-contraction.

2.10. Example. Suppose that (X,d) is a metric space with a graph G and $T : X \to X$ is a G-Chatterjea mapping in the sense that T preserves the edges of G and there exists an $\alpha \in [0, \frac{1}{2})$ such that

$$d(Tx,Ty) \le \alpha\big(d(x,Ty) + d(y,Tx)\big)$$

for all $x,y \in X$ with $(x,y) \in E(G)$ (see [5, 21] for the definition in metric spaces). If $(x,y) \in E(G)$, then an argument similar to that appeared in Example 2.9 establishes that

$$d(Tx,Ty) \le 2\alpha \cdot \max\big\{d(x,y), d(x,Tx), d(y,Ty), d(x,Ty), d(y,Tx)\big\}.$$

Therefore, T satisfies (Q2) and so T is a G-quasi-contraction. Hence every G-Chatterjea mapping is a G-quasi-contraction.

2.11. Example. Suppose that (X,d) is a metric space with a graph G and $T : X \to X$ is a G-Ćirić-Reich-Rus operator in the sense of F. Bojor [3, Definition 7], i.e. T preserves the edges of G and there exist $a,b,c \ge 0$ with $a+b+c < 1$ such that

$$d(Tx,Ty) \le ad(x,y) + bd(x,Tx) + cd(y,Ty)$$

for all $x,y \in X$ with $(x,y) \in E(G)$. If $(x,y) \in E(G)$, then an argument similar to that appeared in Example 2.9 establishes that

$$d(Tx,Ty) \le (a+b+c) \cdot \max\big\{d(x,y), d(x,Tx), d(y,Ty), d(x,Ty), d(y,Tx)\big\}.$$

Therefore, T satisfies (Q2) and so T is a G-quasi-contraction. Hence every G-Ćirić-Reich-Rus operator is a G-quasi-contraction.

Now, suppose that $T : X \to X$ is a Ćirić-Reich-Rus G-contraction in the sense of C. Chifu and G. Petruşel [6, Definition 2.2], i.e. T preserves the edges of G and there exist $\alpha, \beta, \gamma > 0$ with $\alpha + \beta + \gamma < 1$ such that

$$d(Tx,Ty) \le \alpha d(x,y) + \beta d(x,Tx) + \gamma d(y,Ty)$$

for all $x, y \in X$ with $(x, y) \in E(G)$. Then by a similar argument, one can easily see that T is a G-quasi-contraction. Hence every Ćirić-Reich-Rus G-contraction is a G-quasi-contraction.

2.12. Example. Suppose that (X, d) is a metric space and $T : X \to X$ is a λ-generalized contraction in the sense of Lj. B. Ćirić [7, Definition 2.1], i.e. for all $x, y \in X$, there exist four functions $q, r, s, t : X \times X \to [0, \infty)$ with

$$\sup\big\{q(x, y) + r(x, y) + s(x, y) + 2t(x, y) : x, y \in X \times X\big\} = \lambda < 1$$

such that

$$\begin{aligned} d(Tx, Ty) \le\ & q(x, y)d(x, y) + r(x, y)d(x, Tx) + s(x, y)d(y, Ty) \\ & + t(x, y)\big(d(x, Ty) + d(y, Tx)\big) \end{aligned}$$

for all $x, y \in X$. In 1979, B. E. Rhoades [22] studied a more general form of λ-generalized contractions (where the terms $d(x, Ty)$ and $d(y, Tx)$ have different coefficients) in sequentially complete uniform spaces via entourages and the Minkowski's pseudometrics corresponding to them. One can combine Ćirić's and Rhoades' ideas with Jachymski's idea and formulate G-λ-generalized contractions in metric spaces with a graph as follows:

> Let (X, d) be a metric space with a graph G. A mapping $T : X \to X$ is called a G-λ-generalized contraction if T preserves the edges of G and there exist five functions $a_1, a_2, a_3, a_4, a_5 : X \times X \to [0, \infty)$ with
>
> $$\sup\big\{a_1(x, y) + a_2(x, y) + a_3(x, y) + a_4(x, y) + a_5(x, y) : x, y \in X \times X\big\} = \lambda < 1 \tag{2.4}$$
>
> such that
>
> $$\begin{aligned} d(Tx, Ty) \le\ & a_1(x, y)d(x, y) + a_2(x, y)d(x, Tx) + a_3(x, y)d(y, Ty) \\ & + a_4(x, y)d(x, Ty) + a_5(x, y)d(y, Tx) \end{aligned}$$
>
> for all $x, y \in X$ with $(x, y) \in E(G)$.

Now, suppose that (X, d) is a metric space with a graph G and $T : X \to X$ is a G-λ-generalized contraction. If $(x, y) \in E(G)$, then an argument similar to that appeared in Example 2.9 establishes that

$$\begin{aligned} d(Tx, Ty) &\le \Big(\sum_{i=1}^{5} a_i(x, y)\Big) \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), \\ &\qquad\qquad d(x, Ty), d(y, Tx)\big\} \\ &\le \lambda \cdot \max\big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\}, \end{aligned}$$

where $a_1, a_2, a_3, a_4, a_5 : X \times X \to [0, \infty)$ satisfy (2.4). Therefore, T satisfies (Q2) and so T is a G-quasi-contraction. Hence every G-λ-generalized contraction (in particular, every λ-generalized contraction) is a G-quasi-contraction.

2.13. Example. Suppose that E is a nontrivial real Banach space and P is a closed cone in E such that $P \cap (-P) = \{0\}$. It is well-known that P induces a partial order $\preceq_P$ on E given by

$$a \preceq_P b \iff b - a \in P \qquad (a, b \in E).$$

Assume that $d : X \times X \to E$ is a cone metric on X and (X, d) is a cone metric space (see [12, Definition 1]). In 2010, W.-S. Du [11] showed that if the underlying cone P has nonempty interior and $\xi_e : E \to \mathbb{R}$ is the nonlinear scalarization function defined by

$$\xi_e(a) = \inf\big\{t \in \mathbb{R} : a \in te - P\big\} \qquad (a \in E),$$

where e is an interior point of P, then the function $\rho_e : X \times X \to \mathbb{R}$ given by

$$\rho_e(x, y) = \xi_e\big(d(x, y)\big) \qquad (x, y \in X) \tag{2.5}$$

defines a metric on X, and the natural (cone) topology on X induced by the cone metric d and the metric topology on X induced by the metric ρ_e coincide (see [11, Theorems 2.1 and 2.2]).

Now, suppose that $T : (X, d) \to (X, d)$ is a quasi-contraction in the sense of D. Ilić and V. Rakočević [13, Definition 1.2], i.e. there exists a $\lambda \in (0, 1)$ such that

$$d(Tx, Ty) \preceq_P \lambda \cdot u_{x,y}$$

for all $x, y \in X$ and some

$$u_{x,y} \in \big\{d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)\big\}.$$

Suppose further that the underlying cone P has nonempty interior and pick an interior point e of P. If $x, y \in X$, since ξ_e is positively homogeneous (i.e. $a \in E$ and $t \geq 0$ imply $\xi_e(ta) = t\xi_e(a)$) and nondecreasing (i.e. $a, b \in E$ and $a \preceq_P b$ imply $\xi_e(a) \leq \xi_e(b)$) on E (see [11, Lemma 1.1(v) and (vi)]), it follows that

$$\begin{aligned}
\rho_e(Tx, Ty) &= \xi_e\big(d(Tx, Ty)\big)\\
&\leq \xi_e(\lambda \cdot u_{x,y})\\
&= \lambda \cdot \xi_e(u_{x,y})\\
&\leq \lambda \cdot \max\Big\{\xi_e\big(d(x, y)\big), \xi_e\big(d(x, Tx)\big), \xi_e\big(d(y, Ty)\big),\\
&\qquad\qquad \xi_e\big(d(x, Ty)\big), \xi_e\big(d(y, Tx)\big)\Big\}\\
&= \lambda \cdot \max\big\{\rho_e(x, y), \rho_e(x, Tx), \rho_e(y, Ty), \rho_e(x, Ty), \rho_e(y, Tx)\big\}.
\end{aligned}$$

Therefore, $T : (X, \rho_e) \to (X, \rho_e)$ is also a quasi-contraction and in particular, a G_0-quasi-contraction. Hence every quasi-contraction on a cone metric space is a G_0-quasi-contraction whose domain is a suitable metric space with the complete graph G_0 provided that the underlying cone has nonempty interior.

The following proposition is an immediate consequence of the definition of G-quasi-contractions and gives a simple procedure to construct new G-quasi-contractions from older ones.

2.14. Proposition. *Let (X, d) be a metric space with a graph G and $T : X \to X$ be a mapping.*

- **a)** *If T preserves the edges of G, then T preserves the edges of G^{-1} and $\widetilde{G}$.*
- **b)** *If T satisfies* (Q2) *for the graph G, then T satisfies* (Q2) *for both the graphs G^{-1} and $\widetilde{G}$.*
- **c)** *If T is a G-quasi-contraction with a quasi-contractive constant $\lambda \in [0, 1)$, then T is both a G^{-1}-quasi-contraction and a $\widetilde{G}$-quasi-contraction with a quasi-contractive constant λ.*

To prove the existence of a fixed point for a G-quasi-contraction in a complete metric space with a graph, we need some lemmas. The first one is the graph version of [9, Lemma 1] proved by Lj. B. Ćirić and the proof appears here is very similar to Ćirić's proof. Nevertheless, for convenience of the reader, we repeat the detailed proof here.

2.15. Lemma. *Let (X, d) be a metric space with a graph G and $T : X \to X$ be a G-quasi-contraction with a quasi-contractive constant λ. Then*

$$d(T^i x, T^j x) \leq \lambda \cdot \operatorname{diam}\big(O(x; n)\big) \qquad i, j = 1, \ldots, n$$

for all $x \in C_T$ and all $n \in \mathbb{N}$.

Proof. Let $x \in C_T$ and $n \in \mathbb{N}$ be given. If i and j are arbitrary positive integers no more than n, then $(T^{i-1}x, T^{j-1}x) \in E(\widetilde{G})$. By Proposition 2.14(c), T is also a $\widetilde{G}$-quasi-contraction with a quasi-contractive constant λ. In particular, T satisfies (Q2) for the graph $\widetilde{G}$. Therefore,

$$\begin{aligned} d(T^i x, T^j x) &= d(TT^{i-1}x, TT^{j-1}x) \\ &\leq \lambda \cdot \max\big\{d(T^{i-1}x, T^{j-1}x), d(T^{i-1}x, T^i x), d(T^{j-1}x, T^j x), \\ &\qquad d(T^{i-1}x, T^j x), d(T^{j-1}x, T^i x)\big\} \\ &\leq \lambda \cdot \operatorname{diam}\big(O(x;n)\big). \end{aligned}$$ □

The next example shows that both the integers i and j must be positive in Lemma 2.15. In other words, neither i nor j is allowed to be zero.

2.16. Example. Consider the set $\mathbb{R}$ of real numbers with the usual (Euclidean) metric and the complete graph G_0, and define a mapping $T : \mathbb{R} \to \mathbb{R}$ by the rule $Tx = \frac{x}{2}$ for all $x \in \mathbb{R}$. Then T is a G_0-quasi-contraction with a quasi-contractive constant $\lambda = \frac{1}{2}$. In addition, $T^n x = \frac{x}{2^n}$ and $\operatorname{diam}(O(x;n)) = |x|(1 - \frac{1}{2^n})$ for all $x \in \mathbb{R}$ and all $n \in \mathbb{N} \cup \{0\}$. Now, let x_0 be a positive real number and put $n = 2$, $i = 0$ and $j = 1$ in Lemma 2.15. Then we have

$$|x_0 - Tx_0| = \frac{x_0}{2} > \frac{x_0}{2} \cdot \big(1 - \frac{1}{2^2}\big) = \lambda \cdot \operatorname{diam}\big(O(x_0;2)\big).$$

2.17. Lemma. *Let (X,d) be a metric space with a graph G and $T : X \to X$ be a G-quasi-contraction. Then for each $x \in C_T$ and each $n \in \mathbb{N}$, there exists a positive integer k no more than n such that*

$$\operatorname{diam}\big(O(x;n)\big) = d(x, T^k x).$$

Proof. Let $x \in C_T$ and $n \in \mathbb{N}$ be given. If $\operatorname{diam}(O(x;n)) = 0$, then $O(x;n)$ is singleton. In particular, x is a fixed point for T and $d(T^i x, T^j x) = 0$ for all $i, j = 0, \ldots, n$. Thus, the statement holds trivially for any positive integer k no more than n.

Otherwise, since $O(x;n)$ is a finite set, it follows that there exist distinct nonnegative integers i and j no more that n such that $\operatorname{diam}(O(x;n)) = d(T^i x, T^j x)$. If both the integers i and j are assumed to be positive, then from Lemma 2.15, we have

$$\operatorname{diam}\big(O(x;n)\big) = d(T^i x, T^j x) \leq \lambda \cdot \operatorname{diam}\big(O(x;n)\big),$$

where $\lambda \in [0,1)$ is a quasi-contractive constant of T, a contradiction. Hence either i or j must be zero and the proof is finished. □

2.18. Remark. Combining Lemmas 2.15 and 2.17, one can easily obtain that if (X,d) is a metric space with a graph G and $T : X \to X$ is a G-quasi-contraction with a quasi-contractive constant λ, then for each $x \in C_T$ and each $n \in \mathbb{N}$, there exists a positive integer k no more than n such that

$$d(T^i x, T^j x) \leq \lambda \cdot \operatorname{diam}\big(O(x;n)\big) = \lambda \cdot d(x, T^k x) \qquad i, j = 1, \ldots, n.$$

2.19. Lemma. *Let (X,d) be a metric space with a graph G and $T : X \to X$ be a G-quasi-contraction with a quasi-contractive constant λ. Then*

$$\operatorname{diam}\big(O(x;n)\big) \leq \frac{1}{1-\lambda} \cdot d(x, Tx)$$

for all $x \in C_T$ and all $n \in \mathbb{N} \cup \{0\}$.

Proof. Let $x \in C_T$ and $n \in \mathbb{N} \cup \{0\}$ be given. If $n = 0$, since $\operatorname{diam}(O(x;0)) = 0$, there remains nothing to prove. Otherwise, from Lemma 2.17, there exists a positive integer k no more than n such that $\operatorname{diam}(O(x;n)) = d(x, T^k x)$. Putting $i = 1$ and $j = k$ in Lemma 2.15, we get

$$\begin{aligned}\operatorname{diam}\big(O(x;n)\big) &= d(x, T^k x)\\ &\leq d(x, Tx) + d(Tx, T^k x)\\ &\leq d(x, Tx) + \lambda \cdot \operatorname{diam}\big(O(x;n)\big).\end{aligned}$$

Now the inequality

$$\operatorname{diam}\big(O(x;n)\big) \leq \frac{1}{1-\lambda} \cdot d(x, Tx)$$

follows immediately. □

2.20. Lemma. *Let (X, d) be a metric space with a graph G and $T : X \to X$ be a G-quasi-contraction. Then $\{T^n x\}$ is Cauchy for all $x \in C_T$.*

Proof. Let $x \in C_T$ be given. If $m, n \in \mathbb{N}$ and $m \geq n \geq 2$, since $T^{n-1}x \in C_T$, it follows that putting $i = m - n + 1$ and $j = 1$ in Lemma 2.15, we get

$$d(T^m x, T^n x) = d(T^{m-n+1}T^{n-1}x, TT^{n-1}x) \leq \lambda \cdot \operatorname{diam}\big(O(T^{n-1}x; m-n+1)\big), \tag{2.6}$$

where $\lambda \in [0, 1)$ is a quasi-contractive constant of T. Moreover, by Lemma 2.17, there exists a positive integer k no more than $m - n + 1$ such that

$$\operatorname{diam}\big(O(T^{n-1}x; m-n+1)\big) = d(T^{n-1}x, T^{k+n-1}x). \tag{2.7}$$

Because $n \geq 2$, it follows that $T^{n-2}x \in C_T$ and so putting $i = 1$ and $j = k+1$ in Lemma 2.15 yields

$$\begin{aligned}d(T^{n-1}x, T^{k+n-1}x) &= d(TT^{n-2}x, T^{k+1}T^{n-2}x)\\ &\leq \lambda \cdot \operatorname{diam}\big(O(T^{n-2}x; m-n+2)\big). \end{aligned}\tag{2.8}$$

Finally, combining (2.6), (2.7) and (2.8), and using induction and Lemma 2.19, we obtain

$$\begin{aligned}d(T^m x, T^n x) &\leq \lambda \cdot \operatorname{diam}\big(O(T^{n-1}x; m-n+1)\big)\\ &= \lambda \cdot d(T^{n-1}x, T^{k+n-1}x)\\ &\leq \lambda^2 \cdot \operatorname{diam}\big(O(T^{n-2}x; m-n+2)\big)\\ &\;\vdots\\ &\leq \lambda^n \cdot \operatorname{diam}\big(O(x;m)\big)\\ &\leq \frac{\lambda^n}{1-\lambda} \cdot d(x, Tx).\end{aligned}$$

Letting $m, n \to \infty$, we find $d(T^m x, T^n x) \to 0$. Hence $\{T^n x\}$ is Cauchy. □

Now we are ready to prove our main theorem on the existence of fixed points for G-quasi-contractions in complete metric spaces with a graph.

2.21. Theorem. *Let (X, d) be a complete metric space with a graph G and $T : X \to X$ be G-quasi-contraction. Then the restriction of T to C_T is a weakly Picard operator if either T is orbitally $\widetilde{G}$-continuous on X or G is a $(\widetilde{\mathrm{C}})$-graph.*

In particular, whenever T is orbitally $\widetilde{G}$-continuous on X or G is a $(\widetilde{\mathrm{C}})$-graph, T has a fixed point in X if and only if $C_T \neq \emptyset$.

Proof. If $C_T = \emptyset$, then there remains nothing to prove. So assume that C_T is nonempty. If $x \in C_T$, since $(T^m x, T^n x) \in E(\widetilde{G})$ for all $m, n \in \mathbb{N} \cup \{0\}$, it follows that $Tx \in C_T$. Thus, C_T is T-invariant, i.e. $T(C_T) \subseteq C_T$.

Now, let $x \in C_T$ be given. By Lemma 2.20, $\{T^n x\}$ is a Cauchy sequence in X and since (X, d) is complete, there exists an $x^\star \in X$ (depending on x) such that $T^n x \to x^\star$. We show that $x^\star$ is a fixed point for T.

To this end, note first that from $x \in C_T$, we have $(T^n x, T^{n+1} x) \in E(\widetilde{G})$ for all $n \in \mathbb{N} \cup \{0\}$. If T is orbitally $\widetilde{G}$-continuous on X, then $T^n x \to x^\star$ implies $T^{n+1} x = T(T^n x) \to Tx^\star$ and by uniqueness of the limit of convergent sequences in metric spaces, we obtain $Tx^\star = x^\star$.

Otherwise, if G is a $(\widetilde{\mathrm{C}})$-graph, since $T^n x \to x^\star$, there exists a strictly increasing sequence $\{n_k\}$ of positive integers such that $(T^{n_k} x, x^\star) \in E(\widetilde{G})$ for all $k \in \mathbb{N}$. On the other hand, if $\lambda \in [0, 1)$ is a quasi-contractive constant of T, then by Proposition 2.14(c), T is a $\widetilde{G}$-quasi-contraction with a quasi-contractive constant λ. In particular, T satisfies (Q2) for the graph $\widetilde{G}$. Therefore,

$$
\begin{aligned}
d(T^{n_k+1} x, Tx^\star) &= d(TT^{n_k} x, Tx^\star) \\
&\leq \lambda \cdot \max\big\{d(T^{n_k} x, x^\star), d(T^{n_k} x, T^{n_k+1} x), d(x^\star, Tx^\star), \\
&\qquad\qquad d(T^{n_k} x, Tx^\star), d(x^\star, T^{n_k+1} x)\big\}
\end{aligned}
\tag{2.9}
$$

for all $k \in \mathbb{N}$. For a fixed positive integer k, one of the five terms appeared in the right side of (2.9) is the maximum. So we consider the following five possible cases:

Case 1: If the first term is the maximum, then

$$d(T^{n_k+1} x, Tx^\star) \leq \lambda \cdot d(T^{n_k} x, x^\star);$$

Case 2: If the second term is the maximum, then

$$d(T^{n_k+1} x, Tx^\star) \leq \lambda \cdot d(T^{n_k} x, T^{n_k+1} x);$$

Case 3: If the third term is the maximum, then

$$
\begin{aligned}
d(T^{n_k+1} x, Tx^\star) &\leq \lambda \cdot d(x^\star, Tx^\star) \\
&\leq \lambda \cdot \big(d(x^\star, T^{n_k+1} x) + d(T^{n_k+1} x, Tx^\star)\big).
\end{aligned}
$$

Therefore,

$$d(T^{n_k+1} x, Tx^\star) \leq \frac{\lambda}{1-\lambda} \cdot d(x^\star, T^{n_k+1} x) = \frac{\lambda}{1-\lambda} \cdot d(T^{n_k+1} x, x^\star);$$

Case 4: If the forth term is the maximum, then

$$
\begin{aligned}
d(T^{n_k+1} x, Tx^\star) &\leq \lambda \cdot d(T^{n_k} x, Tx^\star) \\
&\leq \lambda \cdot \big(d(T^{n_k} x, T^{n_k+1} x) + d(T^{n_k+1} x, Tx^\star)\big).
\end{aligned}
$$

Therefore,

$$d(T^{n_k+1} x, Tx^\star) \leq \frac{\lambda}{1-\lambda} \cdot d(T^{n_k} x, T^{n_k+1} x);$$

Case 5: Finally, if the fifth term is the maximum, then

$$d(T^{n_k+1} x, Tx^\star) \leq \lambda \cdot d(x^\star, T^{n_k+1} x) = \lambda \cdot d(T^{n_k+1} x, x^\star).$$

Clearly, at least one of the above five cases happens for infinitely many indices k. Hence $\{T^{n_k+1} x\}$ has a subsequence converging to $Tx^\star$, and again by the uniqueness of the limit of convergent sequences in metric spaces, we obtain $Tx^\star = x^\star$.

Finally, since C_T contains all fixed points of T, it follows that $x^\star \in C_T$. Consequently, $T|_{C_T}\colon C_T \to C_T$ is a weakly Picard operator. $\square$

Before listing some important consequences of Theorem 2.21, it is worth having a discussion on the hypotheses of Theorem 2.21.

2.22. Remark. In [9, Theorem 1], Lj. B. Ćirić has used a weaker type of completeness of metric spaces which had been defined by himself in [8] as follows:

> Let (X, d) be a metric space and $T : X \to X$ be a mapping. The metric space (X, d) is called T-orbitally complete if each Cauchy sequence of the iterates of a point of X under T is convergent.

It is clear that every complete metric space (X, d) is T-orbitally complete for all mappings $T : X \to X$, but the converse is not true in general. For instance, the set $\mathbb{Q}$ consisting of all rational numbers with the usual (Euclidean) metric is not a complete metric space whereas $\mathbb{Q}$ is T-orbitally complete, where $T : \mathbb{Q} \to \mathbb{Q}$ is defined by the rule $Tx = \frac{x}{2}$ for all $x \in \mathbb{Q}$.

The notion of T-orbital completeness of a metric space can be generalized to metric spaces with a graph in several different ways. However, by a subtle look at the proof of Theorem 2.21, it is easily realized that we have only used the following weaker type of T-orbital completeness (called, e.g., weak $\widetilde{G}$-T-orbital completeness) in metric spaces with a graph as follows:

> Let (X, d) be a metric space with a graph G and $T : X \to X$ be a mapping. The metric space (X, d) is called "weak $\widetilde{G}$-T-orbitally complete" if for each $x \in C_T$, the sequence $\{T^n x\}$ is convergent whenever $\{T^n x\}$ is Cauchy and satisfies $(T^n x, T^{n+1} x) \in E(\widetilde{G})$ for all $n \in \mathbb{N}$.

Obviously, by replacing this new notion with the standard notion of completeness, a new version of Theorem 2.21 is obtained.

2.23. Remark. By a subtle look at the proof of Theorem 2.21 in the case that the mapping T is orbitally $\widetilde{G}$-continuous on X, it is easily realized that not the whole but a weaker type of the hypothesis of orbital $\widetilde{G}$-continuity of T is used. Indeed, the sequence $\{b_n\}$ of positive integers in Definition 1.2 is replaced with the sequence $\{n\}$, i.e. the sequence of all positive integers. Using this, a weaker type of orbital $\widetilde{G}$-continuity (called, e.g., weak orbital $\widetilde{G}$-continuity) can be defined as follows:

> Let (X, d) be a metric space with a graph G. A mapping $T : X \to X$ is called "weakly orbitally $\widetilde{G}$-continuous" on X if $T^n x \to y$ implies $T^{n+1} x \to Ty$ for all $x, y \in X$ such that $(T^n x, T^{n+1} x) \in E(\widetilde{G})$ for all $n \in \mathbb{N}$.

Obviously, by replacing this new notion with the notion of orbital $\widetilde{G}$-continuity, Theorem 2.21 is strengthened.

Now we present three important consequences of Theorem 2.21 where the graph G is replaced with the special graphs. Firstly, we put $G = G_0$ in Theorem 2.21 and we get Ćirić's fixed point theorem [9, Theorem 1] on single-valued quasi-contractions in complete metric spaces instead of T-orbitally complete metric spaces as follows:

2.24. Corollary. *Every quasi-contraction defined on a complete metric space is a Picard operator.*

Proof. Let (X, d) be a complete metric space and $T : X \to X$ be a quasi-contraction. The set C_T is nonempty because $C_T = X$. Therefore, by Theorem 2.21, the mapping $T = T\,|_{C_T}$ is a weakly Picard operator. In particular, T has a fixed point in X. To see that T is a Picard operator, it sufficies to show that T has a unique fixed point in X. To this end, suppose that $x^\star$ and $x^{\star\star}$ are two fixed points for T in X. Then from (2.1) we have

$$\begin{aligned} d(x^\star, x^{\star\star}) &= d(Tx^\star, Tx^{\star\star}) \\ &\le \lambda \cdot \max\Big\{ d(x^\star, x^{\star\star}), \underbrace{d(x^\star, Tx^\star)}_{=0}, \underbrace{d(x^{\star\star}, Tx^{\star\star})}_{=0}, \\ &\qquad\qquad \underbrace{d(x^\star, Tx^{\star\star})}_{=d(x^\star, x^{\star\star})}, \underbrace{d(x^{\star\star}, Tx^\star)}_{=d(x^{\star\star}, x^\star)} \Big\} \\ &= \lambda \cdot d(x^\star, x^{\star\star}), \end{aligned}$$

where $\lambda \in [0, 1)$ is a constant. Hence $d(x^\star, x^{\star\star}) = 0$ or equivalently, $x^\star = x^{\star\star}$. □

2.25. Remark. By a subtle look at the proof of Corollary 2.24, and use an argument similar to that appeared there, we see that both the ends of any link of G cannot be fixed points for a G-quasi-contraction, i.e. if $x \neq y$, $Tx = x$ and $Ty = y$, then $(x, y) \notin E(G)$. Roughly speaking, no G-quasi-contraction can keep both the ends of a link of G fixed. In particular, the following results on the number of the fixed points of G-quasi-contractions are obtained:

- No quasi-contraction can have two distinct fixed points.
- If $\preccurlyeq$ is a partial order on X, then neither a G_1-quasi-contraction nor a G_2-quasi-contraction can have two distinct fixed points which are comparable elements of $(X, \preccurlyeq)$.
- If $\varepsilon > 0$, then no G_3-quasi-contraction can have two distinct fixed points which are ε-close elements of (X, d).

Secondly, we consider a partial order on the metric space (X, d) and put $G = G_1$ or $G = G_2$ in Theorem 2.21. Having done this, the following partially ordered version of Ćirić's fixed point theorem on ordered quasi-contractions in complete metric spaces equipped with a partial order is obtained:

2.26. Corollary. *Let $(X, \preccurlyeq)$ be a partially ordered set and d be a metric on X such that (X, d) is a complete metric space. Let $T : X \to X$ be a mapping which maps comparable elements of $(X, \preccurlyeq)$ onto comparable elements and satisfies* (2.2)*. Then the restriction of T to the set of all points $x \in X$ whose every two iterates under T are comparable elements of $(X, \preccurlyeq)$ is a weakly Picard operator if either T is orbitally G_2-continuous on X or the triple $(X, d, \preccurlyeq)$ satisfies $(*)$.*

In particular, whenever T is orbitally G_2-continuous on X or the triple $(X, d, \preccurlyeq)$ satisfies $(*)$, T has a fixed point in X if and only if there exists an $x \in X$ such than $T^m x$ and $T^n x$ are comparable elements of $(X, \preccurlyeq)$ for all $m, n \in \mathbb{N} \cup \{0\}$.

Finally, we put $G = G_3$ in Theorem 2.21 and we get the following version of Ćirić's fixed point theorem on quasi-contractions in complete metric spaces:

2.27. Corollary. *Let (X, d) be a complete metric space and $\varepsilon > 0$ be a fixed real number. Let $T : X \to X$ be a mapping which maps ε-close elements of (X, d) onto ε-close elements and satisfies* (2.3)*. Then the restriction of T to the set of all points $x \in X$ whose every two iterates under T are ε-close elements of (X, d) is a weakly Picard operator.*

In particular, T has a fixed point in X if and only if there exists an $x \in X$ such that $T^m x$ and $T^n x$ are ε-close elements of (X, d) for all $m, n \in \mathbb{N} \cup \{0\}$.

Since Banach G-contractions, G-Kannan mappings, G-Chatterjea mappings, G-Ćirić-Reich-Rus operators, Ćirić-Reich-Rus G-contractions and G-λ-generalized contractions are all a G-quasi-contraction, we have also the following fixed point theorem for these contractions as a consequence of Theorem 2.21:

2.28. Corollary. *Let (X, d) be a complete metric space with a graph G and $T : X \to X$ be a Banach G-contraction (a G-Kannan mapping, a G-Chatterjea mapping, a G-Ćirić-Reich-Rus operator, a Ćirić-Reich-Rus G-contraction, or a G-λ-generalized contraction). Then the restriction of T to C_T is a weakly Picard operator if either T is orbitally $\widetilde{G}$-continuous on X or G is a $(\widetilde{\mathrm{C}})$-graph.*

In particular, whenever T is orbitally $\widetilde{G}$-continuous on X or G is a $(\widetilde{\mathrm{C}})$-graph, T has a fixed point in X if and only if $C_T \neq \emptyset$.

By comparing Corollary 2.28 as a version of Theorem 2.21 for several types of contractions with some recent results in graph metric fixed point theory, one can get the followings:

- If we employ Corollary 2.28 for Banach G-contractions, then we obtain a simple and weaker version of [14, Theorems 3.2(4^o) and 3.3(2^o)] and [3, Corollary 2];
- If we employ Corollary 2.28 for G-Kannan mappings, then we obtain another version of [2, Theorem 3] and [3, Corollary 3] without imposing the assumption of weak T-connectedness on the graph (see [3, Definition 8]);
- If we employ Corollary 2.28 for G-Chatterjea mappings, then we obtain a new version of Chatterjea's fixed point theorem [5] in complete metric spaces with a graph;
- If we employ Corollary 2.28 for either G-Ćirić-Reich-Rus operators or Ćirić-Reich-Rus G-contractions, then we obtain another version of [3, Theorem 6] without imposing the assumption of weak T-connectedness on the graph and another version of [6, Theorem 2.2 and Lemma 2.7];
- Finally, if we employ Corollary 2.28 for G-λ-generalized contractions, then we obtain a new version of [7, Theorem 2.5] and a weaker version of [22, Theorem 1] in complete metric spaces with a graph.

Because convergence of sequences in a cone metric space has already been defined in [12, Definition 2], Picard operators can be generalized naturally from metric to cone metric spaces in the following way:

> Let E be a nontrivial real Banach space, P be a closed cone in E such that $P \cap (-P) = \{0\}$, and (X, d) be a cone metric space. A mapping $T : X \to X$ is called a Picard operator if T has unique fixed point $x^\star \in X$ and $T^n x \to x^\star$ for all $x \in X$.

Similar to the Cauchy property of sequences in metric spaces and using the idea of formulating convergent sequences in cone metric spaces, the Cauchy property of sequences is defined in cone metric spaces (see [12, Definition 3]). So it is natural to say that a cone metric space is complete if every Cauchy sequence is convergent (see [12, Definition 4]). Hence we have also the following consequence of Corollary 2.24 in complete cone metric spaces where the underlying cone has nonempty interior. This result is another version of [20, Theorem 2.1] and generalizes [12, Theorem 1], [13, Theorem 2.1] and [16, Theorems 2.2 and 2.3].

2.29. Corollary. *Every quasi-contraction defined on a complete cone metric space is a Picard operator provided that the underlying cone has nonempty interior.*

Proof. Let E be a nontrivial real Banach space, P be a closed cone in E with nonempty interior such that $P \cap (-P) = \{0\}$, and (X, d) be a complete cone metric space. Pick any interior point e of P and consider the metric ρ_e given by (2.5). Since the cone metric space (X, d) is complete, it follows from [11, Theorem 2.2(iii)] that the metric space (X, ρ_e) is also complete.

Now, let $T : (X, d) \to (X, d)$ be a quasi-contraction. As it was shown in Example 2.13, $T : (X, \rho_e) \to (X, \rho_e)$ is also a quasi-contraction. Therefore, by Corollary 2.24,

$T : (X, \rho_e) \to (X, \rho_e)$ is a Picard operator, i.e. T has a unique fixed point $x^\star \in X$ and $T^n x \to x^\star$ in (X, ρ_e) for all $x \in X$.

On the other hand, it follows from [11, Theorem 2.2(i)] that a sequence $\{x_n\}$ consisting of points of X converges to an $x \in X$ in the cone metric space (X, d) if and only if $\{x_n\}$ converges to the same point x in the metric space (X, ρ_e). Hence $T^n x \to x^\star$ in (X, d) for all $x \in X$. Consequently, $T : (X, d) \to (X, d)$ is a Picard operator. $\square$

Acknowledgments

The authors are thankful to the Payame Noor University for supporting this research and also to the referee for his/her valuable comments to improve the paper.

References

[1] Aleomraninejad, S.M.A. Rezapour, Sh. and Shahzad, N. *Some fixed point results on a metric space with a graph*, Topology Appl. **159** (3), 659-663, 2012.

[2] Bojor, F. *Fixed points of Kannan mappings in metric spaces endowed with a graph*, An. Ştiinţ. Univ. "Ovidius" Constanţa Ser. Mat. **20** (1), 31-40, 2012.

[3] Bojor, F. *Fixed point theorems for Reich type contractions on a metric spaces with a graph*, Nonlinear Anal. **75** (9), 3895-3901, 2012.

[4] Bondy, J.A. and Murty, U.S.R. *Graph theory*, (Springer, 2008).

[5] Chatterjea, S.K. *Fixed-point theorems*, C. R. Acad. Bulgare Sci. **25**, 727–730, 1972.

[6] Chifu, C. and Petruşel, G. *Generalized contractions in metric spaces endowed with a graph*, Fixed Point Theory Appl. **2012**:161, 9 pages, 2012.

[7] Ćirić, Lj.B. *Generalized contractions and fixed-point theorems*, Publ. Inst. Math. (Beograd) (N.S.) **12(26)**, 19-26, 1971.

[8] Ćirić, Lj.B. *On contraction type mappings*, Math. Balkanica **1**, 52-57, 1971.

[9] Ćirić, Lj.B. *A generalization of Banach's contraction principle*, Proc. Amer. Math. Soc. **45** 267-273, 1974.

[10] Ćirić, Lj.B. *Quasi-contractions in Banach spaces*, Publ. Inst. Math. (Beograd) (N.S.) **21(35)**, 41-48, 1977.

[11] Du, W.-S. *A note on cone metric fixed point theory and its equivalence*, Nonlinear Anal. **72** (5), 2259-2261, 2010.

[12] Long-Guang, H. and Xian, Z. *Cone metric spaces and fixed point theorems of contractive mappings*, J. Nonlinear Anal. Appl. **332** (2), 1468-1476, 2007.

[13] Ilić, D. and Rakočević, V. *Quasi-contraction on a cone metric space*, Appl. Math. Lett. **22** (5), 728-731, 2009.

[14] Jachymski, J. *The contraction principle for mappings on a metric space with a graph*, Proc. Amer. Math. Soc. **136** (4), 1359-1373, 2008.

[15] Kadelburg, Z. Pavlović, M. and Radenović, S. *Common fixed point theorems for ordered contractions and quasicontractions in ordered cone metric spaces*, Comput. Math. Appl. **59** (9), 3148-3159, 2010.

[16] Kadelburg, Z. Radenović, S. and Rakočević, V. *Remarks on "Quasi-contraction on a cone metric space"*, Appl. Math. Lett. **22** (11), 1674-1679, 2009.

[17] Khamsi, M.A. *Quasicontraction mappings in modular spaces without Δ_2-condition*, Fixed Point Theory Appl. **2012**, Art. ID 916187, 9 pages.

[18] Petruşel, A. and Rus, I.A. *Fixed point theorems in ordered L-spaces*, Proc. Amer. Math. Soc. **134** (2), 411-418, 2006.

[19] Ran, A.C.M. and Reurings, M.C.B. *A fixed point theorem in partially ordered sets and some applications to matrix equations*, Proc. Amer. Math. Soc. **132** (5), 1435-1443, 2004.

[20] Rezapour, Sh. Haghi, R.H. and Shahzad, N. *Some notes on fixed points of quasi-contraction maps*, Appl. Math. Lett. **23** (4), 498-502, 2010.

[21] Rhoades, B.E. *A comparison of various definitions of contractive mappings*, Trans. Amer. Math. Soc. **226**, 257-290, 1977.

[22] Rhoades, B.E. *Fixed point theorems in a uniform space*, Publ. Inst. Math. (Beograd) (N.S.) **25(39)**, 153-156, 1979.

[23] Rus, I.A. *Picard operators and applications*, Sci. Math. Jpn. **58** (1), 191-219, 2003.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1049 – 1060

Application of modified optimal homotopy perturbation method to higher order boundary value problems in a finite domain

Fazal Ghani*, Saeed Islam †, Cenap Ozel, ‡ Liaqat Ali§, M. M. Rashidi ¶, Tayebeh Hajjari ‖

Abstract

This research focuses on the solution of higher order boundary value problems by our proposed method "Modified Optimal Homotopy Perturbation Method" (MOHPM). A homotopy with an embedding parameter and Daftardar-Jafari polynomials are used. To control the convergence of solution, some auxiliary functions which depend upon variables and some constants are used. The proposed method is simple, rapid, effective and accurate. The accuracy has been proved by comparing our results with the solutions of optimal homotopy perturbation method (OHPM), optimal homotopy asymptotic method (OHAM), variational iterative method (VIM),variational iteration method using He's polynomials (VIMHP), homotopy perturbation method (HPM), Adoman decomposition method (ADM) and Quintic B-Spline Galerkin's scheme.

Keywords: Boundary value problems, Modified Optimal Homotopy Perturbation Method, Daftardar-Jafari Polynomials

Received : 19.02.2015 *Accepted* : 27.07.2015 *Doi* : 10.15672/HJMS.20164513109

*Abdul Wali Khan University Mardan, pakistan

†Department of mathematics, Abdull Wali Khan University, Mardan, Pakistan.

‡Department of Mathematics, Dokuz Eylul University, Buca/Izmir,35160, Turkey, and KAU King Abdulaziz University Jeddah 21589 S. Arabia Email: cenap.ozel@gmail.com

§Abdul Wali Khan University Mardan, Pakistan & Department of Electrical Engineering, CECOS University Peshawer KPK, Pakistan

¶Shanghai Automotive Wind Tunnel Center, Tongji Clean Energy Institute of advanced Study, China

‖Department of Mathematics, Firoozkooh Branch, Islamic Azad University, Firoozkooh, Iran

1. INTRODUCTION

Nonlinear boundary value problems have a significant contribution in today's modern fields of science and technology. They take place from steady state solutions of transient problems. The significance of higher order boundary value problems (BVPS) can be judge from their extensive use in mathematical modeling of different entities such as visco-elastic flows, hydrodynamic stability problems, non-Newtonian fluids and convection of heat etc [1]. In general, an nth order BVP can be represented as:

$$w^{(p)} = \xi(w,\, w',\, \ldots w^{(p-1)}) + \Theta(s), \;\; f < s < h,$$

having boundary conditions

$w^{(q)}(f) = \eta_i$ and $w^{(q)}(h) = \lambda_i$,

where$(q < p)$is a non-negative integer, $\eta_i$and λ_i are real finite constants and $\Theta(s)$ is a continuous function on$[f,\, h\,]$.

Finding a solution for the above type differential equations is a tedious job. One may find an exact solution, but if the degree of non linearity is high, it becomes impossible to get an exact solution. Researchers have therefore focused on analytic solutions of such type problems. In literature we come across different analytic methods. Some of the interesting analytic methods which can be applied to a wide range of high order differential equations are homotopy analysis method (HAM) [2-4], homotopy perturbation method (HPM) [5-6], Adomian decomposition method (ADM) [7], optimal homotopy asymptotic method (OHAM) [8-12], optimal homotopy perturbation method (OHPM) [13-15] and variational iteration method (VIM) [16-18]. In order to obtain best approximate solution of differential equations, researchers in the field modify the existing analytical methods time to time. One such modification of (HPM) has been done by V. Marinca et al [13-15].The basis for their new method, which is known as (OHPM) is He's homotopy perturbation method. This method is developed on the same lines as was done earlier in He's homotopy perturbation method. A visible change in (OHPM) is that the non linear function is extended in series form for the parameter involved and auxiliary functions are inducted within the coefficients of this truncated series. In (OHPM) the auxiliary functions have unknown parameters which can be determined optimally. All these techniques give (OHPM) an edge over the conventional (HPM).

Our purpose in this paper is to obtain a new version of OHPM, which produces more accurate and reliable results than OHPM. The target is achieved here by introducing Daftardar-Jafferi polynomials in OHPM. The modified version of OHPM thus obtained will have its name as modified optimal homotopy perturbation method (MOHPM). It is important to note that these polynomials were defined in Daftatdar-Jafferi Method (DJM)[19], and basically are the non linear terms of the Taylor's series. S. Bhalekar et al. [20] have proved the convergence of these polynomials. It can be observed from the solved problems in section 3 that MOHPM is a powerful method as it converts a complex problem into a simpler one, which can then be solved easily. This method has great potential to solve ordinary differential equations of any order. The same technique can also be extended to solve partial differential equations, Integro-differential equations and system of differential equations of physical phenomenon. In our coming papers we will be showing the application of MOHPM to these types of problems. Here we have solved some linear and non linear higher order BVPS by MOHPM and OHPM to confirm the difference in obtained solutions. The results of MOHPM are also compared with those of exact solutions and the errors are compared with the already existing well-known results of OHAM, VIM, HPM, VIMHP, ADM and B-Spline. Numerical results show that MOHPM is found best in giving better and more accurate results.

This manuscript is arranged as follows: In part 2, introduction of MOHPM is given. Part 3 is devoted to the application of MOHPM to higher order BVPS. In the same section results of numerical simulation using Mathematica 7.0 are given. In next and final section a concluding remarks are given for the obtained results.

2. INTRODUCTION OF MOHPM

Consider the problem

$$\xi\left(\kappa(s)\right)+\zeta\left(\kappa(s)\right)-\Gamma(s)=0,\quad s\in\Omega \tag{2.1}$$

$$\Delta\left(\kappa,\frac{\partial\kappa}{\partial s}\right)=0, s\in\Pi, \tag{2.2}$$

where ξ and ζ are linear and nonlinear operators respectively, Δ is the boundary operator,Π is the boundary of the domainΩ, Γ is the analytic function and the differential along the normal drawn outwards from Ω is represented by $\frac{\partial}{\partial s}$. According to homotopy scheme, we create a homotopy, $\tilde{\kappa}(s,\,\vartheta):\,\Omega\times[0,1]\to R$ by

$$\begin{aligned}\Sigma\left(\tilde{\kappa},\vartheta\right) &= (1-\vartheta)\left(\xi(\tilde{\kappa}(s,\,\vartheta))-\xi(\kappa_{ini}(s,\vartheta))\right)\\ &\quad+\vartheta\left(\xi(\tilde{\kappa}(s,\vartheta))+\zeta(\tilde{\kappa}(s,\vartheta))-\Gamma(s)\right)=0,\end{aligned} \tag{2.3}$$

where $\vartheta\in[0,1]$ is known as the embedding parameter and the initial guess for the solution of (2.1) by $\kappa_{ini}(s,\vartheta)$, which satisfies the boundary conditions. It is quite easy to note that, when $\vartheta=0$ and $\vartheta=1$ equation (2.3) holds and takes the form respectively as

$$\Sigma\left(\tilde{\kappa},\,0\right)=\xi(\tilde{\kappa}(s,\,0))-\xi(\kappa_{ini}(s,\,0))=0, \tag{2.4}$$

$$\Sigma\left(\tilde{\kappa},\,1\right)=\xi(\tilde{\kappa}(s,\,1))+\zeta(\tilde{\kappa}(s,\,1))-\Gamma(s)=0, \tag{2.5}$$

thus change in ϑ from zero to one, will change the trivial solution for (2.4) to the solution of (2.5) continuously. That is, if ϑ changes from zero to one then $\tilde{\kappa}$ changes from κ_{ini}. to κ, this is known as deformation in topology. The paths $\xi(\tilde{\kappa}(s,\,0))-\xi(\kappa_{ini}(s,\,0))$ and $\xi(\tilde{\kappa}(s,\,1))+\zeta(\tilde{\kappa}(s,\,1))-\Gamma(s)$ are homotopic to each other. At this stage assume the perturbation series

$$\tilde{\kappa}(s)=\tilde{\kappa}_0(s)+\vartheta\,\tilde{\kappa}_1(s)+\vartheta^2\,\tilde{\kappa}_2+\cdots. \tag{2.6}$$

For MOHPM, the nonlinear function $\zeta(\tilde{\mu}(r,\theta))$ decomposes as

$$\begin{aligned}&\zeta\left(\tilde{\kappa}(s,\,\vartheta)\right)=\zeta(\tilde{\kappa}_0(s))+\vartheta\left(\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s))-\zeta(\tilde{\kappa}_0(s))\right)\\ &+\vartheta^2\left(\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s)+\tilde{\kappa}_2(s))-\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s))\right)+\cdots.\end{aligned} \tag{2.7}$$

The terms $\zeta(\tilde{\kappa}_0(s))$, $\{\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s))-\zeta(\tilde{\kappa}_0(s))\}$, $\{\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s)+\tilde{\kappa}_2(s))-\zeta(\tilde{\kappa}_0(s)+\tilde{\kappa}_1(s))\}$ and so on, appearing in equation (2.7) on the right hand side are Daftardar-Jafari polynomials defined in [19]. Equation (2.7) can be written in a more compact form if we write $\zeta_0=\zeta(\tilde{\kappa}_0(s))$and $\zeta_m=\zeta\left(\sum_{i=0}^{m}\tilde{\kappa}_i(s)\right)-\zeta\left(\sum_{i=0}^{m-1}\tilde{\kappa}_i(s)\right)$. Thus, the expression (2.7) reduces to

$$\zeta(\tilde{\kappa}(s,\,\vartheta))=\zeta_0+\sum_{j=1}^{\infty}\vartheta^j\zeta_j. \tag{2.8}$$

putting back, equation (2.8) for equation (2.3), also by introducing a number of unknown auxiliary functions,$\varepsilon_i(s,c_i)$; for$i=0,1,2,3,\ldots$ that depend on the variable s and some constants $c_0,c_1,c_2,\ldots$, we get a new homotopy for (2.1) as:

$$\begin{aligned}&\sum(\tilde{\kappa},\vartheta)=(1-\vartheta)[\xi(\tilde{\kappa}(s,\,\vartheta))-\xi(\kappa_{ini}(s,\,\vartheta))]\\ &+\vartheta[\xi(\tilde{\kappa}(s,\vartheta)+\varepsilon_0(s,c_0)\zeta_0+\sum_{k=1}^{\infty}\varepsilon_k(s,c_k)\vartheta^k\,\zeta_k-\Gamma(s)]=0,\end{aligned} \tag{2.9}$$

along with the boundary conditions:

$$\Delta(\tilde{\kappa}(s,\,\vartheta), \frac{\partial}{\partial s}(\tilde{\kappa}(s,\,\vartheta)\,)) = 0.$$

Now, comparing the coefficients of similar powers of ϑ in (2.9), we get linear differential equations of zeroth order, first order, second order and so on, which can be solved very easily.

Zeroth order problem:

$$\xi(\tilde{\kappa}_0(s)) = \xi(\kappa_{ini}(s)), \qquad \Delta\left(\tilde{\kappa}_0, \frac{d\tilde{\kappa}_0}{ds}\right) = 0. \tag{2.10}$$

First order problem:

$$\xi(\tilde{\kappa}_1(s)) + \varepsilon_0(s, c_o)\zeta_0 - \Gamma(s) = 0, \quad \Delta\left(\tilde{\kappa}_1, \frac{d\tilde{\kappa}_1}{ds}\right) = 0. \tag{2.11}$$

Second order problem:

$$\xi(\tilde{\kappa}_2(s)) + \varepsilon_1(s, c_1)(\zeta_1) = 0, \quad \Delta\left(\tilde{\kappa}_2, \frac{d\tilde{\kappa}_2}{ds}\right) = 0. \tag{2.12}$$

Third order problem:

$$\xi(\tilde{\kappa}_3(s)) + \varepsilon_2(s, c_2)(\zeta_2) = 0, \quad \Delta\left(\tilde{\kappa}_3, \frac{d\tilde{\kappa}_3}{ds}\right) = 0. \tag{2.13}$$

and so on.

Where $\varepsilon_i(s, c_i)$; $i = 0, 1, 2, 3, \ldots$, are auxiliary functions. The parameters c_i's are used to control the convergence and can itself be determined optimally. This can be done over the domain of the problem by minimizing the residual functional. In order to get an accurate result, solutions up to the higher order problems can be made but a solution up to third order will be sufficient. For $\vartheta = 1$, if the series (2.7) converges, then the approximate solution is given by

$$\kappa(s) = \tilde{\kappa}(s) = \tilde{\kappa}_0(s) + \tilde{\kappa}_1(s,\, c_0) + \tilde{\kappa}_2(s,\, c_0, c_1) + \tilde{\kappa}_3(s,\, c_0, c_1, c_2) + \cdots. \tag{2.14}$$

The resulting residual can be obtained by backward substitution of equation (2.13) into equation (2.1) as

$$\bar{R}(s,\, c_0, c_1, c_2, \ldots.) = \xi(\tilde{\kappa}(s)) + \zeta(\tilde{\kappa}(s)) - \Gamma(s). \tag{2.15}$$

The exact solution $\tilde{\kappa}$, will be obtained if $\bar{R} = 0$. In most of the problems usually$\bar{R} \neq 0$, and a minimization is needed over the domain of the problem. This can be done by using either least square's method, Galerkin's method or collocation method. When applying the method of least squares, we first introduce the functional

$$\psi(c_0, c_1, c_2, \ldots) = \int_t^h \bar{R}^2 dx, \tag{2.16}$$

and then minimizing it, we obtain

$$\frac{\partial\psi}{\partial c_0} = \frac{\partial\psi}{\partial c_1} = \frac{\partial\psi}{\partial c_2} = \cdots = 0. \tag{2.17}$$

For auxiliary constants we have to solve the following system, when applying Galerkin's method:

$$\int_t^h \bar{R}\frac{\partial\tilde{\kappa}}{\partial c_0}ds = 0, \quad \int_t^h \bar{R}\frac{\partial\tilde{\kappa}}{\partial c_1}ds = 0, \quad \int_t^h \bar{R}\frac{\partial\tilde{\kappa}}{\partial c_2}ds = 0, \ldots .. \tag{2.18}$$

3. APPLICATION OF OHPM AND MOHPM

In this section high accuracy of MOHPM is shown over the existing methods in the literature. The proposed method is applied to some linear and non linear differential equations of different orders. As a result, we see that MOHPM gives best approximation and takes very less time to produce the solution.
Problem 1. Consider fifth order linear boundary value problem [10].

$$\frac{d^5u}{ds^5} - u + 15\, e^s + 10\, \mathrm{s}\, e^s = 0, \quad 0 < s < 1, \tag{3.1}$$

$$u(0) = 0,\ u\ (1) = 0,\ \ u'(0) = 1,\ \ u'(1) = -e,\ \ u''(0) = 0. \tag{3.2}$$

The exact solution for this problem is $u(s) = s(1-s)e^s$.
To apply MOHPM, we take:

$$u(s,\vartheta) \quad = \quad u_0(s) + \vartheta\, u_1(s) + \vartheta^2 u_2(s), \tag{3.3}$$

$$\xi(\tilde{\kappa}(s,\vartheta)) = \frac{d^5u(s,\vartheta)}{ds^5},\ \xi(\kappa_{ini}(s,\vartheta)) = 0, \varepsilon_0(s,c_0) = 1,\ \varepsilon_1(s,c_1) = 1, \tag{3.4}$$

$$\xi(\tilde{\kappa}(s,\ \vartheta)) + \zeta(\tilde{\kappa}(s,\ \vartheta)) - \Gamma(s) = \frac{d^5u}{ds^5} - u + 15\, e^s + 10\, \mathrm{s}\, e^s. \tag{3.5}$$

Now put the above values in (2.9) and compare the coefficients of like powers of ϑ we get as:
Zeroth order problem:

$$(u_0)^{(5)}(s) = 0,\ u_0(0) = 0,\ u_0(1) = 0\,,\ u_0'(0) = 0,\ u_0'(1) = -e\,,\ u_0''(0) = 0. \tag{3.6}$$

First order problem:

$$\begin{aligned} &15\,\mathrm{e}^s + 10\,\mathrm{e}^s s - s\,u_0(s) + (u_1)^{(5)}(s) = 0,\\ &u_1(0) = 0, u_1(1) = 0, u_1'(0) = 0, u_1'(1) = 0, u_1''(0) = 0. \end{aligned} \tag{3.7}$$

Second order problem:

$$\begin{aligned} &- s\,u_1(s) + (u_2)^{(5)}(s) = 0,\ u_2(0) = 0,\ u_2(1) = 0,\\ &u_2'(0) = 0,\ u_2'(1) = 0\,,\ u_2''(0) = 0. \end{aligned} \tag{3.8}$$

Solve the above equations we obtain: $u_0(s), u_1(s), u_2(s)$, put these values in (3.3) and also $\vartheta = 1$, we get the following solution for t = 0and $h = 1$:

$$\begin{aligned} u(s) &= s - 0.5000000000\,s^3 - 0.3333333333\,s^4 - 0.125\,s^5\\ &- 0.0333333333\,s^6\ - 0.0069444444\,s^7\\ &- 0.001190476\,s^8 - 0.000173611\,s^9 - 0.000022047s^{10}\ - 0.000002480\,s^{11}\\ &- 2.5052\times10^{-7}s^{12} - 2.2967\times10^{-8}s^{13} - 1.9261\times10^{-9}s^{14}. \end{aligned} \tag{3.9}$$

The results for problem 1 are shown in table-1 and figure-1 as follows:

Problem 2. Fifth order non-linear boundary value problem [10].

$$\frac{d^5u}{ds^5} - u^2e^{-s} = 0, \quad 0 < s < 1, \tag{3.10}$$

$$u(0) = 1,\ \ u'(0) = 1,\ \ u''(0) = 1,\ u\ (1) = e,\ \ u'(1) = e\,. \tag{3.11}$$

Having exact solution $u(s) = e^s$. To solve this problem, we consider the second order approximation

$$u(s) = u_0(s) + u_1(s,c_0) + u_2(s,c_0,c_1). \tag{3.12}$$

Let, $u(s,\vartheta)=u_0(s)+\vartheta\, u_1(s)+\vartheta^2 u_2(s), \xi(\tilde{\kappa}(s,\vartheta))=\frac{d^5 u(s,\vartheta)}{ds^5}, \xi(\kappa_{ini}(s,\vartheta))=0, \zeta(\tilde{\kappa}(s,\vartheta))=u^2(s)e^{-s}, \varepsilon_0(s,c_0)=c_0,\ \varepsilon_1(s,c_1)=c_1, \varsigma_0=u_0^2(s),\ \varsigma_1=2u_0(s)u_1(s)+u_1^2(s)$.

Put the above values in (2.9) and compare coefficients of like powers of ϑ we get as:

Zeroth order problem:

$$(u_0)^{(5)}(s)=0,\ u_0(0)=1,\ u_0'(0)=1,\ u_0''(0)=1,\ u_0(1)=e,\ u_0'(1)=e. \tag{3.13}$$

First order problem:

$$\begin{aligned}&-\mathrm{e}^{-s}c_0\, u_0(s)^2+(u_1)^{(5)}(s)=0,\ u_1(0)=0,\\ &u_1(1)=0,\ u_1'(0)=0,\ u_1'(1)=0,\ u_1''(0)=0.\end{aligned} \tag{3.14}$$

Second order problem:

$$\begin{aligned}&-2\mathrm{e}^{-s}c_1\, u_0(s)\, u_1(s)-\mathrm{e}^{-s}c_1\, u_1(s)^2+(u_2)^{(5)}(s)=0,\\ &u_2(0)=0,\ u_2(1)=0,\ u_2'(0)=0,\ u_2'(1)=0,\ u_2''(0)=0.\end{aligned} \tag{3.15}$$

Solution of the above gives $u_0(s), u_1(s,c_0), u_2(s,c_0,c_1)$.

Now use (3.12) and apply the Galerkin's method consist of (2.15) and (2.18) we get the following values of $c_i's$ for t = 0and $h=1$,

$$c_0=0.999999240,\quad c_1=0.999758960.$$

The MOHPM approximate solution becomes:

$$\begin{aligned}u(s)=&1+s+\tfrac{s^2}{2}+0.1666666667s^3+0.0416666667s^4+\tfrac{s^5}{120}+\tfrac{s^6}{720}+\tfrac{s^7}{5040}\\ &+0.000024802\, s^8+0.000002756\, s^9+2.7557\times10^{-7}s^{10}+2.5050\times10^{-8}s^{11}\\ &+2.0928\times10^{-9}s^{12}+1.5506\times10^{-10}s^{13}+1.4690\times10^{-11}s^{14}.\end{aligned} \tag{3.16}$$

Now to check the accuracy of OHPM, we apply OHPM to (3.10) and obtain

$$c_0=1.000553563,\quad c_1=-0.417820368.$$

The approximate solution by OHPM is then given as

$$\begin{aligned}u(s)&=1+s+\frac{s^2}{2}+0.166668432s^3+0.041661159s^4+0.008337946s^5\\ &+0.00138968s^6+0.000198523\, s^7+0.000019825\, s^8+0.000006842\, s^9\\ &-5.0613\times10^{-7}s^{10}-4.1383\times10^{-8}s^{11}-1.6480\times10^{-8}s^{12}\\ &+1.4653\times10^{-8}s^{13}-9.8279\times10^{-9}s^{14}\end{aligned} \tag{3.17}$$

The results for problem 2 are shown in table-2 and figure-2 as follows:

Problem 3. Sixth order linear boundary value problem [17]

$$\frac{d^6u}{ds^6}-(u-6\,e^s)=0,\quad 0<s<1, \tag{3.18}$$

$$u(0)=1,\ \mathrm{u}(1)=0,\ \mathrm{u}''(0)=-1,\ \mathrm{u}''(1)=-2e,\ \mathrm{u}''''(0)=-3,\ \mathrm{u}''''(1)=-4e. \tag{3.19}$$

The exact solution for this problem is:

$$u(s)=(1-s)e^s.$$

To apply MOHPM, we use the steps used in problem-1 and in problem-2, we obtain approximate solution for t = 0and $h=1$ as

(3.20)

$$\begin{aligned}u(s)&=1.-3.5811\times10^{-12}s-0.5\,s^2-0.333333333\,s^3-0.125\,s^4-0.0333333333\,s^5\\ &-0.006944444\,s^6-0.001190476\,s^7-0.000173611\,s^8-0.000022046\,s^9\\ &-0.000002480\,s^{10}-2.5052\times10^{-7}\,s^{11}-2.2964\times10^{-8}\,s^{12}\end{aligned}$$

The results for problem 3 are shown in table-3 and figure-3 as follows:

(1) Table 3
(2) Figure 3

Problem 4. Sixth order nonlinear boundary value problem [24]

$$\frac{d^6u}{ds^6} - u^2e^s = 0, \quad 0 < s < 1, \tag{3.21}$$

$$u(0) = 1,\ u'(0) = -1,\ u''(0) = 1,\ u(1) = e^{-1},\ u'(1) - e^{-1},\ u''(1) = e^{-1}. \tag{3.22}$$

The exact solution is given as $u(s) = e^{-s}$.

To apply MOHPM, we consider the following second order approximation$u(s) = u_0(s) + u_1(s, c_0) + u_2(s, c_0, c_1)$.

Now we use the steps mentioned in problem-1 and problem-2.

Using the Galerkin's method which consist of (2.15) and (2.18) , we obtain the following values of $c_i^{,}s$for t $= 0$and $h = 1:$

$$c_0 = 0.999781503, \quad c_1 = 0.568319310\,.$$

The approximate solution by MOHPM becomes:

$$\begin{aligned} u = 1 - s + \tfrac{s^2}{2} &- 0.1666666676\,s^3 + 0.041666667s^4 - 0.008333333\,s^5 + \tfrac{s^6}{720} \\ &- \tfrac{s^7}{5040} + \tfrac{s^8}{40320} - 0.000002756\,s^9 + 2.7556 \times 10^{-7}s^{10} - 2.5048 \times 10^{-8}s^{11} \\ &+ 2.0862 \times 10^{-9}s^{12} - 1.5896 \times 10^{-10}s^{13} + 9.6037 \times 10^{-12}s^{14}. \end{aligned} \tag{3.23}$$

Also when we apply OHPM, we obtain the values of $c_i^{,}s$ as

$$c_0 = 0.999986482, \quad c_1 = 0.991801518\,.$$

The approximate solution given by OHPM is

$$\begin{aligned} u = 1 - s + \tfrac{s^2}{2} &- 0.1666666667\,s^3 + 0.0416666667\,s^4 - 0.0083333334\,s^5 \\ &+ 0.001388889\,s^6 - 0.000198413\,s^7 + 0.000024802\,s^8 - 0.000002756\,s^9 \\ &+ 2.7561 \times 10^{-7}s^{10} - 2.5071 \times 10^{-8}s^{11} + 2.0905 \times 10^{-9}s^{12} \\ &- 1.5728 \times 10^{-10}s^{13} + 7.1664 \times 10^{-12}s^{14}. \end{aligned} \tag{3.24}$$

Results for problem 4 are given in table 4 and figure 4 as follows:

(1) Table 4
(2) Figure 4

4. CONCLUSION

In this paper a new idea has been developed and effectively applied to four higher order boundary value problems of fifth and sixth orders which provide very accurate results as compared to other well known methods in practice. Our proposed method has great potential to solve ordinary differential equations of any order. The same technique can also be extended to the solutions of partial differential equations, integro-differential equations and system of differential equations, the results obtained for these types of differential equations will be revealed in our coming papers. The merit of MOHPM is that it requires only a few terms to obtain accurate approximate solutions. This technique has a great robust, to attract engineer, scientists and researchers of every field.

References

[1] S. Chandrasekhar, Hydrodynamic and hydro-magnetic stability, Dover Press, New York (1981).

[2] S. Liao, Beyond Perturbation: Introduction to the homotopy analysis method, Chapman & Hall/CRC Press, Florida (2004).

[3] M.M. Rashidi, O. Anwar Beg, M.T. Rastegari, A Study of Non-Newtonian Flow and Heat Transfer over a Non-Isothermal Wedge using the Homotopy Analysis Method, Chemical Engineering Communications 199 (2012) 231-256.

[4] O. Anwar Beg, M.M. Rashidi, T.A. Beg, M. Asadi, Homotopy Analysis of Transient Magneto-Bio-Fluid Dynamics of Micropolar Squeeze Film in a Porous Medium: a Model for Magneto-Bio-Rheological Lubrication, Journal of Mechanics in Medicine and Biology, **12** (2012).

[5] H.J. He, Homotopy perturbation method for solving boundary value problems, Physics Letters A, **350** (2006) 87-88.

[6] M.M. Rashidi, D.D. Ganji, S. Dinarvand, Explicit Analytical Solutions of the Generalized Burger and Burger–Fisher Equations by Homotopy Perturbation Method, Numerical Methods for Partial Differential Equations, **25** (2009) 409-417.

[7] G. Adomian, A review of the decomposition method and some recent results for nonlinear equations, Computers and Mathematics with Applications, **21** (1991) 101-127.

[8] V. Marinca, N. Herisanu, An optimal homotopy asymptotic approach applied to non-linear MHD Jeffery-Hamel flow, Math. Probl. Eng., Art. N 169056, published 2011.

[9] V. Marinca, N. Herisanu andI. Nemes, Optimal homotopy asymptotic method for solving Blasius.equation, Appl. Math. Comput., 231 (2014) 134-139.

[10] J. Ali, S. Islam, S. Shah and H. Khan,The optimal homotopy asymptotic method for the solution of fifth and sixth order boundary value problems, World Applied Sciences Journal, **15** (2011) 1120-1126.

[11] N. Herisanu, V. Marinca, Gh.Madescu, An Analytical approach to non-linear dynamical model of a permanet magnet synchronous generator, Wind Energy, 2015,DOI:10.1002/we.1785

[12] N. Herisanu,V. Marinca, On the flow of a Walters-type B' viscoelastic fluid in a vertical channel with porous wall, Int.J.Heat and Mass Transfer, **79** (2014) 146-165.

[13] V. Marinca and N. Herisanu, Nonlinear dynamic analysis of an electrical machine rotor-bearing system by optimal homotopy perturbation method, Computers and Mathematics with Applications, **61** (2011) 2019-2024.

[14] N. Herisanu and V. Marinca, Optimal homotopy perturbation method for non-conservative dynamical system of a rotating electrical machine, Zeitschrift fur Naturforschung A, **67**.(2012) 509-516

[15] V. Marinca and N. Herisanu, Optimal homotopy perturbation approach to thin film flow of a fourth grade fluid, AIP Conference Proceedings, 1479 (2012) 2383-2386.

[16] M.A. Noor and S. T. Mohyud-Din,Variational iteration method for fifth order boundary value problems using He's polynomial, Mathematical Problems in Engineering, **10** (2008) ID 954794.

[17] M.A. Noor, K.I. Noor and S.T. Mohyud-Din,Variational iteration method for solving sixth order boundary value problems, communications in non linear science and numerical simulations, **14** (2009) 2571-2580.

[18] M.A. Noor and S.T. Mohyud-Din, Variational iteration technique for solving higher order boundary value problems, Applied Mathematics and Computation, **189** (2007) 1929–1942.

[19] V. Daftardar-Gejji and H. Jafari, An iterative method for solving nonlinear functional equations, Journal of Mathematical Analysis and Applications, **316** (2006) 753–763.

[20] S. Bhalekar and V. Daftardar-Gejji,Convergence of the New Iterative Method, International Journal of Differential Equations, **11** (2011) Article ID 989065,10pages

[21] J. Ali S. Islam,I. Ali and H. Khan, Application of optimal homotopy asymptotic method to higher order boundary value problems, Abstract and Applied Analysis, **12** Article ID 401217, 14 pages.

Table 1. Table1 shows comparison of the errors obtained by (MOHPM)(3.9), (OHAM) [10], (VIMHP) [16], (VIM) [18], (B-Spline) [22], with the exact solution. We observe that the result of MOHPM is better and more accurate than the above mentioned methods.

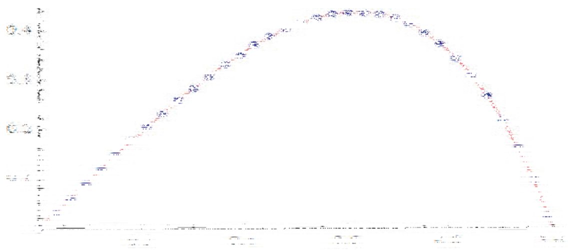
s	Exact	MOHPM	E* (MOHPM)	E* (OHAM)	E* (VIMHP)	E* (VIM)	E* (B Spline)
0.0	0.000000	0.000000	0.	0.	0.	0.	0.
0.1	0.099465	0.099465	5.4x10^{-14}	9.0x10^{-11}	- 3.0x10^{-11}	1.0x10^{-09}	- 7.0x10^{-04}
0.2	0.195424	0.195424	3.7x10^{-13}	4.0x10^{-10}	- 2.0x10^{-10}	2.0x10^{-09}	- 7.2x10^{-04}
0.3	0.283470	0.283470	1.0x10^{-12}	5.0x10^{-10}	- 4.0x10^{-10}	1.0x10^{-09}	4.1x10^{-04}
0.4	0.358037	0.358037	1.9x10^{-12}	2.0x10^{-11}	- 8.0x10^{-10}	2.0x10^{-09}	4.6x10^{-04}
0.5	0.412180	0.412180	2.7x10^{-12}	1.0x10^{-09}	- 1.0x10^{-09}	3.1x10^{-08}	4.7x10^{-04}
0.6	0.43730	0.437308	3.0x10^{-12}	2.0x10^{-09}	- 2.0x10^{-09}	3.7x10^{-08}	4.8x10^{-04}
0.7	0.422888	0.422888	2.1x10^{-12}	2.0x10^{-09}	- 2.0x10^{-09}	4.1x10^{-08}	3.9x10^{-04}
0.8	0.356086	0.356086	3.7x10^{-12}	1.0x10^{-09}	- 2.0x10^{-09}	3.1x10^{-08}	3.1x10^{-04}
0.9	0.221364	0.221364	- 3.2x10^{-11}	4.0x10^{-10}	-1.0x1- $_{-09}$	1.4x10^{-08}	1.6x10^{-04}
1.0	0.000000	2.5x10^{-8}	- 1.6x10^{-10}	0.	0.	0.	0.

[22] S.U. Islam and M.A. Khan, A numerical method based on polynomials sextic spline functions for the solution of special fifth-order boundary value problems, Applied Mathematics and Computation, **181** (2006) 356-361.

[23] A.M. Wazwaz, The numerical solution of sixth-order boundary value problems by the modified decomposition method, Applied Mathematics and Computation, **118** (2001) 311-325.

[24] M.A. Noor and S.T. Mohyud-Din, Homotopy perturbation method for solving sixth order boundary value problems, Computers and Mathematics with Applications, **55** (2008) 2953-2972.

Table 1



E*=Exact-Approx.

Figure 1: Dotted curve-sol: (MOHPM) and solid curve-sol: (Exact).

TABLE 2

E*=Exact-Approx.

Table 2. Table 2 shows comparison of the solutions obtained by MOHPM (3.16), OHPM (3.17), OHAM [10], VIM [18] and B-Spline [22]. From the numerical results it is clear that MOHPM is more efficient and more accurate.

s	Exact	MOHPM	E* MOHPM	E* OHPM	E* OHAM	E* VIM	E* B-Spline
0.0	1.000000	1.000000	0.0	0.0	0.0	0.0000	0.0000
0.1	1.105170	1.105170	3.1x10^{-15}	1.2x10^{-09}	1.9x10^{-10}	- 3.0x10^{-11}	- 8.0x10^{-03}
0.2	1.221402	1.221402	1.9x10^{-14}	6.8x10^{-09}	1.2x10^{-09}	- 2.2x10^{-10}	- 1.2x10^{-03}
0.3	1.349858	1.349858	5.4x10^{-14}	1.4x10^{-08}	3.3x10^{-09}	- 4.0x10^{-10}	- 5.0x10^{-03}
0.4	1.491824	1.491824	1.0x10^{-13}	2.0x10^{-08}	6.3x10^{-09}	- 8.0x10^{-10}	3.0x10^{-03}
0.5	1.648721	1.648721	1.4x10^{-13}	2.1x10^{-08}	9.3x10^{-09}	- 1.2x10^{-09}	8.0x10^{-03}
0.6	1.822118	1.822118	1.6x10^{-13}	1.7x10^{-08}	1.1x10^{-08}	- 209x10^{-09}	6.0x10^{-03}
0.7	2.013752	2.013752	1.5x10^{-13}	1.2x10^{-08}	1.1x10^{-08}	- 2.2x10^{-09}	-0.0000
0.8	2.225540	2.225540	9.9x10^{-14}	7.0x10^{-09}	8.2x10^{-09}	- 1.9x10^{-09}	9.0 10^{-03}
0.9	2.459603	2.459603	1.1x10^{-14}	2.0x10^{-09}	1.9x10^{-09}	- 1.4x10^{-09}	- 9.0x10^{-03}
1.0	2.718281	2.718281	1.0E-13	3.0E-09	0.00	0.000	0.0000

Figure 2: Dotted curve-sol: (MOHPM) and solid curve-sol: (Exact).

Table 3
E*=Exact-Approx.

Table 3. Table3shows comparison of the errors obtained by MOHPM(3.20), OHAM [10], ADM [23], VIM [18] and HPM [24], with the exact solution. We observe that our results of MOHPM are better and more accurate than the above mentioned methods.

s	Exact	MOHPM	E* MOHPM	E* OHAM	E* ADM	E* VIM	E* HPM
0.0	1.	1.	0.	0.	0.	0.	0.
0.1	0.994653	0.994653	$3.5x10^{-13}$	$2.1x10^{-08}$	$-4.1x10^{-04}$	$-4.1x10^{-04}$	$-4.1x10^{-04}$
0.2	0.977122	0.977122	$6.7x10^{-13}$	$4.0x10^{-08}$	$-7.8x10^{-04}$	$-7.8x10^{-04}$	$-7.8x10^{-04}$
0.3	0.944901	0.944901	$9.2x10^{-13}$	$5.7x10^{-08}$	$-1.1x10^{-03}$	$-1.1x10^{-03}$	$-1.1x10^{-03}$
0.4	0.895094	0.895094	$1.1x10^{-12}$	$7.0x10^{-08}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$
0.5	0.824360	0.824360	$9.0x10^{-13}$	$7.6x10^{-08}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$
0.6	0.728847	0.728847	$-1.6x10^{-12}$	$7.5x10^{-08}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$	$-1.3x10^{-03}$
0.7	0.604125	0.604125	$-1.9x10^{-11}$	$6.5x10^{-08}$	$-1.1x10^{-03}$	$-1.1x10^{-03}$	$-1.1x10^{-03}$
0.8	0.445108	0.445108	$-1.1x10^{-10}$	$4.8x10^{-08}$	$-4.1x10^{-04}$	$-4.1x10^{-04}$	$-4.1x10^{-04}$
0.9	0.245960	0.245960	$-5.3x10^{-10}$	$2.5x10^{-08}$	$-7.8x10^{-04}$	$-7.8x10^{-04}$	$-7.8x10^{-04}$
1.0	0.	2.0E-O9	$-2.1x10^{-09}$	$-2.1x10^{-09}$	0.0	0.0	0.0

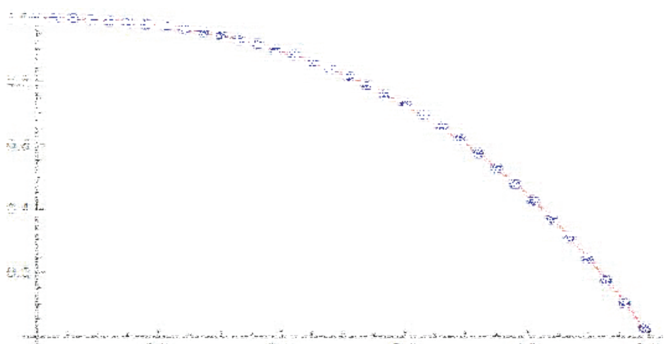


Figure 3: Dotted curve-sol: (MOHPM) and solid curve-sol: (Exact).
Table 4
E*=Exact-Approx.

Table 4. Table 4 shows comparison of errors obtained by MOHPM (3.23), OHPM (3.24), OHAM [21], ADM [23], VIM [18] and HPM [24], with the exact solution. Results indicate clearly that MOHPM gives better and accurate approximations than the above mentioned methods.

s	Exact	MOHPM	E* MOHPM	E* OHPM	E* OHAM	E* VIM	E* ADM	E* HPM
0.0	1.000000	1.000000	0.0	0.0	0	0	0	0
0.1	0.904837	0.904837	4.1x10^{-15}	1.2x10^{-09}	-4.82 x10^{-10}	- 2.3x10^{-07}	- 1.2x10^{-04}	- 1.2x10^{-04}
0.2	0.818730	0.818730	2.4x10^{-14}	6.8x10^{-09}	- 4.92x10^{-10}	- 1.3x10^{-06}	- 2.3x10^{-04}	- 2.3x10^{-04}
0.3	0.740818	0.740818	5.5x10^{-14}	1.4x10^{-08}	- 2.37x10^{-11}	- 3.3x10^{-06}	- 3.2x10^{-04}	- 3.2x10^{-04}
0.4	0.670320	0.670320	8.3x10^{-14}	2.0x10^{-08}	5.11x10^{-10}	- 5.2x10^{-06}	- 3.8x10^{-04}	- 3.8x10^{-04}
0.5	0.606530	0.606530	9.3x10^{-14}	2.1x10^{-08}	6.42x10^{-10}	- 6.1x10^{-06}	- 4.0x10^{-04}	- 4.0x10^{-04}
0.6	0.548811	0.548811	8.2x10^{-14}	1.7x10^{-08}	2.02x10^{-10}	- 5.7x10^{-06}	- 3.9x10^{-04}	- 3.9x10^{-04}
0.7	0.496585	0.496585	5.5x10^{-14}	1.2x10^{-08}	- 5.37x10^{-10}	- 4.0x10^{-06}	- 3.3x10^{-04}	- 3.3x10^{-04}
0.8	0.449328	0.449328	2.9x10^{-14}	7.0x10^{-09}	- 1.02x10^{-09}	- 1.9x10^{-06}	- 2.4x10^{-04}	- 2.4x10^{-04}
0.9	0.406569	0.406569	3.1x10^{-14}	2.0x10^{-09}	- 8.23x10^{-10}	- 3.5x10^{-07}	- 1.2x10^{-04}	- 1.2x10^{-04}
1.0	0.367879	0.367879	1.1x10^{-13}	3.0x10^{-09}	- 2.05x10^{-12}	- 5.0x10^{-10}	2.0x10^{-09}	2.0x10^{-09}

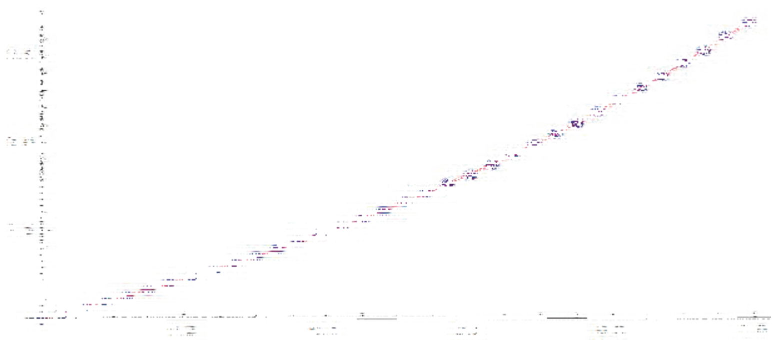


Figure 4: Dotted curve-sol: (MOHPM) and solid curve-sol: (Exact).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1061 – 1066

On some questions regarding projectivity criteria

Jang Hyun Jo *

Abstract

We investigate questions which are related to projectivity criteria and give some partial answers and related results to them.

Keywords: Complete cohomology, free, projective, stably free

2000 AMS Classification: 16D40, 18G05, 18G20

Received : 13.04.2015 *Accepted* : 27.08.2015 *Doi* : 10.15672/HJMS.20164513112

1. Introduction

The purpose of this paper is to investigate questions related to projectivity criteria. It is well-known that if G is a finite group, then a $\mathbb{Z}G$-module M is projective if and only if M is $\mathbb{Z}$-free and $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ (cf. [5]). In [16] we investigated whether or not only finite groups satisfy the criterion above, and showed that this is true in the class of groups $\mathbf{LH}\mathfrak{F}$. For the definitions of $\mathbf{LH}\mathfrak{F}$ and some other terminologies below in this section, see Section 2.

Note that if G is a virtually torsion-free group with $\operatorname{vcd} G = n$ and M is a $\mathbb{Z}$-free $\mathbb{Z}G$-module, then $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ if and only if $\operatorname{proj.dim}_{\mathbb{Z}G} M \leq n$ ([5, Theorem X.5.3]). This result was generalized in [15, Theorem 4.7] as follows: if G is a $\mathbf{H}\mathfrak{F}$-group and $\operatorname{spli} G < \infty$, then $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ if and only if $\operatorname{proj.dim}_{\mathbb{Z}G} M \leq \operatorname{pccd} G$.

It is also known that if H is a subgroup of finite index in G, then a $\mathbb{Z}G$-module M is $\mathbb{Z}G$-projective if and only if M is $\mathbb{Z}H$-projective and $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ ([6, Lemma 4.1 (a)]).

In these viewpoints, we may ask the following questions:

1.1. Question. Let n be a nonnegative integer. Suppose that G satisfies the following property: for any $\mathbb{Z}$-free $\mathbb{Z}G$-module M, $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ if and only if $\operatorname{proj.dim}_{\mathbb{Z}G} M \leq n$. Is it true that $\operatorname{pccd} G \leq n$?

1.2. Question. Let H be a subgroup of G. Suppose that (G, H) satisfies the property that for any $\mathbb{Z}G$-module M, M is $\mathbb{Z}G$-projective if and only if M is $\mathbb{Z}H$-projective and $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$. Is it true that $|G : H| < \infty$?

*Department of Mathematics, Sogang University, Seoul, 121-742 KOREA,
Email: `jhjo@sogang.ac.kr`

It is known in [16, Corollary 2.7] that Question 1.1 has a positive answer for any $\mathbf{LH}\mathfrak{F}$-group and $n = 0$.

On the other hand, recall the following conjecture (a special case of Moore's conjecture ([1, p 64]), which is a far reaching generalization of Serre's theorem [1, p 65]) on cohomological dimension of groups.

1.3. Conjecture. *Let G be a torsion-free group and H a subgroup of finite index in G. Then every $\mathbb{Z}G$-module M which is $\mathbb{Z}H$-projective is also $\mathbb{Z}G$-projective.*

In the same sprit as the questions above, we also naturally ask the following:

1.4. Question. Let G be a torsion-free group and H a subgroup of G. Suppose that (G, H) satisfies the property that every $\mathbb{Z}H$-projective $\mathbb{Z}G$-module is $\mathbb{Z}G$-projective. Is it true that $|G : H| < \infty$?

It can be seen that for a torsion-free group G and its subgroup H, if Question 1.4 has an affirmative answer for (G, H), then Question 1.2 has also an affirmative answer for (G, H).

We give some partial answers and related results to Questions 1.1, 1.2, and 1.4 in Theorems 3.7, 3.8, 3.9, and 3.10 and Corollaries 3.5 and 3.6.

2. Preliminaries

In this section, we briefly introduce some definitions and preliminary results. For more details, we recommend each reference below.

1. ([18, 3]) The class $\mathbf{H}\mathfrak{F}$ is the smallest class of groups containing the class of finite groups and which contains a group G whenever G admits a finite dimensional contractible G-C-complex whose stabilizers are already in $\mathbf{H}\mathfrak{F}$. The class $\mathbf{LH}\mathfrak{F}$ is the class of groups such that all of its finitely generated subgroups are in $\mathbf{H}\mathfrak{F}$. The class $\mathbf{LH}\mathfrak{F}$ is extension closed, closed under ascending unions, and closed under amalgamated free products and HNN extensions. The class $\mathbf{LH}\mathfrak{F}$ contains, for example, all elementary amenable groups and all linear groups.

2. The cohomological dimension of G, denoted $\operatorname{cd} G$, is the projective dimension of the trivial G-module $\mathbb{Z}$ over $\mathbb{Z}G$. For a virtually torsion-free group G, i.e., G has a torsion-free subgroup of finite index, it was well-known that all torsion-free subgroups of G of finite index have the same cohomological dimension (cf. [5]). The common cohomological dimension of the torsion-free subgroups of finite index is called the virtual cohomological dimension of G and is denoted by $\operatorname{vcd} G$. The finiteness of $\operatorname{vcd} G$ ensures that the Farrell cohomology of a group is well defined. There are other well-known invariants of a group which have been accompanied with the Ikenaga's generalized cohomology ([14]) and the complete cohomology ([4, 12, 19]):

(1) $\underline{\operatorname{cd}}\, G := \sup\{\, n \,:\, \operatorname{Ext}^n_{\mathbb{Z}G}(M, F) \neq 0,\ M : \mathbb{Z}\text{-free},\ F : \mathbb{Z}G\text{-free}\}$ ([14]).
(2) $\operatorname{spli} G := \sup\{n \,:\, \operatorname{Ext}^n_{\mathbb{Z}G}(I, -) \neq 0,\ I : \mathbb{Z}G\text{-injective}\}$ ([11]).
(3) $\operatorname{silp} G := \sup\{n \,:\, \operatorname{Ext}^n_{\mathbb{Z}G}(-, P) \neq 0,\ P : \mathbb{Z}G\text{-projective}\}$ ([11]).
(4) $\operatorname{fin.dim} G := \sup\{n \,:\, \operatorname{proj.dim}_G M = n < \infty\}$ ([20]).
(5) $\operatorname{pccd} G := \sup\{\, n \,:\, H^n(G, P) \neq 0,\ P : \mathbb{Z}G\text{-projective}\}$ ([15]).
(6) $\operatorname{Gcd} G := \operatorname{Gpd}_{\mathbb{Z}G}\mathbb{Z}$, the Gorenstein projective dimension of the trivial $\mathbb{Z}G$-module $\mathbb{Z}$ ([2, 3]).

It is well known from [2, 3, 7, 11, 13, 14, 15, 17, 22] that for any group G,

(a) $\operatorname{pccd} G \leq \underline{\operatorname{cd}}\, G = \operatorname{Gcd} G \leq \operatorname{silp} G = \operatorname{spli} G \leq \underline{\operatorname{cd}}\, G + 1 = \operatorname{Gcd} G + 1$.
(b) $-1 \leq \operatorname{pccd} G \leq \infty$.
(c) If G is the Thompson group T, $\bigoplus_{i=1}^{\infty} \mathbb{Z}$, or $GL_n(K)$, where K is a subfield of the algebraic closure of $\mathbb{Q}$, then $\operatorname{pccd} G = -1$.

(d) If $G = *_{n\in\mathbb{N}} G_n$, where $G_n := \bigoplus_{i=1}^{n} \mathbb{Z}$, then $\operatorname{pccd} G = \infty$.
(e) If $\operatorname{Gcd} G < \infty$, then $\operatorname{Gcd} G = \operatorname{pccd} G$ and so $-1 < \operatorname{pccd} G < \infty$.
(f) $\operatorname{fin.dim} G \leq \operatorname{spli} G$, the equality holds when $G \in \mathbf{LH}\mathfrak{F}$ or $\operatorname{spli} G < \infty$.

3. Main results

In what follows, let G be an arbitrary discrete group and $\mathbb{Z}G$ its group ring. We write "G-module", "G-projective", etc. instead of "$\mathbb{Z}G$-module", "$\mathbb{Z}G$-projective", etc.

3.1. Lemma. *Let G be a group satisfying the following property: for any $\mathbb{Z}$-free G-module M,*

$$\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty \text{ if and only if } \operatorname{proj.dim}_{\mathbb{Z}G} M \leq n.$$

Then $\operatorname{fin.dim} G \leq n+1$.

Proof. Let N be a G-module with $\operatorname{proj.dim}_{\mathbb{Z}G} N < \infty$. Consider an exact sequence of G-modules $0 \to K \to P \to N \to 0$, where P is G-projective. It is clear that K is $\mathbb{Z}$-free and $\operatorname{proj.dim}_{\mathbb{Z}G} K < \infty$. Thus $\operatorname{proj.dim}_{\mathbb{Z}G} K \leq n$ by the assumption and so $\operatorname{proj.dim}_{\mathbb{Z}G} N \leq n+1$. Hence we conclude that $\operatorname{fin.dim} G \leq n+1$. □

In [8] Dembegioti and Talelli proposed the following conjecture and gave some example of groups satisfying it.

3.2. Conjecture. *For any group G,* $\operatorname{spli} G = \underline{\operatorname{cd}}\, G + 1$.

In [3] Bahlekeh, Dembegioti, and Talelli proposed the following conjecture.

3.3. Conjecture. *For any group G,* $\operatorname{fin.dim} G = \operatorname{Gcd} G + 1$.

Note that $\operatorname{Gcd} G = \underline{\operatorname{cd}}\, G$ for any group G, and $\operatorname{fin.dim} G = \operatorname{spli} G$ when G is an $\mathbf{LH}\mathfrak{F}$-group. Thus Conjecture 3.2 is equivalent to Conjecture 3.3 when G is an $\mathbf{LH}\mathfrak{F}$-group.

3.4. Theorem. *If Conjecture 3.3 is true, then Question 1.1 has an affirmative answer.*

Proof. Assume that G satisfies the property in Question 1.1. Then $\operatorname{fin.dim} G \leq n+1$ by Lemma 3.1. By the assumption, it follows that $\operatorname{Gcd} G \leq n$. Hence $\operatorname{pccd} G \leq n$. □

3.5. Corollary. *Suppose that G satisfies the one of the following:*
(1) $\underline{\operatorname{cd}}\, G = 0$ *or* 1.
(2) *duality group.*
(3) *fundamental group of graph of finite groups.*
(4) *fundamental group of certain finite graph of group of type FP_∞ in* [8, Theorem 3.5].

Then Question 1.1 has an affirmative answer for G.

Proof. It is known from [8, 10] that if a group G is one of the list above, then G satisfies Conjecture 3.3. Hence the result follows from Theorem 3.4. □

The following corollary shows that the validity of Conjecture 3.3 settles Question A in [16] completely.

3.6. Corollary. *Let G be a group with the property that every $\mathbb{Z}$-free G-module of finite projective dimension is projective. If G satisfies Conjecture 3.3, then G is finite.*

Proof. Note that G is finite if and only if $\operatorname{pccd} G = 0$ ([15, Proposition 3.9]). Hence the result follows immediately from Theorem 3.4. □

3.7. Theorem. *Let G be a virtually torsion-free group. If G is an* $\mathbf{LH}\mathfrak{F}$*-group, then Question 1.1 has an affirmative answer for G.*

Proof. Assume that G satisfies the property in Question 1.1. Then $\operatorname{fin.dim} G \leq n+1$ by Lemma 3.1. Since G is an $\mathbf{LH}\mathfrak{F}$-group, it follows from [22, Corollary 2] that $\operatorname{spli} G \leq n+1$. Since $\underline{\operatorname{cd}}\, G \leq \operatorname{silp} G = \operatorname{spli} G$, it follows that $\underline{\operatorname{cd}}\, G \leq n+1$. Let H be a torsion-free subgroup of finite index in G. Since G is an $\mathbf{LH}\mathfrak{F}$-group, it follows from [22, Corollary 2] and [11, 5.2] that $\operatorname{fin.dim} G = \operatorname{spli} G = \operatorname{spli} H = \operatorname{fin.dim} H < \infty$. By [22, Corollary 1] we have $\operatorname{cd} H < \infty$. Then $\underline{\operatorname{cd}}\, G = \underline{\operatorname{cd}}\, H = \operatorname{cd} H = \operatorname{vcd} G$ by [14, Proposition 3, Proposition 5] and so $\operatorname{vcd} G \leq n+1$. Suppose that $\operatorname{vcd} G = n+1$. Then $\operatorname{cd} H = \operatorname{proj.dim}_{\mathbb{Z}H} \mathbb{Z} = n+1$. But this contradicts to the property in Question 1.1. Hence $\operatorname{pccd} G = \operatorname{vcd} G \leq n$ as required. □

3.8. Theorem. *Let H be a normal subgroup of G. Suppose that every H-projective, G-module M with* $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ *is H-free. Then Question 1.2 has an affirmative answer for (G,H).*

Proof. Assume that a G-module M is H-projective and $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$. Let $Q = G/H$. Since M is H-projective, it follows from a spectral sequence argument as in the proof of [6, Lemma 4.1 (a)] that for any G-module N,

$$\operatorname{Ext}^i_{\mathbb{Z}G}(M,N) \cong H^i(Q, \operatorname{Hom}_{\mathbb{Z}H}(M,N)).$$

Suppose that there exists a projective Q-module L and $k > 0$ such that $H^k(Q,L) \neq 0$. We can regard L as a G-module via the quotient map $q : G \twoheadrightarrow Q$. By the assumption, we see that M is G-projective and thereby for any $i > 0$,

$$\operatorname{Ext}^i_{\mathbb{Z}G}(M,L) \cong H^i(Q, \operatorname{Hom}_{\mathbb{Z}H}(M,L)) = 0.$$

Since M is H-free by the assumption, it follows that

$$\operatorname{Hom}_{\mathbb{Z}H}(M,L) \cong \operatorname{Hom}_{\mathbb{Z}H}(\oplus \mathbb{Z}H, L) \cong \prod \operatorname{Hom}_{\mathbb{Z}H}(\mathbb{Z}H, L) \cong \prod L$$

as Q-modules (cf. [21, Thorem 2.31]). Thus we have

$$H^k(Q, \operatorname{Hom}_{\mathbb{Z}H}(M,L)) \cong H^k(Q, \prod L) \cong \prod H^k(Q,L) \neq 0$$

(cf. [21, Proposition 7.22]), which makes a contradiction. Hence $H^i(Q,S) = 0$ for each $i > 0$ and any projective Q-module S, and therefore $\operatorname{pccd} Q \leq 0$. But since $\operatorname{pccd} Q \neq -1$, we see that $\operatorname{pccd} Q = 0$ and so Q is finite by [15, Proposition 3.9]. Hence we conclude that $|G : H| < \infty$. □

3.9. Theorem. *Let H be a normal subgroup of G. Suppose that every H-stably free, G-module M with* $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$ *is H-free. Assume further that for any H-stably-free, G-module M with* $\operatorname{proj.dim}_{\mathbb{Z}G} M < \infty$*, there exist H-free modules M' and F such that $M \oplus M' \cong F$ as H-modules and the H-free rank of M' is different from that of F. Then Question 1.2 has an affirmative answer for (G,H).*

Proof. Let $Q = G/H$. By the proof of Theorem 3.8, we see that for any G-module N,

$$\operatorname{Ext}^i_{\mathbb{Z}G}(M,N) \cong H^i(Q, \operatorname{Hom}_{\mathbb{Z}H}(M,N)).$$

Suppose that there exists a projective Q-module L and $k > 0$ such that $H^k(Q,L) \neq 0$. By the assumption, it follows that M is G-projective and therefore we have that for any $i > 0$,

$$\operatorname{Ext}^i_{\mathbb{Z}G}(M,L) \cong H^i(Q, \operatorname{Hom}_{\mathbb{Z}H}(M,L)) = 0.$$

Note that $\mathrm{Hom}_{\mathbb{Z}H}(F, L) \cong \prod_I L$ and $\mathrm{Hom}_{\mathbb{Z}H}(M', L) \cong \prod_J L$ as Q-modules, where the cardinalities of I and J are the H-free ranks of F and M', respectively. Note also that

$$\mathrm{Hom}_{\mathbb{Z}H}(F, L) \cong \mathrm{Hom}_{\mathbb{Z}H}(M, L) \oplus \mathrm{Hom}_{\mathbb{Z}H}(M', L).$$

Thus we have

$$\begin{aligned}\prod_J H^k(Q, L) &\cong H^k(Q, \mathrm{Hom}_{\mathbb{Z}H}(M, L)) \oplus (\prod_J H^k(Q, L)) \\ &\cong H^k(Q, \mathrm{Hom}_{\mathbb{Z}H}(M, L)) \oplus H^k(Q, \mathrm{Hom}_{\mathbb{Z}H}(M', L)) \\ &\cong H^k(Q, \mathrm{Hom}_{\mathbb{Z}H}(F, L)) \cong H^k(Q, \prod_I L) \cong \prod_I H^k(Q, L).\end{aligned}$$

But this makes a contradiction, since the H-free rank of F is different from that of M'. Hence we can conclude that Q is finite by the same argument of the proof of Theorem 3.8. Therefore $|G : H| < \infty$. □

3.10. Theorem. *Let G be a torsion-free group and H a normal subgroup of G. Suppose that* $\mathrm{pccd}\,(G/H) > -1$ *and (G, H) satisfies one of the following:*

(*a*) *Every H-projective, G-module M with* $\mathrm{proj.dim}_{\mathbb{Z}G}\, M < \infty$ *is H-free.*
(*b*) *Every H-stably free, G-module M with* $\mathrm{proj.dim}_{\mathbb{Z}G}\, M < \infty$ *is H-free, and for any H-stably-free, G-module M with* $\mathrm{proj.dim}_{\mathbb{Z}G}\, M < \infty$*, there exist H-free modules M' and F such that $M \oplus M' \cong F$ as H-modules and the H-free rank of M' is different from that of F.*

Then Question 1.4 has an affirmative answer for (G, H).

Proof. It can be proved by the same argument of the proof of Theorems 3.8 and 3.9. □

3.11. Remark. Let X be a CW-complex such that the universal cover $\widetilde{X}$ is $(m-1)$-connected. It is known from [9, Proposition 1.4] that if $m \geq 3$, then X has the m-type of a finite m-complex if and only if its Swan-Wall class $SW_m[X] = 0$, where $SW_m[X] := C_m(\widetilde{X})/B_m(\widetilde{X}) \in C(\pi_1(X))$, and where $C(\pi_1(X))$ is the abelian monoid of stable equivalence classes of finitely generated $\pi_1(X)$-modules. Recall that for an abelian group A and positive integer m, a CW-complex Y is called a Moore space of type $M(A, m)$ if $H_0(Y) = \mathbb{Z}$, $H_m(Y)$ is isomorphic to A, and $H_i(Y) = 0$ for $i \neq 0, m$.

Suppose now that G is a finite group. Let X be a finite dimensional, finite type CW-complex X with $\pi_1(X) \cong G$ such that $\widetilde{X}$ is a Moore space of type $M(A, m)$. Then we see that $\mathrm{proj.dim}_{\mathbb{Z}G}\,(C_m(\widetilde{X})/B_m(\widetilde{X})) < \infty$, since

$$0 \to C_{\dim X}(\widetilde{X}) \to \cdots \to C_m(\widetilde{X}) \to C_m(\widetilde{X})/B_m(\widetilde{X}) \to 0$$

is a G-free resolution of $C_m(\widetilde{X})/B_m(\widetilde{X})$. It is clear that the sequence of G-modules

$$0 \to Z_m(\widetilde{X})/B_m(\widetilde{X}) \to C_m(\widetilde{X})/B_m(\widetilde{X}) \to C_m(\widetilde{X})/Z_m(\widetilde{X}) \to 0$$

is exact. Since $C_m(\widetilde{X})/Z_m(\widetilde{X}) \cong B_{m-1}(\widetilde{X}) \subset C_m(\widetilde{X})$ and $C_m(\widetilde{X})$ is $\mathbb{Z}$-free, it follows that $C_m(\widetilde{X})/Z_m(\widetilde{X})$ is $\mathbb{Z}$-free. Thus we see that $C_m(\widetilde{X})/B_m(\widetilde{X})$ is finitely generated G-projective and so $C_m(\widetilde{X})/B_m(\widetilde{X}) = 0 \in \widetilde{K}_0(\mathbb{Z}\pi_1(X))$. By [9, Proposition 1.4], it follows that X has the m-type of a finite CW-complex. Consequently, we can conclude that if G is a finite group, then every finite dimensional, finite type CW-complex X with $\pi_1(X) \cong G$ such that $\widetilde{X}$ is a Moore space of type $M(A, m)$ has the m-type of a finite CW-complex. But we do not yet know whether the converse of this holds.

References

[1] E. Aljadeff, Profinite group, profinite completions and a conjecture of Moore, *Adv. Math.* **201** (2006), 63–76.

[2] J. Asadollahi, A. Bahlekeh, and S. Salarian, On the hierarchy of cohomological dimension of groups, *J. Pure Appl. Algebra* **213** (2009), 1795–1803.

[3] A. Bahlekeh, F. Dembegioti, and O. Talelli, Gorenstein dimension and proper actions, *Bull. Lond. Math. Soc.* **41** (2009), 859–871.

[4] D. J. Benson and J. F. Carlson, Products in negative cohomology, *J. Pure Appl. Algebra* **82** (1992), 107–129.

[5] K. S. Brown, *Cohomology of groups*, Springer-Verlag, Berlin-Heidelberg-New York, 1982.

[6] F. Connolly and T. Koźniewski, Finiteness properties of classifying spaces of proper Γ-actions, *J. Pure Appl. Algebra* **41** (1986), 17–36.

[7] J. Cornick and P. H. Kropholler, Homological finiteness conditions for modules over group algebras, *J. London Math. Soc. (2)* **58** (1998), 49–62.

[8] F. Dembegioti and O. Talelli, On a relation between certain cohomological invariants, *J. Pure Appl. Algebra* **212** (2008), 1432–1437.

[9] M. N. Dyer, Homotopy classification of (π, m)-complexes, *J. Pure Appl. Algebra* **7** (1976), 249–282.

[10] I. Emmanouil, On certain cohomological invariants of groups, *Adv. Math.* **225** (2010), 3446–3462.

[11] T. V. Gedrich and K.W.Gruenberg, Complete cohomological functors of groups, *Topology Appl.*, **25** (1987), 203–223.

[12] F. Goichot, Homologie de Tate-Vogel équivariante, *J. Pure Appl. Algebra*, **82** (1992), 39–64.

[13] H. Holm, Gorenstein homological dimensions, *J. Pure Appl. Algebra*, **189** (2004), 167–193.

[14] B. M. Ikenaga, Homological dimension and Farrell cohomology, *J. Algebra* **87** (1984), 422–457.

[15] J. H. Jo, Projective complete cohomological dimension of a group, *Int. Math. Res. Not.* **13**, (2004), 621–636.

[16] J. H. Jo, A criterion for projective modules, *Comm. Alg.* **35**, (2007), 1577–1587.

[17] J. H. Jo, Complete homology and related dimensions of groups, *J. Group Theory* **12** (2009), 431–448.

[18] P. H. Kropholler, On groups of type $(\mathrm{FP})_\infty$, *J. Pure Appl. Algebra* **90** (1993), 55–67.

[19] G. Mislin, Tate cohomology for arbitrary groups via satellites, *Topology and its Appl.* **56** (1994), 293–300.

[20] G. Mislin and O.Talelli, On groups which act freely and properly on finite dimensional homotopy spheres, Computational and geometric aspects of modern algebra (Edinburgh, 1998), *London Math. Soc. Lecture Note Ser.* **275**, Cambridge Univ. Press, Cambridge, 2000.

[21] J. J. Rotman, *An introduction to homological algebra*, Second edition, Universitext, Springer, New York, 2009.

[22] O. Talelli, A characterization of cohomological dimension for a big class of groups, J. Algebra **326** (2011), 238–244.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1067 – 1082

Numerical solution of fourth order parabolic partial differential equation using parametric septic splines

Arshad Khan* and Talat Sultana†

Abstract

In this paper, we report three level implicit method of high accuracy schemes for the numerical solution of fourth order nonhomogeneous parabolic partial differential equation, that governs the behavior of a vibrating beam. Parametric septic spline is used in space and finite difference discretization in time. The linear stability of the presented method is investigated. The computed results for three examples are compared wherever possible with those already available in literature which shows the superiority of the proposed method.

Keywords: Parametric septic splines; Fourth order parabolic equation; Stability analysis; Vibrating beam; Finite difference scheme.

Received : 11.05.2015 *Accepted* : 09.09.2015 *Doi* : 10.15672/HJMS.20164513115

1. Introduction

In this paper, we consider the problem of undamped transverse vibration of a flexible straight beam in such a way that its support do not contribute to the strain energy of the system and is represented by the fourth order parabolic partial differential equation of the form

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^4 u}{\partial x^4} = f(x,t), \quad a \le x \le b, \quad t > 0, \tag{1.1}$$

subject to the initial conditions

$$\left.\begin{aligned} u(x,0) &= g_0(x), \quad a \le x \le b, \\ u_t(x,0) &= g_1(x), \quad a \le x \le b \end{aligned}\right\} \tag{1.2}$$

*Department of Mathematics, Jamia Millia Islamia, New Delhi-25, INDIA.
Email : akhan1234in@rediffmail.com
†Department of Mathematics, Lakshmibai College, Univ. of Delhi, New Delhi-52, INDIA.
Email : talat17m@gmail.com

and the boundary conditions

$$\left.\begin{aligned} u(a,t) = f_0(t), \quad u(b,t) = f_1(t), \quad t \geq 0, \\ u_{xx}(a,t) = q_0(t), \quad u_{xx}(b,t) = q_1(t), \quad t \geq 0, \end{aligned}\right\} \tag{1.3}$$

where u is the transverse displacement of the beam, $g_0(x),\ g_1(x),\ f_0(t),\ f_1(t),\ q_0(t),\ q_1(t)$ are continuous functions, t and x are time and distance variables respectively and $f(x,t)$ is dynamic driving force per unit mass [10,21,22,35].

Numerical methods for the solution of equation (1.1) have been carried out by many authors. Jain et al. [24], Danaee and Evans [1], Evans [8], Collatz [20], Andrade and Mckee [7] and Evans and Yousif [9] used finite difference methods for the numerical solution of transverse vibrations. Fairweather and Gourlay [13] derived explicit and implicit finite difference methods based on the semi explicit method. Parametric quintic spline methods are given by Rashidinia and Aziz [16] using nodal points. Collatz [20], Crandall [33], Jain [23], Conte [32], Jain et al. [24] and Todd [18] have proposed both explicit and implicit methods successfully. Five level, unconditionally stable, explicit method with truncation error of $O(k^2 + h^2 + (\frac{k}{h})^2)$ has been given by Albrecht [15]. All the above authors considered the homogeneous case of equation (1.1) with a constant coefficients. The analytical solution of homogeneous case of equation (1.1) has been obtained by using Adomain decomposition method by Wazwaz [3,4]. The nonhomogeneous problem with constant coefficients has been studied by Aziz et al. [34] based on parametric quintic spline and by Khan et al. [2] based on sextic spline by using nodal points. Khaliq and Twizell [6] and Twizell and Khaliq [11] developed a family of numerical methods, which are second order accurate in space and time, based on exact recurrence relation for a homogeneous case of equation (1.1) with a variable coefficient. Rashidinia and Mohammadi [17] developed three level implicit methods of $O(k^2 + h^4)$ and $O(k^4 + h^4)$ for the numerical solution of equation (1.1) with variable coefficients by using sextic spline. Wazwaz [5] has developed analytical solution of variable coefficient fourth order parabolic partial differential equation in two and three space dimensions. Khan et al. [25] have introduced a new algorithm, namely Laplace Decomposition Algorithm for fourth order parabolic partial differential equations with variable coefficients. In [26], the homotopy analysis method (HAM) is applied to solve such problems. Khan et al. [27] have studied numerical solution of time fractional fourth order partial differential equations with variable coefficients. They have implemented reliable series solution techniques namely, Adomian Decomposition Method (ADM) and He's Variational Iteration Method (HVIM). A family of B-spline methods have been considered by Caglar [14]. In [28], Mittal and Jain discussed two methods. In Method-I, they decomposed equation (1.1) in a system of second order equations and have solved them by using cubic B-spline and in Method-II, they have solved equation (1.1) directly by using quintic B-spline method. Talwar et al. [19] and Mohanty et al. [29-31] have used high accuracy spline scheme for solving one dimensional partial differential equations.

In this paper, parametric septic spline relations have been derived using nodal points. We have used parametric septic spline functions to develop a new numerical method for obtaining smooth approximations to the solution of nonhomogeneous parabolic partial differential equations dealing with vibrations of beams. In section 2, parametric septic spline and spline relations are developed. In section 3, we have presented the formulation of our method. Development of boundary equations are given in section 4. In section 5, truncation error and class of methods are given. Stability analysis is discussed in section 6. Finally in section 7, three examples are given to demonstrate the practical usefulness and superiority of our method.

2. Parametric septic spline

Let a set of grid points in the interval $[a, b]$ such that

$$x_j = a + jh,\ j = 0(1)N,\ h = \frac{(b-a)}{N}. \tag{2.1}$$

A function $S_\Delta(x,\tau)$ of class $C^6[a,b]$ which interpolates $u(x)$ at the mesh point x_j depends on a parameter τ, and as $\tau \to 0$ it reduces to septic spline $S_\Delta(x)$ in $[a,b]$ is termed as parametric septic spline function. Since the parameter τ can occur in $S_\Delta(x)$ in many ways such a spline is not unique.

If $S_\Delta(x,\tau) = S_\Delta(x)$ is a piecewise function satisfying the following differential equation in the interval $[x_{j-1}, x_j]$

$$\begin{aligned} S_\Delta^{(6)}(x) - \tau^2 S''_\Delta(x) &= (Q_j - \tau^2 M_j)\frac{x - x_{j-1}}{h} + (Q_{j-1} - \tau^2 M_{j-1})\frac{x_j - x}{h} \\ &= A_j z + A_{j-1}\bar{z}, \end{aligned} \tag{2.2}$$

where

$$z = \frac{x - x_{j-1}}{h},\ \bar{z} = 1 - z,\ A_i = Q_i - \tau^2 M_i,$$

$$S''_\Delta(x_i,\tau) = M_i,\quad S_\Delta^{(6)}(x_i,\tau) = Q_i,\quad i = j-1, j;\quad \tau > 0,$$

then it is termed as parametric septic spline II.
Solving equation (2.2), we get

$$\begin{aligned} S_\triangle(x) &= A_1 + A_2 x + A_3 \cosh\sqrt{\tau}x + A_4 \sinh\sqrt{\tau}x + A_5 \cos\sqrt{\tau}x + A_6 \sin\sqrt{\tau}x \\ &\quad - \frac{1}{\tau^2}\left\{(Q_j - \tau^2 M_j)\frac{(x - x_{j-1})^3}{6h} + (Q_{j-1} - \tau^2 M_{j-1})\frac{(x_j - x)^3}{6h}\right\} \end{aligned} \tag{2.3}$$

To develop the consistency relations between the value of spline and its derivatives at knots, let

$$\left.\begin{aligned} &S_\Delta(x_j) = u_j,\ S_\Delta(x_{j+1}) = u_{j+1}, \\ &S''_\Delta(x_j) = M_j,\ S''_\Delta(x_{j+1}) = M_{j+1}, \\ &S_\Delta^{(4)}(x_j) = F_j,\ S_\Delta^{(4)}(x_{j+1}) = F_{j+1}. \end{aligned}\right\} \tag{2.4}$$

To define spline in terms of u_j's, M_j's and F_j's, the coefficients introduced in Eq.(2.3) are calculated as

$$\begin{aligned}
A_1 &= u_{j-1} + \frac{h^2}{6\tau^2}(Q_{j-1} - \tau^2 M_{j-1}) - \frac{F_{j-1}}{\tau^2} \\
&\quad - \frac{x_{j-1}}{h}\left[(u_j - u_{j-1}) - \frac{h^2}{6\tau^2}(Q_{j-1} - \tau^2 M_{j-1}) + \frac{h^2}{6\tau^2}(Q_j - \tau^2 M_j) + \frac{1}{\tau^2}(F_{j-1} - F_j)\right], \\
A_2 &= \frac{1}{h}(u_j - u_{j-1}) + \frac{h}{6\tau^2}\left[-(Q_{j-1} - \tau^2 M_{j-1}) + (Q_j - \tau^2 M_j)\right] + \frac{1}{\tau^2 h}(F_{j-1} - F_j), \\
A_3 &= \frac{1}{\tau^2 \sinh\sqrt{\tau}h}\left[\frac{1}{2}\sinh\sqrt{\tau}x_j\left(F_{j-1} - \frac{Q_{j-1}}{\tau}\right) - \frac{1}{2}\sinh\sqrt{\tau}x_{j-1}\left(F_j - \frac{Q_j}{\tau}\right)\right. \\
&\quad \left. - \frac{1}{\tau}\sinh\sqrt{\tau}x_{j-1}Q_j + \frac{1}{\tau}\sinh\sqrt{\tau}x_j Q_{j-1}\right], \\
A_4 &= \frac{1}{\tau^2 \sinh\sqrt{\tau}h}\left[-\frac{1}{2}\cosh\sqrt{\tau}x_j\left(F_{j-1} - \frac{Q_{j-1}}{\tau}\right) + \frac{1}{2}\cosh\sqrt{\tau}x_{j-1}\left(F_j - \frac{Q_j}{\tau}\right)\right. \\
&\quad \left. + \frac{1}{\tau}\cosh\sqrt{\tau}x_{j-1}Q_j - \frac{1}{\tau}\cosh\sqrt{\tau}x_j Q_{j-1}\right], \\
A_5 &= \frac{1}{2\tau^2 \sinh\sqrt{\tau}h}\left[\sin\sqrt{\tau}x_j\left(F_{j-1} - \frac{Q_{j-1}}{\tau}\right) - \sin\sqrt{\tau}x_{j-1}\left(F_j - \frac{Q_j}{\tau}\right)\right], \\
A_6 &= \frac{1}{2\tau^2 \sinh\sqrt{\tau}h}\left[-\cos\sqrt{\tau}x_j\left(F_{j-1} - \frac{Q_{j-1}}{\tau}\right) + \cos\sqrt{\tau}x_{j-1}\left(F_j - \frac{Q_j}{\tau}\right)\right].
\end{aligned} \tag{2.5}$$

Substituting these values in (2.3), we get

$$\begin{aligned}
S_\Delta(x) &= z u_j + \bar{z} u_{j-1} + \frac{h^2}{6}\left[p(z)M_j + p(\bar{z})M_{j-1}\right] + \frac{h^4}{2}\left[r(z)F_j + r(\bar{z})F_{j-1}\right] \\
&\quad + \frac{h^6}{6}\left[q(z)Q_j + q(\bar{z})Q_{j-1}\right],
\end{aligned} \tag{2.6}$$

where

$$p_1(z) = z^3 - z, \quad q_1(z) = \frac{z}{\omega^4} - \frac{z^3}{\omega^4} + \frac{3\sinh\omega z}{\omega^6 \sinh\omega} - \frac{3\sin\omega z}{\omega^6 \sin\omega},$$
$$r_1(z) = \frac{-2z}{\omega^4} + \frac{\sinh\omega z}{\omega^4 \sinh\omega} + \frac{\sin\omega z}{\omega^4 \sin\omega} \quad \mathit{and} \quad \omega = \sqrt{\tau}h. \tag{2.7}$$

Applying the first, third and fifth derivative continuities at the knots, i.e. $S_\Delta^{(\mu)}(x_j^-) = S_\Delta^{(\mu)}(x_j^+)$, $\mu = 1, 3$ and 5, the following consistency relations are derived:

$$\begin{aligned}
M_{j+1} + 4M_j + M_{j-1} &= \frac{6}{h^2}(u_{j+1} - 2u_j + u_{j-1}) + 3h^2(\alpha_2 F_{j+1} + 2\beta_2 F_j + \alpha_2 F_{j-1}) \\
&\quad + h^4(\alpha_1 Q_{j+1} + 2\beta_1 Q_j + \alpha_1 Q_{j-1}), \quad j = 1(1)N-1.
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
M_{j+1} - 2M_j + M_{j-1} &= \frac{h^2}{6}[(1 - \omega^4\alpha_1)F_{j+1} + 2(2 - \omega^4\beta_1)F_j + (1 - \omega^4\alpha_1)F_{j-1}] \\
&\quad - \frac{h^4}{2}(\alpha_2 Q_{j+1} + 2\beta_2 Q_j + \alpha_2 Q_{j-1}), \quad j = 1(1)N-1.
\end{aligned} \tag{2.9}$$

$$h^2[(1-\omega^4\alpha_1)Q_{j+1}+2(2-\omega^4\beta_1)Q_j+(1-\omega^4\alpha_1)Q_{j-1}] =$$
$$3[(\omega^4\alpha_2+2)F_{j+1}+2(\omega^4\beta_2-2)F_j+(\omega^4\alpha_2+2)F_{j-1}], \quad j=1(1)N-1, \tag{2.10}$$

where

$$\begin{aligned}
\alpha_1 &= \frac{1}{\omega^4}+\frac{3}{\omega^5\sinh\omega}-\frac{3}{\omega^5\sin\omega},\\
\beta_1 &= \frac{2}{\omega^4}-\frac{3}{\omega^5}\coth\omega+\frac{3}{\omega^5}\cot\omega,\\
\alpha_2 &= \frac{-2}{\omega^4}+\frac{1}{\omega^3\sinh\omega}+\frac{1}{\omega^3\sin\omega},\\
\beta_2 &= \frac{2}{\omega^4}-\frac{1}{\omega^3}\coth\omega-\frac{1}{\omega^3}\cot\omega.
\end{aligned} \tag{2.11}$$

As $\tau\to 0$ that is $\omega\to 0$ then $(\alpha_1,\beta_1,\alpha_2,\beta_2)\to(\frac{-31}{2520},\frac{-4}{315},\frac{7}{180},\frac{2}{45})$.

Using equations (2.8)-(2.10), we obtain the following scheme

$$(e_1u_{j-3}+e_2u_{j-2}+e_3u_{j-1}+e_4u_j+e_3u_{j+1}+e_2u_{j+2}+e_1u_{j+3})$$
$$=\frac{h^4}{6}(p_1F_{j-3}+p_2F_{j-2}+p_3F_{j-1}+p_4F_j+p_3F_{j+1}+p_2F_{j+2}+p_1F_{j+3}),\ j=3(1)N-3, \tag{2.12}$$

where the coefficients (e_1,e_2,e_3,e_4) and (p_1,p_2,p_3,p_4) of the developed scheme are given by

$$\begin{aligned}
e_1 &= 1-3\omega^4\alpha_1+3\omega^8\alpha_1^2-\omega^{12}\alpha_1^3,\\
e_2 &= 4\omega^4\alpha_1-2\omega^4\beta_1-8\omega^8\alpha_1^2+4\omega^8\alpha_1\beta_1-2\omega^{12}\alpha_1^2\beta_1,\\
e_3 &= 7(1-\omega^4\alpha_1)^3-8(1-\omega^4\alpha_1)^2(2-\omega^4\beta_1),\\
e_4 &= 12(1-\omega^4\alpha_1)^2(2-\omega^4\beta_1)-8(1-\omega^4\alpha_1)^3,\\
p_1 &= c_1(1-\omega^4\alpha_1)^2,\\
p_2 &= 2c_1(1-\omega^4\alpha_1)(2-\omega^4\beta_1)+c_2(1-\omega^4\alpha_1)^2-3d_1(1-\omega^4\alpha_1)(2+\omega^4\alpha_2),\\
p_3 &= (c_1+c_3)(1-\omega^4\alpha_1)^2+6d_1(1-\omega^4\alpha_1)(2-\omega^4\beta_2)+2c_2(1-\omega^4\alpha_1)(2-\omega^4\beta_1)\\
&\quad -3d_2(1-\omega^4\alpha_1)(2+\omega^4\alpha_2),\\
p_4 &= 2c_2(1-\omega^4\alpha_1)^2-6d_1(1-\omega^4\alpha_1)(2+\omega^4\alpha_2)-6d_1(2-\omega^4\beta_1)(2-\omega^4\beta_2)\\
&\quad +2c_3(1-\omega^4\alpha_1)(2-\omega^4\beta_1)+6d_2(1-\omega^4\alpha_1)(2-\omega^4\beta_2).
\end{aligned} \tag{2.13}$$

Also

$$\begin{aligned}
c_1 &= \frac{1}{6}\omega^8\alpha_1^2-\frac{3}{2}\omega^4\alpha_2^2-\frac{1}{3}\omega^4\alpha_1-6\alpha_1-6\alpha_2+\frac{1}{6},\\
c_2 &= \frac{2}{3}\omega^8\alpha_1^2+\frac{1}{3}\omega^8\alpha_1\beta_1-18\omega^4\alpha_1\alpha_2-3\omega^4\alpha_2\beta_2-6\omega^4\alpha_2^2-2\omega^4\alpha_1-\frac{1}{3}\omega^4\beta_1\\
&\quad -12\alpha_1-6\beta_2+\frac{4}{3},\\
c_3 &= \frac{1}{3}\omega^8\alpha_1^2+\frac{4}{3}\omega^8\alpha_1\beta_1-36\omega^4\alpha_1\beta_2-12\omega^4\alpha_2\beta_2-3\omega^4\alpha_2^2-\frac{10}{3}\omega^4\alpha_1-\frac{4}{3}\omega^4\beta_1\\
&\quad +36\alpha_1+12\alpha_2+12\beta_2+3,\\
d_1 &= \omega^4\alpha_2\beta_1-\omega^4\alpha_1\beta_2+6\omega^4\alpha_1^2-10\alpha_1-2\alpha_2+2\beta_1+\beta_2,\\
d_2 &= 4\omega^4\alpha_2\beta_1-4\omega^4\alpha_1\beta_2+12\omega^4\alpha_1\beta_1-16\alpha_1-18\alpha_2-4\beta_1+4\beta_2.
\end{aligned} \tag{2.14}$$

As $\tau \to 0$ that is $\omega \to 0$, we have

(i)$(e_1, e_2, e_3, e_4) \longrightarrow (1, 0, -9, 16)$,
(ii)$(c_1, c_2, c_3, d_1, d_2) \longrightarrow \left(\frac{1}{140}, \frac{17}{14}, \frac{249}{70}, \frac{9}{140}, \frac{4}{35}\right)$,
(iii)$(p_1, p_2, p_3, p_4) \longrightarrow \left(\frac{1}{140}, \frac{6}{7}, \frac{1191}{140}, \frac{604}{35}\right)$.

[Remarks:] For these values our scheme reduces to the polynomial septic spline for fourth order boundary value problem which is given as equation (7) in G. Akram and S. S. Siddiqi [12].
Here, we have taken $(e_1, e_2, e_3, e_4) = (1, 0, -9, 16)$, therefore scheme (2.12) becomes

$$\begin{aligned} &p_1(F_{j-3} + F_{j+3}) + p_2(F_{j-2} + F_{j+2}) + p_3(F_{j-1} + F_{j+1}) + p_4 F_j \\ &= \frac{6}{h^4}\left[(u_{j-3} + u_{j+3}) - 9(u_{j-1} + u_{j+1}) + 16u_j\right], \ j = 3(1)N - 3. \end{aligned} \tag{2.15}$$

We can also write (2.15) as

$$\Lambda_x F_j = \frac{6}{h^4}(6\delta_x^4 + \delta_x^6)u_j, \tag{2.16}$$

where δ is the central difference operator and operator Λ_x for any function W is defined by

$$\Lambda_x W_j = p_1(W_{j-3} + W_{j+3}) + p_2(W_{j-2} + W_{j+2}) + p_3(W_{j-1} + W_{j+1}) + p_4 W_j. \tag{2.17}$$

3. Derivation of the method

Let the region $R = [a, b] \times [0, \infty)$ be discretized by a set of points $R_{h,k}$ which are the vertices of a grid points (x_j, t_m), where $x_j = jh, j = 0(1)N, Nh = b - a$ and $t_m = mk, \ m = 0, 1, 2, 3, ...$ The quantities h and k are mesh sizes in the space and time directions respectively.

We have developed an approximation for (1.1) in which the time derivative is replaced by a finite difference approximation and space derivative is replaced by the parametric septic spline function approximation. We need the following finite difference approximation for the time partial derivative of u :

$$\overline{u}_{tt_j}^m = k^{-2}\delta_t^2(1 + \sigma\delta_t^2)^{-1}u_j^m, \tag{3.1}$$

where σ is a parameter such that the finite difference approximation to the time derivative is $O(k^2)$ for arbitrary σ and $O(k^4)$ for $\sigma = 1/12$. u_j^m is the approximate solution of (1.1) at (x_j, t_m) and δ_t is the central difference operator with respect to t so that

$$\delta_t^2 u_j^m = u_j^{m+1} - 2u_j^m + u_j^{m-1}.$$

At the grid point (j, m) the differential equation may be discretized by

$$\overline{u}_{tt_j}^m + \overline{u}_{xxxx_j}^m = f_j^m, \tag{3.2}$$

where $\overline{u}_{xxxx_j}^m$ is the fourth order spline derivative at (x_j, t_m) denoted by $F_j^m = S_\Delta^{(4)}(x_j, t_m)$ with respect to the space variable $f_j^m = f(x_j, t_m)$. Using (3.1) and replacing fourth order spline derivative by F_j^m, we have

$$k^{-2}\delta_t^2(1 + \sigma\delta_t^2)^{-1}u_j^m + F_j^m = f_j^m. \tag{3.3}$$

Operating Λ_x on both sides of (3.3) and using (2.16), we obtain

$$\begin{aligned} &\delta_t^2[p_1(u_{j-3}^m + u_{j+3}^m) + p_2(u_{j-2}^m + u_{j+2}^m) + p_3(u_{j-1}^m + u_{j+1}^m) + p_4 u_j^m] + 6r^2(1 + \sigma\delta_t^2)(6\delta_x^4 + \delta_x^6)u_j^m \\ &= k^2(1 + \sigma\delta_t^2)[p_1(f_{j-3}^m + f_{j+3}^m) + p_2(f_{j-2}^m + f_{j+2}^m) + p_3(f_{j-1}^m + f_{j+1}^m) + p_4 f_j^m], \ j = 3(1)N - 3, \end{aligned} \tag{3.4}$$

where $r = \frac{k}{h^2}$ is the mesh ratio and p_1, p_2, p_3, p_4 are parameters. After simplifying the above equation, we obtain

$$\delta_t^2[(2p_1 + 2p_2 + 2p_3 + p_4) + (9p_1 + 4p_2 + p_3)\delta_x^2 + (6p_1 + p_2 + 36\sigma r^2)\delta_x^4 + (p_1 + 6\sigma r^2)\delta_x^6]u_j^m$$
$$+6r^2(6\delta_x^4+\delta_x^6)u_j^m = k^2(1+\sigma\delta_t^2)[p_1(f_{j-3}^m+f_{j+3}^m)+p_2(f_{j-2}^m+f_{j+2}^m)+p_3(f_{j-1}^m+f_{j+1}^m)+p_4 f_j^m],$$
$$j = 3(1)N-3. \tag{3.5}$$

This scheme (3.5) is finite difference in time and spline scheme in space variable, which on simplification can be written as

$$\begin{aligned}
&[P_1(u_{j-3}^{m+1} + u_{j+3}^{m+1}) + P_2(u_{j-2}^{m+1} + u_{j+2}^{m+1}) + P_3(u_{j-1}^{m+1} + u_{j+1}^{m+1}) + P_4 u_j^{m+1}]\\
&+[S_1(u_{j-3}^{m} + u_{j+3}^{m}) + S_2(u_{j-2}^{m} + u_{j+2}^{m}) + S_3(u_{j-1}^{m} + u_{j+1}^{m}) + S_4 u_j^{m}]\\
&+[P_1(u_{j-3}^{m-1} + u_{j+3}^{m-1}) + P_2(u_{j-2}^{m-1} + u_{j+2}^{m-1}) + P_3(u_{j-1}^{m-1} + u_{j+1}^{m-1}) + P_4 u_j^{m-1}]\\
&= K_1[p_1(f_{j-3}^{m+1} + f_{j+3}^{m+1}) + p_2(f_{j-2}^{m+1} + f_{j+2}^{m+1}) + p_3(f_{j-1}^{m+1} + f_{j+1}^{m+1}) + p_4 f_j^{m+1}]\\
&+K_2[p_1(f_{j-3}^{m} + f_{j+3}^{m}) + p_2(f_{j-2}^{m} + f_{j+2}^{m}) + p_3(f_{j-1}^{m} + f_{j+1}^{m}) + p_4 f_j^{m}]\\
&+K_1[p_1(f_{j-3}^{m-1} + f_{j+3}^{m-1}) + p_2(f_{j-2}^{m-1} + f_{j+2}^{m-1}) + p_3(f_{j-1}^{m-1} + f_{j+1}^{m-1}) + p_4 f_j^{m-1}],\ j = 3(1)N-3.
\end{aligned} \tag{3.6}$$

The final scheme (3.6) may be written in the schematic form as

$$\left.\begin{matrix} P_1 & P_2 & P_3 & P_4 & P_3 & P_2 & P_1 \\ S_1 & S_2 & S_3 & S_4 & S_3 & S_2 & S_1 \\ P_1 & P_2 & P_3 & P_4 & P_3 & P_2 & P_1 \end{matrix}\right\} u_j^m = \left.\begin{matrix} K_1p_1 & K_1p_2 & K_1p_3 & K_1p_4 & K_1p_3 & K_1p_2 & K_1p_1 \\ K_2p_1 & K_2p_2 & K_2p_3 & K_2p_4 & K_2p_3 & K_2p_2 & K_2p_1 \\ K_1p_1 & K_1p_2 & K_1p_3 & K_1p_4 & K_1p_3 & K_1p_2 & K_1p_1 \end{matrix}\right\} f_j^m,$$

where

$$\left.\begin{aligned} P_1 = p_1 + 6\sigma r^2, P_2 = p_2, P_3 = p_3 - 54\sigma r^2, P_4 = p_4 + 96\sigma r^2,\\ S_1 = -2P_1 + 6\sigma r^2, S_2 = -2P_2, S_3 = -2P_3 - 54\sigma r^2, S_4 = -2P_4 + 96\sigma r^2,\\ K_1 = \sigma k^2, K_2 = k^2(1 - 2\sigma). \end{aligned}\right\}$$

4. Development of boundary equations

The relation (3.6) gives $N-5$ linear algebraic equations in $N-1$ unknowns $u_j,\ j = 3(1)N-3$. We need four more equations, two at each end of the range of integration, for the direct computation of $u_j, j = 1(1)N-1$.

(i)
$$\frac{1392}{7}u_1^m - \frac{2340}{7}u_2^m + \frac{20320}{63}u_3^m - \frac{1320}{7}u_4^m + \frac{432}{7}u_5^m - \frac{548}{63}u_6^m = \frac{464}{9}u_0^m - \frac{80}{7}h^2(u_0^m)'', j = 1,$$

(ii)
$$-\frac{10280}{469}u_1^m + \frac{8746}{87}u_2^m - \frac{19309}{97}u_3^m + \frac{11287}{52}u_4^m - \frac{960}{7}u_5^m + \frac{9793}{207}u_6^m - \frac{464}{67}u_7^m$$
$$= -\frac{720}{469}h^2(u_0^m)'', j = 2,$$

(iii)
$$-\frac{464}{67}u_{N-7}^m + \frac{9793}{207}u_{N-6}^m - \frac{960}{7}u_{N-5}^m + \frac{11287}{52}u_{N-4}^m - \frac{19309}{97}u_{N-3}^m + \frac{8746}{87}u_{N-2}^m$$
$$-\frac{10280}{469}u_{N-1}^m = -\frac{720}{469}h^2(u_N^m)'', j = N-2,$$

(iv)
$$-\frac{548}{63}u_{N-6}^m + \frac{432}{7}u_{N-5}^m - \frac{1320}{7}u_{N-4}^m + \frac{20320}{63}u_{N-3}^m - \frac{2340}{7}u_{N-2}^m + \frac{1392}{7}u_{N-1}^m$$
$$= \frac{464}{9}u_N^m - \frac{80}{7}h^2(u_N^m)'', j = N-1.$$

For high accuracy formula of $O(k^6+h^{10})$, we use the following equations for approximating the boundary equations:
(i)

$$\frac{15928}{35}u_1^m-\frac{5141}{5}u_2^m+\frac{32427}{22}u_3^m-\frac{11441}{8}u_4^m+\frac{15875}{17}u_5^m-\frac{17741}{45}u_6^m+\frac{6919}{71}u_7^m-\frac{3853}{359}u_8^m$$

$$=\frac{6985}{72}u_0^m-\frac{12600}{761}h^2(u_0^m)'', j=1,$$

(ii) $$-\frac{4883}{125}u_1^m+\frac{18192}{79}u_2^m-\frac{38050}{61}u_3^m+\frac{21181}{21}u_4^m-\frac{275354}{261}u_5^m+\frac{138411}{191}u_6^m-\frac{25063}{79}u_7^m$$

$$+\frac{3707}{46}u_8^m-\frac{4288}{473}u_9^m=-\frac{967}{625}h^2(u_0^m)'', j=2,$$

(iii) $$-\frac{4288}{473}u_{N-9}^m+\frac{3707}{46}u_{N-8}^m-\frac{25063}{79}u_{N-7}^m+\frac{138411}{191}u_{N-6}^m-\frac{275354}{261}u_{N-5}^m+\frac{21181}{21}u_{N-4}^m$$

$$-\frac{38050}{61}u_{N-3}^m+\frac{18192}{79}u_{N-2}^m-\frac{4883}{125}u_{N-1}^m=-\frac{967}{625}h^2(u_N^m)'', j=N-2,$$

(iv) $$-\frac{3853}{359}u_{N-8}^m+\frac{6919}{71}u_{N-7}^m-\frac{17741}{45}u_{N-6}^m+\frac{15875}{17}u_{N-5}^m-\frac{11441}{8}u_{N-4}^m+\frac{32427}{22}u_{N-3}^m$$

$$-\frac{5141}{5}u_{N-2}^m+\frac{15928}{35}u_{N-1}^m=\frac{6985}{72}u_N^m-\frac{12600}{761}h^2(u_N^m)'', j=N-1.$$

5. Truncation error and class of methods

Expanding (3.5) in Taylor series in terms of $u(x_j,t_m)$ and its derivatives, we obtain the following relations

$$\begin{aligned}
\delta_x^6u(x_j,t_m) &= \Big[h^6D_x^6+\frac{1008}{8!}h^8D_x^8+\frac{105840}{10!}h^{10}D_x^{10}+\frac{1013760}{12!}h^{12}D_x^{12} \\
&\quad +\frac{9369360}{14!}h^{14}D_x^{14}+...\Big]u(x_j,t_m) \\
\delta_x^4u(x_j,t_m) &= \Big[h^4D_x^4+\frac{120}{6!}h^6D_x^6+\frac{505}{8!}h^8D_x^8+\frac{1016}{10!}h^{10}D_x^{10}+\frac{2040}{12!}h^{12}D_x^{12}+...\Big]u(x_j,t_m) \\
\delta_t^2u(x_j,t_m) &= \Big[-r^2h^4D_x^4+\frac{1}{12}r^4h^8D_x^8-\frac{1}{360}r^6h^{12}D_x^{12}+\frac{1}{20160}r^8h^{16}D_x^{16}+...\Big]u(x_j,t_m),
\end{aligned}\tag{5.1}$$

where $(D_t^2 + D_x^4)u(x_j, t_m) = f(x_j, t_m)$. Using (3.5) and (5.1), we obtain the truncation error

$$
\begin{aligned}
T_j^m \;=\;& \Big[(2p_1 + 2p_2 + 2p_3 + p_4) + (9p_1 + 4p_2 + p_3)\delta_x^2 + (6p_1 + p_2 + 36\sigma r^2)\delta_x^4 \\
&+ (p_1 + 6\sigma r^2)\delta_x^6\Big]\delta_t^2 u_j^m + 6r^2(6\delta_x^4 + \delta_x^6)u_j^m \\
&- k^2(1 + \sigma\delta_t^2)\Big[p_1(f_{j-3}^m + f_{j+3}^m) + p_2(f_{j-2}^m + f_{j+2}^m) + p_3(f_{j-1}^m + f_{j+1}^m) + p_4 f_j^m\Big] \\
=\;& \Big[o_1\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right) \\
&+ (o_2h^2D_x^2)\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right) \\
&+ (2o_2 + 24o_3)\frac{h^4D_x^4}{4!}\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right) \\
&+ (2o_2 + 120o_3 + 720o_4)\frac{h^6D_x^6}{6!}\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right) \\
&+ (2o_2 + 504o_3 + 10080o_4)\frac{h^8D_x^8}{8!}\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right) \\
&+ (2o_2 + 1016o_3 + 105840o_4)\frac{h^{10}D_x^{10}}{10!}\left(k^2D_t^2 + \frac{2k^4D_t^4}{4!} + \frac{2k^6D_t^6}{6!} + \frac{2k^8D_t^8}{8!} + ...\right)\Big]u_j^m \\
&+ 6r^2\left(\frac{144}{4!}h^4D_x^4 + \frac{1440}{6!}h^6D_x^6 + \frac{4032}{8!}h^8D_x^8 + \frac{111936}{10!}h^{10}D_x^{10} + \frac{10266000}{12!}h^{12}D_x^{12}\right. \\
&\left. + ...\right)u_j^m - \Big[o_1k^2 + o_1\sigma k^4D_t^2 + \frac{2o_1}{4!}\sigma k^6D_t^4 + \frac{2o_1}{6!}\sigma k^8D_t^6 + \frac{2o_1}{8!}\sigma k^{10}D_t^{10} + ... \\
&\left(o_2h^2D_x^2 + 2(81p_1 + 16p_2 + p_3)\frac{h^4D_x^4}{4!} + 2(729p_1 + 64p_2 + p_3)\frac{h^6D_x^6}{6!}\right. \\
&\left. + 2(6561p_1 + 256p_2 + p_3)\frac{h^8D_x^8}{8!} + ...\right) \times \\
&\left(k^2 + \sigma k^4D_t^2 + \frac{2\sigma k^6D_t^4}{4!} + \frac{2\sigma k^8D_t^6}{6!} + \frac{2\sigma k^{10}D_t^8}{8!} + ...\right)\Big](D_t^2 + D_x^4)u_j^m
\end{aligned}
$$

which may be written as

$$
\begin{aligned}
T_j^m \;=\;& \Big[(36 - o_1)r^2h^4D_x^4 + (12 - o_2)r^2h^6D_x^6 + \left(\frac{3}{5} - \frac{1}{12}(81p_1 + 16p_2 + p_3)\right)r^2h^8D_x^8 \\
&+ \left(\frac{583}{3150} - \frac{1}{360}(729p_1 + 64p_2 + p_3)\right)r^2h^{10}D_x^{10} \\
&+ \left(\frac{2565}{199584} - \frac{1}{20160}(6561p_1 + 256p_2 + p_3)\right)r^2h^{12}D_x^{12} + ... \\
&+ \left(\frac{1}{12} - \sigma\right)o_1k^4D_t^4 + \left(\frac{1}{12} - \sigma\right)o_2h^2k^4D_x^2D_t^4 + \left(\frac{1}{360} - \frac{\sigma}{12}\right)o_2h^2k^6D_x^2D_t^6 \\
&+ \left(\frac{h^4}{12}(o_2 + 12o_3) - o_1\sigma k^2 - \frac{h^4}{12}(81p_1 + 16p_2 + p_3)\right)k^4D_x^4D_t^2 \\
&+ \left(\frac{h^4}{360}(o_2 + 60o_3 + 360o_4) - o_2\sigma k^2 - \frac{\sigma h^4}{360}(729p_1 + 64p_2 + p_3)\right)h^2k^4D_x^6D_t^2
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\frac{h^4}{144}(o_2+12o_3)-\frac{1}{12}o_1\sigma k^2-\frac{\sigma h^4}{12}(81p_1+16p_2+p3)\right)k^4D_x^4D_t^4\\
&+\left(\frac{1}{20160}-\frac{\sigma}{360}\right)o_1k^8D_t^8+\left(\frac{1}{360}-\frac{\sigma}{12}\right)o_1k^6D_t^6+\left(\frac{2}{8!}-\frac{2\sigma}{6!}\right)o_2h^2k^8D_x^2D_t^8\\
&+\left(\frac{h^4}{8!}(2o_2+504o_3+10080o_4)-\frac{\sigma k^2}{12}(81p_1+16p_2+p_3)\right.\\
&\left.-\frac{2\sigma h^4}{8!}(6561p_1+256p_2+p_3)\right)h^4k^2D_x^8D_t^2\\
&\left.+\left(\frac{h^4}{4320}(o_2+12o_3)-\frac{1}{360}o_1\sigma k^2-\frac{\sigma h^4}{144}(81p_1+16p_2+p_3)\right)k^6D_x^4D_t^6+...\right]u_j^m\\
&+...,
\end{aligned}
\tag{5.2}
$$

where

$$
o_1=2p_1+2p_2+2p_3+p_4,\quad o_2=9p_1+4p_2+p_3,
$$
$$
o_3=6p_1+p_2+36\sigma r^2,\quad o_4=p_1+6\sigma r^2,\ D_x=\frac{\partial}{\partial x},\quad D_t=\frac{\partial}{\partial t}. \tag{5.3}
$$

For various values of parameters p_1, p_2, p_3, p_4 and σ, we obtain the following class of methods:

Case 1: If $36-o_1=0$, we obtain various schemes of $O(k^2+h^2)$ for arbitrary values of σ.

Case 2: If $36-o_1=0$ and $12-o_2=0$, we obtain various schemes of $O(k^2+h^4)$ for arbitrary values of σ.

Case 3: If $36-o_1=0,\ 12-o_2=0$, and $\frac{3}{5}-\frac{1}{12}(81p_1+16p_2+p_3)=0$, we obtain various schemes of $O(k^4+h^8)$ for $\sigma\neq\frac{1}{12}$ and $O(k^6+h^8)$ for $\sigma=\frac{1}{12}$.

Case 4: For $(p_1,p_2,p_3,p_4,\sigma)=\left(\frac{16}{75},\frac{-42}{25},\frac{84}{5},\frac{16}{3},\frac{1}{12}\right)$, we obtain a scheme of $O(k^6+h^{10})$.

6. Stability analysis

To investigate the stability analysis of the scheme (3.6), we use the Von Neumann method. We have assumed that the solution of (3.6) at the grid point (x_j, t_m) is of the form

$$
u_j^m=\xi^m e^{ji\theta}, \tag{6.1}
$$

where $i=\sqrt{-1},\ \theta$ is real and ξ in general is complex.

Substituting (6.1) in homogeneous part of (3.6), we obtain a characteristic equation

$$
X\xi^2+Y\xi+Z=0, \tag{6.2}
$$

where

$$
X=Z=P_1\cos 3\theta+P_2\cos 2\theta+P_3\cos\theta+2P_4,
$$
$$
Y=S_1\cos 3\theta+S_2\cos 2\theta+S_3\cos\theta+2S_4.
$$

Under the transformation $\xi=\frac{1+\eta}{1-\eta}$, equation (6.2) becomes

$$
(X-Y+Z)\eta^2+2(X-Z)\eta+(X+Y+Z)=0. \tag{6.3}
$$

The necessary and sufficient condition for $|\xi|\leq 1$ is that $X-Y+Z>0,\ X-Z>0$ and $X+Y+Z>0$.

The conditions $X-Z>0$ and $X+Y+Z>0$ are always satisfied for all real values of θ. From the condition $X-Y+Z>0$, we get that the scheme (3.6) is unconditionally stable if $\sigma\geq\frac{1}{4}$ and conditionally stable if $\sigma<\frac{1}{4}$ for all real values of p_1, p_2, p_3, p_4 and θ.

We summarized the above results in the following theorem:

Theorem: The scheme (3.6) for solving (1.1) is unconditionally stable if $\sigma \geq \frac{1}{4}$ and conditionally stable if $\sigma < \frac{1}{4}$. By using the Lax theorem, we can conclude that the present method is converge as long as stability criterion is satisfied.

7. Numerical results and discussions

We have applied the presented method on the fourth order parabolic partial differential equation and have considered one homogeneous and two nonhomogeneous examples. The proposed method (3.6) is implicit three level method based on parametric septic spline function.

Example 1: Consider a nonhomogeneous fourth order parabolic partial differential equation [2,9,17,34]

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^4 u}{\partial x^4} = (\pi^4 - 1)\sin \pi x \cos t, \quad 0 \leq x \leq 1,\ t \geq 0,$$

subject to the initial conditions

$$u(x,0) = \sin \pi x, \quad u_t(x,0) = 0, \quad 0 \leq x \leq 1$$

and the boundary conditions

$$u(0,t) = u(1,t) = u_{xx}(0,t) = u_{xx}(1,t) = 0, \quad t \geq 0.$$

The analytical solution for this example is

$$u(x,t) = \sin \pi x \cos t.$$

We have solved the above example with $h = 0.05$ and $k = 0.005$ giving $r = 2$ and by choosing $\sigma = \frac{1}{4}, \frac{1}{12}$ with $O(k^4 + h^8), O(k^6 + h^8)$ and $O(k^6 + h^{10})$ for arbitrary choices of parameters p_1, p_2, p_3 and p_4. All computations have been done over ten time steps. The absolute errors at particular points $x = 0.1, 0.2, 0.3, 0.4, 0.5$ and comparison with other existing methods [2,8,15,26] are tabulated in table 1. We repeat the computations for 16 time steps with $r = 0.5$.

Table 1. Absolute errors for example 1

Methods $(p_1, p_2, p_3, p_4, \sigma)$	r	Time steps	$x = 0.1$	$x = 0.2$	$x = 0.3$	$x = 0.4$	$x = 0.5$
$O(k^6 + h^{10})$							
$\left(\frac{16}{75}, \frac{-42}{25}, \frac{84}{5}, \frac{16}{3}, \frac{1}{12}\right)$	2	10	2.05(−6)	4.37(−7)	1.04(−7)	1.19(−8)	5.61(−8)
	0.5	16	4.04(−7)	7.81(−9)	8.34(−9)	4.56(−8)	5.10(−8)
$O(k^6 + h^8)$							
$\left(\frac{12}{25}, \frac{-82}{25}, \frac{104}{5}, 0, \frac{1}{4}\right)$	2	10	2.63(−6)	2.86(−8)	4.34(−8)	7.02(−8)	6.31(−8)
	0.5	16	6.80(−7)	3.52(−7)	9.65(−7)	1.39(−6)	1.53(−6)
$\left(\frac{43}{100}, \frac{-149}{150}, \frac{401}{20}, 1, \frac{1}{4}\right)$	2	10	2.76(−6)	4.65(−8)	3.55(−8)	6.71(−8)	6.06(−8)
	0.5	16	1.08(−6)	2.14(−7)	1.02(−6)	1.59(−6)	1.78(−6)
$O(k^4 + h^8)$							
$\left(\frac{12}{25}, \frac{-82}{25}, \frac{104}{5}, 0, \frac{1}{4}\right)$	2	10	9.02(−6)	8.37(−7)	3.92(−7)	1.60(−8)	9.53(−8)
	0.5	16	3.84(−6)	1.60(−6)	3.04(−7)	1.98(−6)	2.66(−6)
$\left(\frac{43}{100}, \frac{-149}{150}, \frac{401}{20}, 1, \frac{1}{4}\right)$	2	10	9.10(−6)	8.43(−7)	3.96(−7)	1.67(−8)	9.57(−8)
	0.5	16	2.14(−6)	1.10(−6)	1.20(−7)	1.07(−6)	1.51(−6)
[17]							
$O(k^2 + h^4), \sigma = \frac{1}{4}$,	2	10	3.09(−6)	4.04(−6)	1.65(−6)	2.44(−6)	2.73(−7)
	0.5	16	5.25(−7)	2.87(−7)	1.54(−7)	1.64(−7)	1.76(−7)
$O(k^4 + h^4), \sigma = \frac{1}{12}$,	2	10	2.91(−6)	1.73(−6)	1.60(−6)	2.33(−6)	2.60(−7)
	0.5	16	4.47(−7)	2.66(−7)	1.39(−7)	1.55(−7)	1.57(−7)
[2]	2	10	1.87(−6)	2.13(−5)	1.49(−5)	8.60(−6)	5.96(−6)
	0.5	16	9.07(−6)	7.79(−6)	2.75(−6)	1.01(−6)	2.59(−6)
[34]	2	10	1.80(−5)	2.00(−5)	1.40(−5)	8.30(−6)	5.70(−6)
	0.5	16	9.20(−6)	7.90(−6)	2.80(−6)	9.80(−7)	2.50(−6)
[9]	2	10	2.20(−4)	4.10(−4)	5.40(−4)	6.20(−4)	6.50(−4)
	0.5	16	2.50(−5)	4.70(−5)	6.60(−5)	7.80(−5)	8.20(−5)

Example 2: Consider a homogeneous fourth order parabolic partial differential equation [18,33,34]

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^4 u}{\partial x^4} = 0, \quad 0 \le x \le 1, \; t \ge 0,$$

subject to the initial conditions

$$u(x,0) = \frac{x}{12}(2x^2 - x^3 - 1), \quad u_t(x,0) = 0, \quad 0 \le x \le 1$$

and the boundary conditions

$$u(0,t) = u(1,t) = u_{xx}(0,t) = u_{xx}(1,t) = 0, \quad t \ge 0.$$

The analytical solution for this example is

$$u(x,t) = \sum_{s=0}^{\infty} d_s \sin(2s+1)\pi x \cos(2s+1)^2 \pi^2 t,$$

where

$$d_s = \frac{-8}{[(2s+1)^5 \pi^5]}.$$

We have solved this example with $h = 0.1$ for $\sigma = \frac{1}{4}, \frac{1}{12}$. The absolute errors at particular points $x = 0.1, 0.2, 0.3, 0.4, 0.5$ with $r = 2$ and 50 time steps of $O(k^4 + h^8), O(k^6 + h^8)$ and $O(k^6 + h^{10})$ for arbitrary choices of parameters p_1, p_2, p_3 and p_4 and comparison with other existing methods are tabulated in table 2. We repeat the computations for 100 time steps with $r = \sqrt{\frac{1}{6}}, \sqrt{\frac{7}{60}}$ and $r = \sqrt{\frac{1}{84}}$. We have also included results given by unconditionally stable method.

Table 2. Absolute errors for example 2

Methods $(p_1, p_2, p_3, p_4, \sigma)$	r^2	Time steps	$x = 0.1$	$x = 0.2$	$x = 0.3$	$x = 0.4$	$x = 0.5$
$O(k^6 + h^{10})$							
$\left(\frac{16}{75}, \frac{-42}{25}, \frac{84}{5}, \frac{16}{3}, \frac{1}{12}\right)$	4	50	2.66(−12)	5.48(−12)	1.12(−12)	6.07(−13)	7.42(−13)
	$\frac{1}{6}$	100	1.29(−12)	3.24(−13)	1.79(−12)	3.63(−12)	5.06(−12)
	$\frac{7}{60}$	100	1.79(−12)	4.02(−12)	1.38(−11)	2.40(−11)	3.11(−11)
	$\frac{1}{84}$	100	2.17(−13)	1.33(−13)	7.43(−13)	1.35(−12)	1.50(−12)
$O(k^6 + h^8)$							
$\left(\frac{12}{25}, \frac{-82}{25}, \frac{104}{5}, 0, \frac{1}{4}\right)$	4	50	1.69(−12)	1.11(−12)	6.43(−13)	9.50(−13)	1.54(−12)
	$\frac{1}{6}$	100	2.44(−12)	3.47(−13)	3.06(−12)	3.98(−12)	3.71(−12)
	$\frac{7}{60}$	100	3.00(−12)	4.44(−12)	1.55(−11)	2.45(−11)	3.24(−11)
	$\frac{1}{84}$	100	2.74(−13)	1.83(−13)	9.23(−13)	1.39(−12)	1.71(−12)
$\left(\frac{43}{100}, \frac{-149}{150}, \frac{401}{20}, 1, \frac{1}{4}\right)$	4	50	1.54(−12)	1.00(−12)	6.33(−13)	9.57(−13)	1.52(−12)
	$\frac{1}{6}$	100	3.06(−12)	6.12(−13)	3.26(−12)	4.06(−12)	3.46(−12)
	$\frac{7}{60}$	100	3.73(−12)	4.14(−12)	1.57(−11)	2.46(−11)	3.21(−11)
	$\frac{1}{84}$	100	3.39(−13)	1.56(−13)	9.46(−13)	1.40(−12)	1.68(−12)
$O(k^4 + h^8)$							
$\left(\frac{12}{25}, \frac{-82}{25}, \frac{104}{5}, 0, \frac{1}{4}\right)$	4	50	4.77(−13)	1.85(−12)	4.04(−12)	6.80(−12)	9.90(−12)
	$\frac{1}{6}$	100	4.64(−12)	1.43(−12)	3.58(−12)	4.21(−12)	4.01(−12)
$\left(\frac{43}{100}, \frac{-149}{150}, \frac{401}{20}, 1, \frac{1}{4}\right)$	4	50	5.59(−13)	3.20(−13)	5.36(−13)	1.24(−12)	2.71(−12)
	$\frac{1}{6}$	100	8.74(−12)	3.30(−12)	4.79(−12)	4.75(−12)	4.50(−12)
[18], $\sigma = \frac{1}{4}$	4	50	3.19(−4)	6.19(−4)	8.81(−4)	1.87(−3)	1.15(−3)
	$\frac{1}{6}$	100	2.61(−4)	4.43(−4)	5.47(−4)	6.08(−4)	6.33(−4)
[34], $\sigma = \frac{1}{4}$	4	50	3.21(−4)	5.77(−4)	7.24(−4)	7.89(−4)	8.10(−4)
	$\frac{1}{6}$	100	3.81(−4)	3.33(−4)	7.74(−4)	7.81(−4)	7.66(−4)
[33], $\sigma = \frac{1}{12}$	4	50	4.32(−4)	8.34(−4)	1.18(−4)	1.42(−3)	1.52(−3)
	$\frac{1}{6}$	100	2.30(−4)	4.08(−4)	5.40(−4)	6.56(−4)	7.02(−4)
[34], $\sigma = \frac{1}{12}$	4	50	1.00(−5)	5.00(−5)	1.73(−4)	3.33(−4)	4.10(−4)
	$\frac{1}{6}$	100	3.52(−4)	6.30(−4)	7.77(−4)	7.72(−4)	7.38(−4)
	$\frac{7}{60}$	100	1.38(−4)	1.74(−4)	9.05(−5)	3.40(−4)	9.60(−4)
	$\frac{1}{84}$	100	3.53(−5)	6.22(−5)	7.11(−5)	6.11(−5)	5.53(−5)

Example 3: Consider a nonhomogeneous fourth order parabolic partial differential

equation [32]

$$\frac{\partial^2 u}{\partial t^2}+\frac{\partial^4 u}{\partial x^4}=[24-x^2(1-x)^2]\cos t,\quad 0\le x\le 1,\ t>0,$$

subject to the initial conditions

$$u(x,0)=x^2(1-x)^2,\quad u_t(x,0)=0,\quad 0\le x\le 1$$

and the boundary conditions

$$u(0,t)=u(1,t)=0, u_{xx}(0,t)=u_{xx}(1,t)=2\cos t,\quad t\ge 0.$$

The analytical solution for this example is

$$u(x,t)=x^2(1-x)^2\cos t.$$

We have solved this example with $h=0.05$ for $\sigma=\frac{1}{4},\frac{1}{12}$. The absolute errors at particular points $x=0.1,0.2,0.3,0.4,0.5$ with $r=2$ and 10 time steps for $O(k^4+h^8), O(k^6+h^8)$ and $O(k^6+h^{10})$ using arbitrary choices of parameters p_1,p_2,p_3 and p_4 are tabulated in table 3. We repeat the computations for 16 time steps with $r=0.5$.

Table 3. Absolute errors for example 3

Methods (p_1,p_2,p_3,p_4,σ)	r	Time steps	$x=0.1$	$x=0.2$	$x=0.3$	$x=0.4$	$x=0.5$
$O(k^6+h^{10})$							
$\left(\frac{16}{75},\frac{-42}{25},\frac{84}{5},\frac{16}{3},\frac{1}{12}\right)$	2	10	3.16(−4)	2.74(−5)	4.18(−6)	8.92(−7)	1.17(−8)
	0.5	16	1.35(−5)	6.11(−6)	2.99(−6)	1.15(−5)	1.48(−5)
$O(k^6+h^8)$							
$\left(\frac{12}{25},\frac{-82}{25},\frac{104}{5},0,\frac{1}{4}\right)$	2	10	3.97(−4)	2.82(−5)	4.92(−6)	1.20(−6)	1.77(−7)
	0.5	16	1.12(−4)	2.17(−5)	6.58(−5)	9.42(−5)	1.04(−4)
$\left(\frac{43}{100},\frac{-149}{150},\frac{401}{20},1,\frac{1}{4}\right)$	2	10	3.98(−4)	2.90(−5)	5.00(−6)	1.16(−6)	1.26(−7)
	0.5	16	1.36(−4)	1.28(−5)	7.08(−5)	1.10(−4)	1.24(−4)
$O(k^4+h^8)$							
$\left(\frac{12}{25},\frac{-82}{25},\frac{104}{5},0,\frac{1}{4}\right)$	2	10	1.05(−3)	7.46(−5)	5.42(−5)	7.48(−6)	4.98(−6)
	0.5	16	1.50(−4)	9.27(−5)	1.83(−5)	1.20(−4)	1.61(−4)
$\left(\frac{43}{100},\frac{-149}{150},\frac{401}{20},1,\frac{1}{4}\right)$	2	10	1.04(−3)	7.43(−5)	5.43(−5)	7.46(−6)	5.00(−6)
	0.5	16	5.31(−5)	6.32(−5)	6.29(−7)	6.61(−5)	9.31(−5)

Conclusion

The parametric septic spline function have been developed to obtain three level implicit methods for solving fourth order parabolic partial differential equations. The developed methods are tested on three examples. The performance of these methods have been examined by comparing solution of homogeneous and nonhomogeneous fourth order parabolic partial differential equations with available results. In examples 1, 2 and 3, we have computed absolute errors at the points $x=0.1,0.2,0.3,0.4,0.5$ for the sake of comparison with our references and results are tabulated in tables 1-3. Tables show that our results are more accurate than the results obtained by previous methods.

Acknowledgement: The authors would like to thank the editor and reviewers for their valuable comments and suggestions to improve the paper.

References

[1] A. Danaee and D. J. Evans, *Hopscotch procedures for a fourth order parabolic partial differential equation*, Math. Comput. Simul., 24, (1982) 326-329.

[2] A. Khan, I. Khan and T. Aziz, *Sextic spline solution for solving a fourth order parabolic partial differential equation*, Int. J. Comput. Math., 7, 82 (2005) 871-879.

[3] A. M. Wazwaz, *On the solution of fourth order parabolic equation by the decomposition method*, Int. J. Comput. Math., 57 (1995) 213-217.

[4] A. M. Wazwaz, *Analytic treatment for variable coefficient fourth order parabolic partial differential equations*, Appl. Math. Comput., 123 (2001) 219-227.

[5] A. M. Wazwaz, *Exact solution for variable coefficients fourth order parabolic partial differential equations in higher dimensional spaces*, Appl. Math. Comput., 130 (2002) 415-424.

[6] A. Q. M. Khaliq and E. H. Twizell, *A family of second order methods for variable coefficient fourth order parabolic partial differential equations*, Int. J. Comput. Math., 23 (1987) 63-76.

[7] C. Andrade and S. Mckee, *High accuracy A. D. I. methods for fourth order parabolic equations with variable coefficients*, J. Comput. Appl. Math., 1, 3 (1977) 11-14.

[8] D. J. Evans, *A stable explicit method for the finite difference solution of a fourth order parabolic partial differential equation*, Comput. J., 8 (1965) 280-287.

[9] D. J. Evans and W. S. Yousif, *A note on solving the fourth order parabolic equation by the age method*, Int. J. Comput. Math., 40 (1991) 93-97.

[10] D. J. Gorman, *Free Vibrations Analysis of Beams and Shafts*, John Wiley and Sons, New York, (1975).

[11] E. H. Twizell and A. Q. M. Khaliq, *A difference scheme with high accuracy in time for fourth order parabolic equations*, Comput. Methods Appl. Mech. Eng., 41 (1983) 91-104.

[12] G. Akram and S. S. Siddiqi, *End conditions for interpolatory septic splines*, Int. J. Comput. Math., 12, 82 (2005) 1525-1540.

[13] G. Fairweather and A. R. Gourlay, *Some stable difference approximations to a fourth order parabolic partial differential equation*, Math. Comput., 21 (1967) 1-11.

[14] H. Caglar and N. Caglar, *Fifth degree B-spline solution for a fourth order parabolic partial differential equations*, Appl. Math. Comput., 201 (2008) 597-603.

[15] J. Albrecht, *Zum differenzenverfahren bei parabolischen differentialgleichungen*, Z. Angew. Math. Mech., 37 (1957) 202-212.

[16] J. Rashidinia and T. Aziz, *Spline solution of fourth order parabolic partial differential equations*, Int. J. Appl. Sci. Comput., 2, 5 (1998) 139-148.

[17] J. Rashidinia and R. Mohammadi, *Sextic spline solution of variable coefficient fourth order parabolic equations*, Int. J. Comput. Math., 15, 87 (2010) 3443-3454.

[18] J. Todd, *A direct approach to the problem of stability in the numerical solution of partial differential equations*, Commum. Pure Appl. Math., 9 (1956) 597-612.

[19] J. Talwar, R. K. Mohanty and S. Singh, *A new spline in compression approximation for one space dimensional quasilinear parabolic equations on a variable mesh*, Appl. Math. Comput., 260 (2015) 82-96.

[20] L. Collatz, *Hermitian methods for initial value problems in partial differential equations*, in: Topics in Numerical Analysis, J. J. H. Miller, ed., Academic Press, New York, (1973) 41-61.

[21] L. Meirovitch, *Principles and Techniques of Vibrations*, Prentice Hall Inc., New Jersey, (1997).

[22] L. Meirovitch, *Analytical Methods in Vibrations*, Macmillan, New York, (1997).

[23] M. K. Jain, *Numerical Solution of Differential Equations*, 2nd ed., Wiley Eastern, New Delhi, India, (1984).

[24] M. K. Jain, S. R. K. Iyengar and A. G. Lone, *Higher order difference formulas for a fourth order parabolic partial differential equation*, Int. J. Numer. Methods Eng., 10 (1976) 1357-1367.

[25] M. Khan, M. A. Gondal and Y. Khan, *An efficient laplace decomposition algorithm for fourth order parabolic partial differential equations with variable coefficients*, World Appl. Sci. J., 12, 13 (2011) 2463-2466.

[26] N. A. Khan, A. Ara, M. Afzal and A. Khan, *Analytical aspect of fourth order parabolic partial differential equations with variable coefficients*, Math. Comput. Applic., 3, 15 (2010) 481-489.

[27] N. A. Khan, N. Khan, M. Ayaz, A. Mahmood and N. Fatima, *Numerical study of time fractional fourth order differential equations with variable coefficients*, J. King Saud Univ. (Sci.), 23 (2011) 91-98.

[28] R. C. Mittal and R. K. Jain, *B-splines methods with redefined basis functions for solving fourth order parabolic partial differential equations*, Appl. Math. Comput., 217 (2011) 9741-9755.

[29] R.K. Mohanty and M.K. Jain, *High accuracy cubic spline alternating group explicit methods for 1D quasi-linear parabolic equations*, Int. J. Comput. Math., 86 (2009) 1556-1571.

[30] R.K. Mohanty and V. Gopal, *High accuracy cubic spline finite difference approximation for the solution of one-space dimensional non-linear wave equations*, Appl. Math. Comput., 218 (2011) 4234-4244.

[31] R.K. Mohanty and V. Gopal, *High accuracy non-polynomial spline in compression method for one-space dimensional quasi-linear hyperbolic equations with significant first order space derivative term*, Appl. Math. Comput., 238 (2014) 250-265.

[32] S. D. Conte, *A stable implicit finite difference approximation to a fourth order parabolic equation*, J. Assoc. Comput. Mech., 4 (1957) 18-23.

[33] S. H. Crandall, *Numerical treatment of a fourth order partial differential equations*, J. Assoc. Comput. Mech., 1 (1954) 111-118.

[34] T. Aziz, A. Khan and J. Rashidinia, *Spline methods for the solution of fourth order parabolic partial differential equations*, Appl. Math. Comput., 167 (2005) 153-166.

[35] W. T. Thomson, *Theory of Vibration with Applications*, 3rd ed., Prentice Hall, Englewood, New Jersey, (1988).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1083 – 1090

Morita equivalence based on Morita context for arbitrary semigroups

Hongxing Liu*

Abstract

In this paper, we study the Morita context for arbitrary semigroups. We prove that, for two semigroups S and T, if there exists a Morita context (S, T, P, Q, τ, μ) (not necessary unital) such that the maps τ and μ are surjective, the categories US-FAct and UT-FAct are equivalent. Using this result, we generalize Theorem 2 in [2] to arbitrary semigroups. Finally, we give a characterization of Morita context for semigroups.

Keywords: semigroup, S-act, Morita context, functor, category.

2000 AMS Classification: 20M50

Received : 25.07.2015 *Accepted* : 16.09.2015 *Doi* : 10.15672/HJMS.20164520080

1. Introduction

Morita theory characterizes equivalences between module categories over rings with 1. Kyuno [5] studied Morita theory for rings without 1. Knauer [4] and Banschewski [1] independently generalized this theory to monoids. Banschewski [1] proved that for two semigroups S and T, if the two categories S-Act and T-Act are equivalent, then S is isomorphic to T. Talwar [8] extended Morita theory to semigroups with local units. He proved that for two semigroups with local units S and T, the two categories FS-Act and FT-Act are equivalent $\Longleftrightarrow$ there is a unitary Morita context (S, T, P, Q, τ, μ) such that the maps τ and μ are surjective, where $FS\text{-Act} = \{M \in S\text{-Act} | SM = M \text{ and } S \otimes \mathrm{Hom}_S(S, M) \cong M\}$. In [7], Talwar investigated strong Morita equivalence for factorisable semigroups. He got that if there is a unitary Morita context (S, T, P, Q, τ, μ) such that the maps τ and μ are surjective, then S and T are strongly Morita equivalent. Chen and Shum [2] showed that, for factorisable semigroups S and T, if there exists a unitary Morita context (S, T, P, Q, τ, μ) such that the maps τ and μ are surjective, then the categories US-FAct and UT-FAct are equivalent.

In this paper, we mainly use the techniques of paper [5] to study the corresponding problems for arbitrary semigroups. The paper is constructed as follows: In Section 2, we recall some basic notions; In Section 3, we give the main results of the paper. We prove

*School of Mathematical Sciences, Shandong Normal University, 250014, Jinan, P. R. China, Email: lhxshanda@163.com

that, for two semigroups S and T, if there exists a Morita context (S, T, P, Q, τ, μ) (not necessary unital) such that the maps τ and μ are surjective, the categories US-FAct and UT-FAct are equivalent. Also, we extend Theorem 2 in [2] to arbitrary semigroups. In Section 4, we give a characterization of Morita contexts for semigroups.

2. Preliminaries

Let S be a semigroup. A set M is a left S-act if there is a function from $S \times M$ to M, denoted $(s, m) \to sm$, such that $(st)m = s(tm)$ $(\forall s, t \in S, m \in M)$. If M is a left S-act, we write ${}_SM$. A left S-act M is said to be unitary if $M = SM$. Similarly, we can define right acts over semigroups.

Let M and N be two S-acts. A map $f : M \to N$ is an S-morphism if f satisfies $f(sm) = sf(m)$, $(\forall m \in M, s \in S)$. Let $\mathrm{Hom}_S(M, N)$ denote the set of all S-morphisms from ${}_SM$ to ${}_SN$. Denote by $\mathrm{End}_S(M)$ the set of all S-morphisms from M to itself. Let S-Act denote the category of left acts over a semigroup S.

The unital left S-acts together with the S-morphisms form a full subcategory of S-Act, which we shall denote by US-Act.

Let S and T be two semigroups. An S-T-biact is a set M which is both left S-act and right T-act and $(sm)t = s(mt)$ for all $s \in S, t \in T$ and all $m \in M$. A biact is said to be unitary if it is left and right unitary. If M and N are S-T-biact, a map $f : M \to N$ is called biact morphism if f satisfies $f(sm) = sf(m)$ and $f(mt) = f(m)t$ for all $m \in M, s \in S, t \in T$.

Let S be a semigroup and $M \in S$-Act. An equivalence R on S is a congruence if for all $s, t, a \in S$,

$$(s, t) \in R \Rightarrow (as, at) \in R, (sa, ta) \in R.$$

An equivalence ρ on ${}_SM$ is a congruence if for all $s \in S, m, n \in M$,

$$(m, n) \in \rho \Rightarrow (sm, sn) \in \rho.$$

If ρ is a congruence on M, then M/ρ is also a left S-act. The act M/ρ is called a quotient act. Let ϵ be the identity congruence on M.

Let S be a semigroup and $M \in S$-Act. According to [2], we use the following notations.

$$\zeta_M = \{(x, y) \in M \times M | sx = sy, \forall s \in S\};$$

$$US\text{-FAct} = \{M \in US\text{-Act} | \zeta_M = \epsilon\}.$$

Obviously, ζ_M is a congruence on M.

For a right S-act A_S and a left S-act ${}_SB$, the tensor product $A \otimes_S B$ exists. In fact, $A \otimes_S B = (A \times B)/\sigma$, where σ is the equivalence on $A \times B$ generated by

$$\mathcal{R} = \{((xs, y), (x, sy)) | a \in A, b \in B, s \in S\}.$$

We denote the element $(x, y)\sigma$ of $A \otimes_S B$ by $x \otimes y$.

By Proposition 1.4.10 of [3], we have that $a \otimes b = c \otimes d \iff (a, b) = (c, d)$ or there is a sequence

$$(a, b) = (x_1, y_1) \to (x_2, y_2) \to \cdots \to (x_n, y_n) = (c, d)$$

such that either $((x_i, y_i), (x_{i+1}, y_{i+1})) \in T$ or $((x_{i+1}, y_{i+1}), (x_i, y_i)) \in T$, where $1 \leq i \leq n - 1$.

If A is a right S-act and B is an S-T-biact, then $A \otimes_S B$ is a right T-act with

$$(a \otimes b)t = a \otimes bt;$$

similarly, if A is a T-S-biact and B is a left S-biact, then $A \otimes_S B$ is a left T-act with

$$t(a \otimes b) = ta \otimes b$$

(Proposition 3.1, [8]).

3. Morita equivalence for semigroups

In this section, S and T are arbitrary semigroups. If there exists a Morita context (S,T,P,Q,τ,μ), we shall prove that the two categories $F: US\text{-FAct} \rightleftharpoons UT\text{-FAct}: G$ are equivalent. Furthermore, if (S,T,P,Q,τ,μ) is unital, we get that $F \cong (Q\otimes -)/\zeta_{(Q\otimes -)}$ and $G \cong (P\otimes -)/\zeta_{(P\otimes -)}$. This generalizes Theorem 2 in [2].

3.1. Definition. [8] Let S and T be two semigroups. If there exist sets P and Q, such that
1) P is an S-T-biact, Q is a T-S-biact;
2) there are biact morphisms $\tau: P\otimes_T Q \to S$ and $\mu: Q\otimes_S P \to T$ written correspondingly as

$$\tau(p\otimes q) = <p,q>, \quad \mu(q\otimes p) = [q,p]$$

such that

$$<p_1,q>\cdot p_2 = p_1\cdot[q,p_2], \quad [q_1,p]\cdot q_2 = q_1\cdot <p,q_2>$$

for each $p,p_1,p_2 \in P, q,q_1,q_2 \in Q$. Then (S,T,P,Q,τ,μ) is called a Morita context.

By Proposition 3.1 in [8], we have $\tau(p\otimes q)s = \tau((p\otimes q)s) = \tau(p\otimes qs)$, where $p \in P, q\in Q, s\in S$. We will use this fact in the proof of Lemma 3.2 and Lemma 3.4.

3.2. Lemma. *Let (S,T,P,Q,τ,μ) be a Morita context, where τ and μ are surjective. Then*
1) For all $M \in US$-FAct, set $U = Q\times M$. Then $\widetilde{(Q,M)} = (Q\times M)/\rho_{(Q\times M)} \in UT$-FAct, where $\rho_{Q\times M} = \{((q,m),(q^{'},m^{'})) \in U\times U | \tau(p\otimes q)m = \tau(p\otimes q^{'})m^{'}, \forall p\in P\}$.
2) For all $N \in UT$-FAct, set $V = P\times N$. Then $\widetilde{(P,N)} = (P\times N)/\rho_{(P\times N)} \in UT$-FAct, where $\rho_{P\times N} = \{((p,n),(p^{'},n^{'})) \in V\times V | \mu(q\otimes p)n = \mu(q\otimes p^{'})n^{'}, \forall q\in Q\}$.

Proof *1) i) Clearly, ρ_U is an equivalence on U. Set $\widetilde{(Q,M)} = U/\rho_U$. Denote by $\overline{(r,m)}$ the equivalence class $(r,m)\rho_U$. For $t\in T$, we can write $t = \mu(q\otimes p)$ since μ is surjective.*

For all $\overline{(q,m)} \in \widetilde{(Q,M)}, \mu(q^{'}\otimes p^{'}) \in T$, define

$$\mu(q^{'}\otimes p^{'})\overline{(q,m)} = \overline{(q^{'},\tau(p^{'}\otimes q)m)}.$$

If $\overline{(q_1,m_1)} = \overline{(q_2,m_2)}$, for all $p\in P$, we have $<p,q_1>m_1 = <p,q_2>m_2$. Hence, the definition is independent of the choice of equivalence class representative.

If $\mu(q_1\otimes p_1) = \mu(q_2\otimes p_2)$, for all $x\in P$, we have

$$\begin{aligned} <x,q_1><p_1,q>m &= <x,q_1<p_1,q>>m = <x,[q_1,p_1]q>m \\ &= <x,[q_2,p_2]q>m = <x,q_2><p_2,q>m. \end{aligned}$$

Hence,

$$\overline{(q_1,<p_1,q>m)} = \overline{(q_2,<p_2,q>m)}.$$

Therefore, the definition is well-defined.

For all $\mu(q_1\otimes p_1),\mu(q_2\otimes p_2) \in T, \overline{(q,m)} \in \widetilde{(Q,M)}$, we have

$$(\mu(q_1\otimes p_1)\mu(q_2\otimes p_2))\overline{(q,m)} = \mu([q_1,p_1]q_2\otimes p_2)\overline{(q,m)} = \overline{([q_1,p_1]q_2,\tau(p_2\otimes q)m)}$$

and

$$\mu(q_1\otimes p_1)(\mu(q_2\otimes p_2)\overline{(q,m)}) = \mu(q_1\otimes p_1)\overline{(q_2,\tau(p_2\otimes q)m)} = \overline{(q_1,\tau(p_1\otimes q_2)\tau(p_2\otimes q)m)}.$$

Then $(\mu(q_1\otimes p_1)\mu(q_2\otimes p_2))\overline{(q,m)} = \mu(q_1\otimes p_1)(\mu(q_2\otimes p_2)\overline{(q,m)})$. This means that $\widetilde{(Q,M)}$ is a left T-Act.

ii) Suppose $(\overline{(q,m)},\overline{(q^{'},m^{'})}) \in \zeta_{\widetilde{(Q,M)}}$. For all $y\in Q, x\in P$, we have

$$\mu(y\otimes x)\overline{(q,m)} = \mu(y\otimes x)\overline{(q^{'},m^{'})}.$$

That is,

$$\overline{(y, \tau(x \otimes q)m)} = \overline{(y, \tau(x \otimes q^{'})m^{'})}.$$

This implies that

$$\tau(p \otimes y)\tau(x \otimes q)m = \tau(p \otimes y)\tau(x \otimes q^{'})m^{'},$$

for all $p \in P$. *Since* $M \in US$-*FAct, we have*

$$\tau(x \otimes q)m = \tau(x \otimes q^{'})m^{'}.$$

For arbitrary of x, *we get that* $\overline{(q, m)} = \overline{(q^{'}, m^{'})}$.

iii) For all $m \in M$, *since* $M = SM$ *and* τ *is surjective, we have* $m = \tau(p \otimes q^{'})m^{'}$, *where* $m^{'} \in M$. *For all* $\overline{(q, m)} \in \widetilde{(Q, M)}$, *we have*

$$\overline{(q, m)} = \overline{(q, \tau(p \otimes q^{'})m^{'})} = \mu(q \otimes p)\overline{(q^{'}, m^{'})} \in T\widetilde{(Q, M)}.$$

Hence, we get $T\widetilde{(Q, M)} = \widetilde{(Q, M)}$. *Therefore,* $\widetilde{(Q, M)} \in UT$-*FAct.*

2) For all $\overline{(p, n)} \in \widetilde{(P, N)}, \tau(p^{'} \otimes q^{'}) \in S$, *define*

$$\tau(p^{'} \otimes q^{'})\overline{(p, n)} = \overline{(p^{'}, \mu(q^{'} \otimes p)n)}.$$

Similarly, we can prove $\widetilde{(P, N)} \in US$-*FAct.*

3.3. Theorem. *Let* S *and* T *be two semigroups. If* (S, T, P, Q, τ, μ) *is a Morita context with* τ *and* μ *surjective, then we have the category equivalence* $F : US\text{-}FAct \rightleftharpoons UT\text{-}FAct : G$, *where* $F = (Q \times -)/\rho_{(Q\times -)}$ *and* $G = (P \times -)/\rho_{(P\times -)}$.

Proof Let $f : M \longrightarrow N$ be an S-morphism, where $M, N \in US$-FAct. Define $\tilde{f} : \widetilde{(Q, M)} \longrightarrow \widetilde{(Q, N)}$ by

$$\tilde{f}(\overline{(q, m)}) = \overline{(q, f(m))}.$$

Suppose $\overline{(q, m)} = \overline{(q^{'}, m^{'})}$. For all $p \in P$, we have $\tau(p \otimes q)m = \tau(p \otimes q^{'})m^{'}$. This implies that $f(\tau(p \otimes q)m) = f(\tau(p \otimes q^{'})m^{'})$. Since f is an S-morphism, it follows that $\tau(p \otimes q)f(m) = \tau(p \otimes q^{'})f(m^{'})$. Hence, $\overline{(q, f(m))} = \overline{(q^{'}, f(m^{'}))}$. This proves that $\tilde{f}$ is well-defined.

It is easy to check that $\tilde{f}$ is a left T-morphism.

Let $f : U \longrightarrow V$ and $g : V \longrightarrow W$ be two S-morphisms, where $U, V, W \in US$-FAct. Let $\tilde{f} : \widetilde{(Q, U)} \longrightarrow \widetilde{(Q, V)}$ and $\tilde{g} : \widetilde{(Q, V)} \longrightarrow \widetilde{(Q, W)}$ be T-morphisms determined by f and g respectively. Then $\widetilde{gf} = \tilde{g}\tilde{f}$. In fact, since $gf : U \longrightarrow W$ is an S-morphism, we have a T-morphism $\widetilde{gf} : \widetilde{(Q, U)} \longrightarrow \widetilde{(Q, W)}$. This implies that $\mathrm{dom}(\widetilde{gf}) = \widetilde{(Q, U)} = \mathrm{dom}(\tilde{g}\tilde{f})$. For all $\overline{(q, u)} \in \widetilde{(Q, U)}$, we have

$$\widetilde{gf}(\overline{(q, u)}) = \overline{(q, gf(u))} = \tilde{g}\overline{(q, f(u))} = \tilde{g}\tilde{f}\overline{(q, u)}.$$

Define $F : US\text{-FAct} \longrightarrow UT\text{-FAct}$ by $F(M) = (Q \times M)/\rho_{(Q\times M)} = \widetilde{(Q, M)}$ and $F(f) = \tilde{f}$, for all $M, N \in US\text{-FAct}, f \in \mathrm{Hom}_S(M, N)$. Then F is a functor.

Similarly, for $U, V \in UT$-FAct, if $f : U \to V$ is a T-morphism, we can define S-morphism $\bar{f} : \widetilde{(P, U)} \longrightarrow \widetilde{(P, V)}$ by

$$\bar{f}(\overline{(p, u)}) = \overline{(p, f(u))}.$$

Also, for $U, V, W \in UT$-FAct, if $f : U \longrightarrow V$ and $g : V \longrightarrow W$ be two T-morphisms, then $\overline{gf} = \bar{g}\bar{f}$.

We can define a functor $G : UT\text{-FAct} \longrightarrow US\text{-FAct}$ by $G(N) = (P \times N)/\rho_{(P\times N)} = \widetilde{(P, N)}$ and $G(f) = \bar{f}$, for all $N \in UT\text{-FAct}, g \in \mathrm{Hom}_T(M, N)$.

For $M \in US$-FAct, we have

$$GF(M) = G(\widetilde{(Q,M)}) = (P,\widetilde{\widetilde{(Q,M)}}).$$

Define $\eta_M : M \longrightarrow (P,\widetilde{\widetilde{(Q,M)}})$ by

$$\tau(p\otimes q)m \mapsto \overline{(p,\overline{(q,m)})}.$$

For all $p,p^{'} \in P, q,q^{'} \in Q, m,m^{'} \in M$, we have

$$\begin{array}{rl} & \tau(p\otimes q)m = \tau(p^{'}\otimes q^{'})m^{'} \\ \Leftrightarrow & \tau(x\otimes y)\tau(p\otimes q)m = \tau(x\otimes y)\tau(p^{'}\otimes q^{'})m^{'}, \\ & \text{for all } x\in P, y\in Q \text{ (since } M\in US\text{-FAct)} \\ \Leftrightarrow & \overline{(y,\tau(p\otimes q)m)} = \overline{(y,\tau(p^{'}\otimes q^{'})m^{'})}, \text{ for all } y\in Q \\ \Leftrightarrow & \mu(y\otimes p)\overline{(q,m)} = \mu(y\otimes p^{'})\overline{(q^{'},m)}, \text{ for all } y\in Q \\ \Leftrightarrow & \overline{(p,\overline{(q,m)})} = \overline{(p^{'},\overline{(q^{'},m^{'})})}. \end{array}$$

This shows that η_M is well-defined and injective. It is obvious that η_M is surjective. For $m\in M$, write $m = \tau(p^{'}\otimes q^{'})m^{'}$, where $p^{'}\in P, q^{'}\in Q, m^{'}\in M$. For all $p\in P, q\in Q$, we have

$$\begin{array}{rcl} \eta_M(\tau(p\otimes q)\tau(p^{'}\otimes q^{'})m^{'}) & = & \overline{(p,\overline{(q,\tau(p^{'}\otimes q^{'})m^{'})})} = \overline{(p,\mu(q\otimes p^{'})\overline{(q^{'},m^{'})})} \\ & = & \tau(p\otimes q)\overline{(p^{'},\overline{(q^{'},m^{'})})} = \tau(p\otimes q)\eta_M(\tau(p^{'}\otimes q^{'})m^{'}). \end{array}$$

Hence, η_M is an S-isomorphism.

Let $f : M \longrightarrow N$ be an S-morphism. For $m = \tau(p\otimes q)m^{'} \in M$, we have

$$\begin{array}{rcl} GF(f)\eta_M(m) & = & GF(f)\eta_M(\tau(p\otimes q)m^{'}) = GF(f)(\overline{(p,\overline{(q,m^{'})})}) \\ & = & \overline{(p,F(f)(\overline{(q,m^{'})}))} = \overline{(p,\overline{(q,f(m^{'}))})} = \eta_N f(m). \end{array}$$

Hence, we have the following commutative diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ \eta_M\downarrow & & \downarrow\eta_N \\ GF(M) & \xrightarrow{GF(f)} & GF(N). \end{array}$$

Therefore, $GF \cong 1_{US\text{-FAct}}$.

Similarly, we can prove that $FG \cong 1_{UT\text{-FAct}}$. This get the desired result. □

3.4. Lemma. *Let (S,T,P,Q,τ,μ) be a Morita context and $M\in US$-FAct. If $q_1\otimes m_1 = q_2\otimes m_2 \in Q\otimes M$, we have $\overline{(q_1,m_1)} = \overline{(q_2,m_2)}$.*

Proof 1) Suppose $((q_1,m_1),(q_2,m_2)) \in T$. Without loss of generality, We suppose $q_2 = q_1 s, m_1 = sm_2$, where $s\in S$. Then

$$\tau(p\otimes q_1)m_1 = \tau(p\otimes q_1)sm_2 = \tau(p\otimes q_1 s)m_2,$$

for all $p\in P$. Hence, we have $\overline{(q_1,m_1)} = \overline{(q_1 s,m_2)} = \overline{(q_2,m_2)}$.

2) If $q_1\otimes m_1 = q_2\otimes m_2$, By Proposition 1.4.10 of [3], we have that $(q_1,m_1) = (q_2,m_2)$ or for some positive integer $n>1$, there is a sequence

$$(q_1,m_1) = (y_1,x_1) \to (y_2,x_2) \to \cdots \to (y_n,x_n) = (q_2,m_2)$$

in which, for each i in $\{1,2,\cdots,n-1\}$, either $((y_i,x_i),(y_{i+1},x_{i+1})) \in \mathfrak{R}$ or $((y_{i+1},x_{i+1}),(y_i,x_i)) \in \mathfrak{R}$. By part 1), we can easily get that $\overline{(q_1,m_1)} = \overline{(q_2,m_2)}$. □

3.5. Definition. Let S and T be two semigroups. A Morita context (S,T,P,Q,τ,μ) is called unital, if P is a unital S-T-biact and Q is a unital T-S-biact.

3.6. Lemma. *Let (S, T, P, Q, τ, μ) be a unital Morita context and $M \in US$-FAct. Then we have a T-isomorphism $(Q \times M)/\rho_{(Q\times M)} \cong (Q \otimes M)/\zeta_{(Q\otimes M)}$.*

Proof Define a map $\varphi : (Q\times M)/\rho_{(Q\times M)} \to (Q\otimes M)/\zeta_{(Q\otimes M)}$ by $\varphi(\overline{(q,m)}) = (q\otimes m)\zeta$, where $(q \otimes m)\zeta$ represent the congruence class $(q \otimes m)\zeta_{(Q\otimes M)}$.

Suppose $\overline{(q_1, m_1)}, \overline{(q_2, m_2)} \in \widetilde{(Q, M)}$. If $\overline{(q_1, m_1)} = \overline{(q_2, m_2)}$, we have $\tau(p \otimes q_1)m_1 = \tau(p \otimes q_2)m_2$, for all $p \in P$. Then

$$\mu(y\otimes x)(q_1\otimes m_1) = \mu(y\otimes x)q_1\otimes m_1 = y\otimes\tau(x\otimes q_1)m_1 = y\otimes\tau(x\otimes q_2)m_2 = \mu(y\otimes x)(q_2\otimes m_2),$$

for all $y \in Q, x \in P$. This implies that $(q_1 \otimes m_1)\zeta = (q_2 \otimes m_2)\zeta$. Therefore, φ is well-defined. Obviously, φ is surjective.

If $(q_1 \otimes m_1)\zeta = (q_2 \otimes m_2)\zeta$, for all $x \in P, y \in Q$, we have

$$\mu(y \otimes x)(q_1 \otimes m_1) = \mu(y \otimes x)(q_2 \otimes m_2).$$

Since $y \otimes \tau(x \otimes q_1)m_1 = y\tau(x \otimes q_1) \otimes m_1 = \mu(y \otimes x)(q_1 \otimes m_1)$, we get

$$y \otimes \tau(x \otimes q_1)m_1 = y \otimes \tau(x \otimes q_2)m_2.$$

By Lemma 3.4, we have

$$\overline{(y, \tau(x \otimes q_1)m_1)} = \overline{(y, \tau(x \otimes q_2)m_2)}.$$

For all $p \in P$, we have

$$\tau(\tau(p\otimes y)x\otimes q_1)m_1 = \tau(p\otimes y)\tau(x\otimes q_1)m_1 = \tau(p\otimes y)\tau(x\otimes q_2)m_2 = \tau(\tau(p\otimes y)x\otimes q_2)m_2.$$

Since P is unitary and τ is surjective, we get

$$\{\tau(p \otimes y)x|\text{for all } p, x \in P, q \in Q\} = SP = P.$$

Then $\overline{(q_1, m_1)} = \overline{(q_2, m_2)}$. This proves that φ is injective.

For all $\overline{(q, m)} \in \widetilde{(Q, M)}$, $\mu(y \otimes x) \in T$, we have

$$\begin{aligned}
\varphi(\mu(y \otimes x)\overline{(q, m)}) &= \varphi(\overline{(y, \tau(x \otimes q)m)}) = (y \otimes \tau(x \otimes q)m)\zeta \\
&= (y\tau(x \otimes q) \otimes m)\zeta = (\mu(y \otimes x)q \otimes m)\zeta \\
&= \mu(y \otimes x)((q \otimes m)\zeta) = \mu(y \otimes x)\varphi(\overline{(q, m)}).
\end{aligned}$$

Hence, φ is a T-isomorphism. That is, $(Q \times M)/\rho_{(Q\times M)} \cong (Q \otimes M)/\zeta_{(Q\otimes M)}$ as left T-act. □

By Theorem 3.3 and Lemma 3.6, we have the following theorem which generalizes Theorem 2 in paper [2].

3.7. Theorem. *Let S and T be two semigroups. If (S, T, P, Q, τ, μ) be a unital Morita context with τ and μ are surjective, then we have the category equivalence $F : US\text{-}FAct \rightleftharpoons UT\text{-}FAct : G$, where $F = (Q \otimes -)/\zeta_{(Q\otimes -)}$ and $G = (P \otimes -)/\zeta_{(P\otimes -)}$.*

4. Characterization of Morita context

In this section, we give an equivalent condition of Morita context in semigroup settings. Also, we give a characterization of Morita context for factorisable semigroups. Similar to Theorem 1 in [6], we have the following.

4.1. Theorem. *Let P and Q be two sets. We have the following equivalent conditions.*

1) There exist two semigroups S and T such that (S, T, P, Q, τ, μ) is a Morita context.

2) There exist maps $\Gamma : P \times Q \times P \to P$ and $\Delta : Q \times P \times Q \to Q$ such that

I) $\Gamma(\Gamma((p_1, q_1, p_2)), q_2, p_3) = \Gamma((p_1, \Delta((q_1, p_2, q_2)), p_3)) = \Gamma(p_1, q_1, \Gamma((p_2, q_2, p_3)))$;

II) $\Delta(\Delta((q_1, p_1, q_2)), p_2, q_3) = \Delta(q_1, \Gamma((p_1, q_2, p_2)), q_3) = \Delta(q_1, p_1, \Delta((q_2, p_2, q_3)))$.

Proof 1) $\Rightarrow$ 2) : Suppose that (S,T,P,Q,τ,μ) is a Morita context. Define $\Gamma : P \times Q \times P \to P$ and $\Delta : Q \times P \times Q \to Q$ by putting $\Gamma((p_1,q_1,p_2)) = \tau(p_1 \otimes q_1) \cdot p_2$ and $\Delta((q_1,p_1,q_2)) = \mu(q_1 \otimes p_1) \cdot q_2$. We can easily check that Γ and Δ satisfy the conditions I) and II).

2) $\Rightarrow$ 1) : Define $H_a : P \to P$ by putting $H_a(p) = \Gamma((a,p))$ and define $K_b : Q \to Q$ by putting $K_b(q) = \Delta((b,q))$, where $a \in P \times Q$ and $b \in Q \times P$.

We write $\mathcal{X} = \{H_a | a \in P \times Q\}$ and $\mathcal{Y} = \{K_b | b \in Q \times P\}$. For all $H_{(p_1,q_1)}, H_{(p_2,q_2)} \in \mathcal{X}$, for all $p \in P$, we have

$$H_{(p_1,q_1)}H_{(p_2,q_2)}(p) = \Gamma((p_1,q_1,\Gamma(p_2,q_2,p))) = \Gamma((\Gamma(p_1,q_1,p_2),q_2,p)) = H_{(\Gamma(p_1,q_1,p_2),q_2)}(p).$$

That is, $H_{(p_1,q_1)}H_{(p_2,q_2)} = H_{(\Gamma((p_1,q_1,p_2)),q_2)} \in \mathcal{X}$. Then we easily get that $\mathcal{X}$ is a subsemigroup of $\mathrm{End}(P)$. Similarly, we have that $\mathcal{Y}$ is a subsemigroup of $\mathrm{End}(Q)$.

For all $p \in P$, $H_a \in \mathcal{X}$, $K_b \in \mathcal{Y}$, define $H_a \cdot p = \Gamma((a,p))$ and $p \cdot K_b = \Gamma((p,b))$. Then P is a $\mathcal{X}$-$\mathcal{Y}$-biact. Similarly, for all $q \in Q$, we can define $K_b \cdot q = \Delta((b,q))$ and $q \cdot H_a = \Delta((q,a))$. This makes Q to be a $\mathcal{Y}$-$\mathcal{X}$-biact.

Now, we define $\alpha : P \otimes_{\mathcal{Y}} Q \to \mathcal{X}$ and $\beta : Q \otimes_{\mathcal{X}} P \to \mathcal{Y}$ by putting $\alpha(p \otimes q) = H_{(p,q)}$ and $\beta(q \otimes p) = K_{(q,p)}$, where $p \in P$ and $q \in Q$. It is easy to check that α and β are both biact morphisms. Then

$$\alpha(p_1 \otimes q) \cdot p_2 = H_{(p_1,q)} \cdot p_2 = \Gamma((p_1,q,p_2)) = p_1 \cdot K_{(q,p_2)} = p_1 \cdot \beta(q \otimes p_2).$$

Similarly, we have

$$\beta(q_1 \otimes p)q_2 = q_1\alpha(p \otimes q_2).$$

Then $(\mathcal{X},\mathcal{Y},P,Q,\alpha,\beta)$ is a Morita context. $\Box$

4.2. Definition. [7] A semigroup S is called factorisable if $S = S^2$.

4.3. Theorem. *Let P and Q be two sets. We have the following equivalent conditions.*

1) There exist two factorisable semigroups S and T such that (S,T,P,Q,τ,μ) is a unital Morita context and τ and μ are surjective.

In this case, $(Q \otimes -)/\zeta_{(Q\otimes -)} : US\text{-}Act \rightleftharpoons UT\text{-}Act : (P \otimes -)/\zeta_{(P\otimes -)}$ are equivalent functors.

2) There exist surjective maps $\Gamma : P \times Q \times P \to P$ and $\Delta : Q \times P \times Q \to Q$ satisfy the two conditions in part 2) of Theorem 4.1 and

III) For all $p,p^{'} \in P$, $q \in Q$, there exist $p_1,p_2 \in P, q_1,q_2 \in Q$ such that

$$\Gamma(((p,q),p^{'})) = \Gamma((\Gamma(p_1,q_1,p_2),q_2,p^{'})).$$

IV) For all $p \in P$, $q,q^{'} \in Q$, there exist $p_1,p_2 \in P, q_1,q_2 \in Q$ such that

$$\Delta(((q,p),q^{'})) = \Delta((\Delta(q_1,p_1,q_2),p_2,q^{'})).$$

Proof 1) $\Rightarrow$ 2) : Since S is factorisable and τ is surjective, for all $p \in P, q \in Q$, there exist $p_1,p_2 \in P, q_1,q_2 \in Q$ such that $\tau(p \otimes q) = \tau(p_1 \otimes q_1)\tau(p_2 \otimes q_2)$. Hence,

$$\begin{aligned}\Gamma(((p,q),p^{'})) &= \tau(p \otimes q)p^{'} = \tau(p_1 \otimes q_1)\tau(p_2 \otimes q_2)p^{'} = \tau(p_1 \otimes q_1)\Gamma((p_2,q_2,p^{'})) \\ &= \Gamma((p_1,q_1,\Gamma(p_2,q_2,p^{'}))) = \Gamma((\Gamma((p_1,q_1,p_2)),q_2,p^{'})).\end{aligned}$$

Therefore, the condition III) holds. Similarly, we can get IV).

By Theorem 3.7 or Theorem 2 in [2], we have the category equivalence $(Q\otimes -)/\zeta_{(Q\otimes -)} : US\text{-Act} \rightleftharpoons UT\text{-Act} : (P \otimes -)/\zeta_{(P\otimes -)}$.

2) $\Rightarrow$ 1) : For all $H_{(p,q)} \in \mathcal{X}, p^{'} \in P$, by the condition III), we have

$$\Gamma(((p,q),p^{'})) = \Gamma((\Gamma(p_1,q_1,p_2),q_2,p^{'})).$$

This implies that

$$\begin{aligned} H_{(p,q)}(p^{'}) &= \Gamma(((p,q),p^{'})) = \Gamma((\Gamma(p_1,q_1,p_2),q_2,p^{'})) \\ &= \Gamma((p_1,q_1,\Gamma((p_2,q_2,p^{'})))) = H_{(p_1,q_1)}H_{(p_2,q_2)}(p^{'}). \end{aligned}$$

That is, $H_{(p,q)} = H_{(p_1,q_1)}H_{(p_2,q_2)}$. This proves that $\mathcal{X}$ is factorisable.

Similarly, we have that $\mathcal{Y}$ is a factorisable semigroup.

Since Γ and Δ are surjective, we obviously have that P and Q are unital as biacts and α and β are surjective. Hence, $(\mathcal{X},\mathcal{Y},P,Q,\alpha,\beta)$ is a unital Morita context. $\Box$

Acknowledgements The author is grateful to the referees for their valuable comments which improve the paper highly.

References

[1] Banaschewski, B. *Functors into the categories of M-sets*, Abh. Math. Sem. Univ. Hamburg **8**, 49-64, 1972.

[2] Chen, Y. Q. and Shum, K. P. *Morita equivalence for factorisable semigroups*, Acta Mathematica Sinica (English Series), **17**(3), 437-454, 2001.

[3] Howie, J.M. *Fundamentals of semigroup theory*, Clarendon press, Oxford, 1995.

[4] Knauer, U. *Projectivity of acts and Morita equivalence of semigroups*, Semigroup Forum **3**, 359-370, 1972.

[5] Kyuno, S. *Equivalence of module categories*, Math. J. Okayama Univ, **28**, 147-150, 1986.

[6] Paseka, J. *Characterization of Morita equivalence pairs of quantales*, Int. J. Theor. Phys., **44** (7),875-883, 2005.

[7] Talwar, S. *Strong Morita equivalence and a generalisation of the Rees theorem*, J. Algebra, **181**, 371-394, 1996.

[8] Talwar, S. *Morita equivalence for semigroups*, J. Austral. Math. Soc. (Series A), **59**, 81-111, 1995.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1091 – 1100

Soft topological space and topology on the Cartesian product

Matejdes Milan*

Abstract

The paper deals with a soft topological space which is defined over an initial universe set U with a fixed set of parameters E. The main goal is to point out that any soft topological space is homeomorphic to a topological space $(E \times U, \tau)$ where τ is an arbitrary topology on the product $E \times U$, consequently many soft topological notions and results can be derived from general topology. Furthermore, in many papers some notions are introduced by different ways and it would be good to give a unified approach for a transfer of topological notions to a soft set theory and to create a bridge between general topology and soft set theory.

Keywords: Soft set, Soft open (closed) set, Soft interior (closure) of soft set, Soft topological space, Separation axioms, Soft continuity, Soft e-continuity.

2000 AMS Classification: 54A05, 54B10, 54C08, 54C60.

Received : 10.03.2015 *Accepted* : 22.07.2015 *Doi* : 10.15672/HJMS.20164513117

1. Introduction

In 1999, Molodtsov [18], [19], [20] introduced a soft set theory as a new tool for investigation of uncertainties where we can find a large range of applications of soft sets in many different fields. There has been a rapid growth of interest in soft set theory, its applications and its connection with another mathematical branches [1], [2], [4], [5], [7], [8], [12], [13], [14], [15], [16], [23]. Moreover, there are many papers devoted to soft topological spaces [3], [6], [9] ,[10], [11], [17], [21], [22]. The basic topological notions such as the soft open and soft closed sets, soft subspace, soft closure and soft interior, soft boundary, soft separation axioms, soft continuity have been introduced and the investigation of their basic properties has been established.

We continue investigating the soft topological theory based on a corresponce between set valued mappings and binary relations. Their close connection shows that both definitions of a soft set by a set valued mapping and by a relation are equivalent and there is

*Department of Mathematics and Computer Science, Faculty of Education, Trnava University in Trnava, Priemyselná 4, 918 43 Trnava, Slovakia. Email: milan.matejdes@truni.sk

only a formal difference between them. Furthemore, the binary relation view is very comfortable and many results concerning the properties of the operations on soft sets follow from the set theory. On the other hand, the set valued mapping view gives possibilities for a further investigation of the soft set theory in many directions, since the theory of set valued mappings is strong and has many applications in mathematics (general topology, generalized continuities, linear and dynamic programming, differential inclusions, fixe point theory, statistics, economics and so on).

This paper shows that many results concerning soft topological spaces follow from general topology. In particular, some notions introduced in soft topological spaces and their consequences (the properties of soft open (closed) sets, interior and closure of soft set, soft cluster points) are identical with corresponding notions known from general topology. Some of them are different (soft separation axioms, soft continuity) and they are introduced by different ways. The main goal of this paper is to give a unified view for a further development of soft topological spaces based on the results of general topology.

2. Relations, set valued mappings and their correspondence

Any subset S of a Cartesian product $A \times U$ is called a binary relation from a set A to a set U. By $\mathbf{R}(A, U)$, we denote a set of all binary relations from A to U and $S[a] := \{u \in U : [a, u] \in S\}$. The operations of sum $S \cup T$, $\cup_{t \in T} S_t$, intersection $S \cap T$, $\cap_{t \in T} S_t$, complement S^c and difference $S \setminus T$ of relations are defined in the obvious way as in the set theory.

By $F : A \to 2^U$ we denote a set valued mapping from A to power set 2^U of U. The set of all set valued mappings from A to 2^U is denoted by $\mathbf{F}(A, U)$. If F, G are two set valued mappings, then $F \subset G$ $(F = G)$ means $F(a) \subset G(a)$ $(F(a) = G(a))$ for any $a \in A$.

A graph of F is a set $Gr(F) := \{[a, u] \in A \times U : u \in F(a)\}$ and it is a subset of $A \times U$, hence $Gr(F) \in \mathbf{R}(A, U)$. So, any set valued mapping F determines a relation from $\mathbf{R}(A, U)$ denoted by $R_F := \{[a, u] \in A \times U : u \in F(a)\} = Gr(F)$.

On the other hand, any relation $S \in \mathbf{R}(A, U)$ determines a set valued mapping F_S from A to 2^U where $F_S(a) = S[a]$. So, there is one-to-one correspondence between a relation S from $\mathbf{R}(A, U)$ and a set valued mapping G from $\mathbf{F}(A, U)$, i.e.,

$$S \mapsto F_S,\ F_S(a) = S[a], \quad G \mapsto R_G,\ R_G[a] = G(a),$$

$$F_{R_G} = G,\ R_{F_S} = S.$$

2.1. Remark. For $H, G, F_t \in \mathbf{F}(A, U)$, $t \in T$, we define the following obvious set valued mapping operations and we give also their binary relation equivalents.

(1) Sum: $\cup_{t \in T} F_t : A \to 2^U$
$(\cup_{t \in T} F_t)(a) = \cup_{t \in T} F_t(a) = \cup_{t \in T} R_{F_t}[a] = (\cup_{t \in T} R_{F_t})[a]$, $a \in A$,

(2) Intersection: $\cap_{t \in T} F_t : A \to 2^U$
$(\cap_{t \in T} F_t)(a) = \cap_{t \in T} F_t(a) = \cap_{t \in T} R_{F_t}[a] = (\cap_{t \in T} R_{F_t})[a]$, $a \in A$,

(3) Complement: $H^c : A \to 2^U$
$(H^c)(a) = U \setminus H(a) = U \setminus R_H[a] = R_H^c[a]$, $a \in A$,

(4) Difference: $H \setminus G : A \to 2^U$
$(H \setminus G)(a) = H(a) \setminus G(a) = R_H[a] \setminus R_G[a] = (R_H \setminus R_G)[a]$, $a \in A$.

The next lemma is a consequence of Remark 2.1.

2.2. Lemma. *For $H, G, F_t \in \mathbf{F}(A, U)$ and $S, P, R_t \in \mathbf{R}(A, U)$, $t \in T$, the following equations hold.*

(1) $R_{\cup_{t \in T} F_t} = \cup_{t \in T} R_{F_t}$, $F_{\cup_{t \in t} R_t} = \cup_{t \in T} F_{R_t}$,

(2) $R_{\cap_{t\in T}F_t} = \cap_{t\in T} R_{F_t}$, $F_{\cap_{t\in t}R_t} = \cap_{t\in T} F_{R_t}$,
(3) $R_{H^c} = {R_H}^c$, $F_{S^c} = {F_S}^c$,
(4) $R_{H\setminus G} = R_H \setminus R_G$, $F_{S\setminus P} = F_S \setminus F_P$.

3. Set valued mapping and binary relation representation of soft set

In this section we will consider soft sets over a common initial universe set U and a fixed set of parameters E and a definition of a soft set is introduced by a set valued mapping (see references).

3.1. Definition. If $F : E \to 2^U$ is a set valued mapping, then a pair (F, E) is called a soft set over U with respect to a set of parameters E. The family of all soft sets over U with respect to a set of parameters E is denoted by $SS(E, U)$.

As we said above there is no difference between the graph of a set valued mapping F and a relation $Gr(F) \subset E \times U$, which is a member of $\mathbf{R}(E, U)$. So, a soft set can be defined equivalently as follows.

3.2. Definition. A soft set over U with respect to a set E is any subset of $E \times U$. So, in this case a soft set is a member of $\mathbf{R}(E, U)$.

From Definition 3.2 we can see a benefit of both equivalent interpretations of a soft set. Any operation known from a set theory setting can be used for a soft set (soft sets) from $\mathbf{R}(E, U)$. In this case we deal with the soft sets as subsets of $E \times U$ and it is not necessary to use the different notations (symbols) for operations and many proofs can be omitted. For example, the next operations on the soft sets form $\mathbf{R}(E, U)$, $R \subset S$, $R = S$, $R \setminus S$, $R \cap S$, $R \cup S$, $\cup_{t\in T} R_t$, $\cap_{t\in T} R_t$, R^c are the set theory ones and all known properties from set theory hold in the soft set setting (for example associativity, commutativity, distributivity, de Morgan laws and so on).

Equivalently, if a soft set is understood as a pair (F, U) $(F \in \mathbf{F}(E, U))$, we can define standard operations on the set valued mappings (sum, intersection, complement, difference) which have equivalent binary relation forms, as we see from Remark 2.1 and Lemma 2.2.

3.3. Lemma. *Let $S \in \mathbf{R}(E, U)$, $G \in \mathbf{F}(E, U)$ and $(H_1, E), (H_2, E) \in SS(E, U)$. Then*
(1) *$G(a) \subset S[a]$ for all $a \in E$ iff $G \subset F_S$ iff $R_G \subset S$ iff a soft set (G, E) is a soft subset of (F_S, E),*
(2) *$S[a] \subset G(a)$ for all $a \in E$ iff $S \subset R_G$ iff $F_S \subset G$ iff a soft set (F_S, E) is a soft subset of (G, E),*
(3) *$G(a) = S[a]$ for all $a \in E$ iff $G = F_S$ iff $R_G = S$ iff a soft set (F_S, E) is equal to a soft set (G, E),*
(4) *a soft set (H_1, E) is a soft subset of (H_2, E) iff $R_{H_1} \subset R_{H_1}$ iff $H_1 \subset H_2$,*
(5) *a soft set (H_1, E) is equel to a soft set (H_2, E) iff $R_{H_1} = R_{H_1}$ iff $H_1 = H_2$.*

4. Special soft sets, their notation and terminology

Let $A \subset E$, $X \subset U$. Then $A \times X$ is called a rectangle soft set. It represents a constant soft set (a constant set valued mapping F with values $F(a) = X$ if $a \in A$ and $F(a) = \emptyset$ if $a \notin A$) denoted also $c(A, X)$. Maximal (minimal) rectangle soft set with respect to the set inclusion is $E \times U$ $(\emptyset \times \emptyset)$ called a full soft set (a null soft set). For the special cases of a constant soft set $c(A, X) \in \mathbf{R}(A, U)$ we introduce the next notation and terminology. Let $A \subset U$, $X \subset U$, $e \in E$, $x \in U$.

(1) $c(A,x) := A \times \{x\}$ - a horizontal x-line on A,
(2) $c(E,x) := E \times \{x\}$ - a full horizontal x-line,
(3) $c(e,X) := \{e\} \times X$ - a vertical e-line on X,
(4) $c(e,U) := \{e\} \times U$ - a full vertical e-line,
(5) $c(e,x) := [e,x]$ - a point, denoted also $P[e,x]$ or briefly P.

4.1. Lemma. *Let $S \in \mathbf{R}(E,U)$, $G \in \mathbf{F}(E,U)$. Then*

(1) $S = \cup_{e\in E}\big(\{e\} \times S[e]\big) = \cup_{e\in E}[S \cap c(e,U)] = \cup_{x\in U}[S \cap c(E,x)]$,
(2) $R_G = \cup_{e\in E}\big(\{e\} \times G(e)\big) = \cup_{e\in E}[R_G \cap c(e,U)] = \cup_{x\in U}[R_G \cap c(E,x)]$.

5. Soft topological space

By [9],[10],[21] a soft topological space is a triplet (E,U,τ), where $\tau \subset SS(E,U)$ is a topology. So, τ is represented by a family of set valued mappings F from $\mathbf{F}(E,U)$ each of them has a binary relation representation R_F from $\mathbf{R}(E,U)$. Put $\tau_{E\times U} := \{R \in \mathbf{R}(E,U) : (F_R,E) \in \tau\}$.

On the other hand, if $(E \times U, \tau_{E\times U})$ is a topological space, then $(E,U,\tau_{E,U})$ is a soft topological space, where $\tau_{E,U} = \{(G,E) \in SS(E,U) : R_G \in \tau_{E\times U}\}$. Then a soft topological space can be characterized (defined) as follows:

5.1. Proposition. *A triplet $(E,U,\tau_{E,U})$ is a soft topological space if and only if $(E \times U, \tau_{E\times U})$ is a topological space. For brevity we will denote $\tau_{E\times U}$ as well as $\tau_{E,U}$ by τ and the difference between both topologies is clear from notation (E,U,τ) (a soft topological space, $\tau \subset SS(E,U)$) and $(E \times U, \tau)$ (a topological space, $\tau \subset \mathbf{R}(E,U)$).*

Again, there is a one-to-one correspondence between the soft topological spaces and the topological spaces and any soft topological space can be considered as an arbitrary topological space on the product of two sets. The members from τ are called open sets in a topological space $(E \times U, \tau)$ or soft open sets in a soft topological space (E,U,τ). A complement of an open set (a soft open set) is called a closed set (a soft closed set). It can be formulated by the following lemma.

5.2. Lemma. *A soft set (G,E) is soft open (soft closed) in a soft topological space (E,U,τ) iff R_G is open (closed) in a topological space $(E \times U, \tau)$ and S is open (closed) in a topological space $(E \times U, \tau)$ iff (F_S,E) is soft open (soft closed) in a soft topological space (E,U,τ).*

Any topological notion known from general topology on the product of two sets can be introduced (formulated) also for a soft topological space by direct reformulation. From Proposition 5.1 we can see that many results in a soft topological space follow from general topology, provided they are directly reformulated. It is not necessary to prove the properties of the soft open and soft closed sets, the properties of the soft interior and soft closure operators, soft cluster points, soft interior points, soft topological subspaces, soft boundary sets, the soft separation axioms and so on, provided they are defined by the same way as in topological spaces. But some notions are defined by a different way and we will discuss them below.

On the other hand, many interesting and important results follow from general topology, provided τ is the Tychonoff (product) topology on $E \times U$. But if a topology τ on $E \times U$ is not Tychonoff, many results do not hold. So, there are many open problems in the soft topology setting and before researchers is a huge challenge.

6. Comparison of some topological notions and soft topological ones

In this section we show a few correspondences between some topological notions and soft topological ones. For example, our expectation is that the soft closure of a soft set (G, E) in a soft topological space (E, U, τ) agrees with the closure of R_G in a topological space $(E \times U, \tau)$.

Recall a few basic topological notions. Let $(E \times U, \tau)$ be a topological space. A set $H \in \tau$ is an open neighborhood of a point $P[a, x] \in E \times U$, if $P[a, x] \in H$ and $P[a, x]$ is a cluster point of $S \subset E \times U$, if any open neighborhood of $P[a, x]$ intersects S. The set of all cluster points of S in $(A \times U, \tau)$ is equal to the closure of S denoted by $cl(S)$, which is the smallest closed set containing S or it is the intersection of all closed sets containing S. A point $P[a, x] \in E \times U$ is an interior point of S, if there is an open set $H \in \tau$ such that $P[a, x] \in H \subset S$ and S is open, if any its point is an interior point of S. The sum of all open subsets of S is called the interior of S denoted by $int(S)$.

1. Open (closed) sets and interior (cluster) points of a set versus soft open (closed) sets and a-interior (a-cluster) points of a soft set

By [9], an a-soft open neighborhood of x is any open soft set (G, E) such that $x \in G(a)$, equivalently R_G is open and $x \in R_G[a]$ or $P[a, x] \in R_G$. A point $x \in U$ is said to be an a-cluster point of $(H, E) \in SS(E, U)$ if for every a-soft open neighborhood (G, E) of x, (H, E) and (G, E) are not soft disjoint (there is $e \in E$ such that $H(e) \cap G(e) \neq \emptyset$). The set of all a-cluster points of (H, A) is denoted by $cl(H, a)$, see [9]. Similarly, $int(H, a)$ is a set of all a-interior points of (H, E), see [9].

6.1. Lemma.
(1) *$x \in cl(H, a)$ if and only if $P[a, x] \in cl(R_H)$, so $cl(H, a) = cl(R_H)[a]$,*
(2) *$x \in int(H, a)$ if and only if $P[a, x] \in int(R_H)$, so $int(H, a) = int(R_H)[a]$.*
Consequently, $cl(H, a)$ ($int(H, a)$) is a set of all cluster (interior) points of R_H in $(E \times U, \tau)$ from the full vertical a-line ($cl(H, a) = cl(R_H) \cap c(a, U)$ ($int(H, a) = int(R_H) \cap c(a, U)$)).

Proof. (1)
"$\Rightarrow$" Let $x \in cl(H, a)$ and $P[a, x] \in S \in \tau$. Then $x \in F_S(a)$. That means (F_S, E) is a-open neighborhood of x, so (H, E) and (F_S, E) are not soft disjoint. Hence, there are $e \in E, y \in U$ such that $y \in H(e) \cap F_S(e)$ or $P[e, y] \in R_H \cap S$. That means $R_H \cap S \neq \emptyset$ or $P[a, x] \in cl(R_H)$.

"$\Leftarrow$" Let $P[a, x] \in cl(R_H)$ and (G, E) be a-soft open neighborhood of x. Then $P[a, x] \in R_G \in \tau$ and $R_H \cap R_G \neq \emptyset$. So, there are $e \in E$ and $y \in U$ such that $y \in R_H[e] \cap R_G[e]$, hence (H, E) and (G, E) are not soft disjoint, so $x \in cl(H, a)$.
Item (2) is similar.

□

6.2. Lemma. *Let $cl(G, E)$, $int(G, E)$ be a soft closure, a soft interior of a soft set (G, E), respectively (see [9]). Then*
(1) *$cl(G, E) = (F_{cl(R_G)}, E)$,*
(2) *$int(G, E) = (F_{int(R_G)}, E)$.*

Proof. (1) By the definition of the soft closure ([9]) and by Lemma 5.2, $cl(G, E)$ is the intersection of all soft closed supersets $(G_t, E), t \in T$ of (G, E) if and only if $\cap_{t \in T} R_{G_t}$ is the intersection of all closed (in $(E \times U, \tau)$) supersets R_{G_t} of R_G. That means $\cap_{t \in T} R_{G_t} =$

$cl(R_G)$ is the graph of a multivalued mapping H for which $cl(G, E) = (H, E)$, so $F_{cl(R_G)} = H$.

(2) is similar.

□

Since $(F_{cl(R_G)}, E) = cl(G, E)$, $cl(G, E)$ is a soft set given by a set valued mapping with the values $cl(R_G)[a]$, $a \in E$ which is equal to a set valued mapping $R_{G,E}$ defined in [9] as $R_{G,E}(a) = G(a) \cup cl(G, a) = cl(G, a) = cl(R_G)[a]$ (see Lemma 6.1). So, $cl(G, E) = (R_{G,E}, E)$. Similarly, $R_{G,E}(a) = G(a) \cap int(G, a) = int(R_G)[a]$ for $int(G, E) = (R_{G,E}, E)$ see [9]. So Proposition 3.9 and 3.12 of [9] are clear.

2. Separation axioms

In the literature ([11], [17], [21]) we can see notation $x \in (F, E)$, where F is a set valued mapping from E to 2^U and $x \in U$. It means $x \in F(e)$ for any $e \in E$. So, the notation $x \in (F, E)$ is in fact the inclusion $c(E, x) \subset R_F$. It was used in the definitions of soft separation axioms in a soft topological space. In general topology, the separation axioms separate two different points or a closed set and a point or two disjoint closed sets. For example, by [21], (E, U, τ) is called soft T_2, if for every distinct points x, y of U there are two soft open sets (F, E) and (G, E) such that $x \in (F, E)$, $y \in (G, E)$ and (F, E) and (G, E) are soft disjoint. That means it separates two full horizontal lines $c(E, x)$ and $c(E, y)$. Further, if (F, E) is a soft closed set and $x \notin (F, E)$, a soft regularity introduced in [21] separates two sets, namely $c(E, x)$ and R_F which need not be disjoint. It is a very strict definition as we see from the next lemma.

6.3. Lemma. *Let (E, U, τ) be a soft topological space and (F, E) be a soft closed set. If there are $e_1, e_2 \in E$ and a point $y \in U$ such that $y \in F(e_1)$ and $y \notin F(e_2)$ (it is sufficient (E, U, τ) is not indiscrete), then (E, U, τ) is not soft regular (in the sense of [21]). Consequently, if some soft topological space over U is soft regular, then any soft closed set (F, E) is constant, i.e., there is a set $X \subset U$ such that $F(e) = X$ for any $e \in E$.*

Proof. Suppose (E, U, τ) is soft regular. It is clear $y \notin (F, E)$. Then there are two soft open and soft disjoint sets (G, E) containing y and (H, E) containing (F, E), but $y \in G(e_1) \cap H(e_1)$, a contradiction.

□

The next theorem shows that soft regularity in the sense of [21] seems to be a rather strong definition.

6.4. Theorem. *If a non indiscrete soft topological space is soft regular, then any soft closed set is a constant soft set (it is of the form $c(E, X)$).*

In [10], the authors introduced other soft separation notions, namely T_0, T_1, T_2, T_3. We recall only two of them (for further see [10]). A soft topological space (E, U, τ) is called soft T_2, if for any distinct points x and y of U and for every $a \in E$ there exist two soft open sets (G, E) and (H, E) such that $x \in_a (G, E)$, $y \in_a (H, E)$ and $G(a) \cap H(a) = \emptyset$ ($z \in_a (G, E)$ means that $z \in G(a)$ and $z \notin_a (G, E)$ means that $z \notin G(a)$). In this case we separate a couple of the points $P[a, x]$ and $P[a, y]$ from full vertical a-line $c(a, U)$ by two soft open sets "disjoint at a". This is a different definition of a soft T_2-space introduced in [21]. Further by [10], a soft topological space (E, U, τ) is called a soft T_3-space if for every point $x \in U$, for every $a \in E$ and for every soft closed set (Q, E) such that $x \notin_a (Q, E)$ there exist two soft open sets (G, E) and (H, E) such that $x \in_a (G, E)$, $Q(a) \subset H(a)$, and $G(a) \cap H(a) = \emptyset$. In this case the sets $G(a) \cap c(a, U)$ and $H(a) \cap c(a, U)$ are two subsets of $E \times U$ which are open in a subspace $(c(a, U), \tau_a)$ of $(A \times U, \tau)$.

Generally, for any full vertical line $c(e,U)$, topology τ induces a relative topology τ_e on $c(e,U)$, so also on U (for different e_1, e_2 the induced topological spaces (U, τ_{e_1}) and (U, τ_{e_2}) can be different). So we have the next theorem (which does not hold generally. " $\Rightarrow$ " follows from hereditary of $T_i, i = 0,1,2,3$ and " $\Leftarrow$ " follows from a character of the definitions of the soft separation axioms in [10]).

6.5. Theorem. *A soft topological space (E, A, τ) is soft T_i (in the sense of [10]) if and only if the topological space (U, τ_e) is T_i $(i = 0,1,2,3)$ for any $e \in E$.*

Any subset S of $(E \times U, \tau)$ induces relative topology τ_S on S. Since the properties T_i $(i = 0,1,2,3)$ are hereditary, Proposition 3.13 of [10] holds for any $S \subset E \times U$ not only for $E \times Y$ (see Definition 3.12 in [10] or [21]).

3. Soft e-continuity of f and continuity of $e \times f$, soft continuity of Φ_{ef}

In [3], [9], [10], [13], [22] for two functions $e : E_1 \to E_2$, $f : U_1 \to U_2$, a function Φ_{ef} from $SS(E_1, U_1)$ to $S(E_2, U_2)$ was defined (denoted also f_{pu} in [22], φ_ϕ in [3]). We will show a connection between Φ_{ef} and the product $e \times f : E_1 \times U_1 \to E_2 \times U_2$, defined as $(e \times f)([e_1, x_1]) = [e(e_1), f(x_1)]$. Define two soft mappings:
$\mathbf{S}_{e\times f} : SS(E_1, U_1) \to SS(E_2, U_2)$ as $\mathbf{S}_{e\times f}((H, E_1)) = (F_{(e\times f)(R_H)}, E_2)$,
$\mathbf{S}^{-1}_{e\times f} : SS(E_2, U_2) \to SS(E_1, U_1)$ as $\mathbf{S}^{-1}_{e\times f}((G, E_2)) = (F_{(e\times f)^{-1}(R_G)}, E_1)$.

6.6. Theorem. *Let $(H, E_1) \in SS(E_1, U_1)$ and $(G, E_2) \in SS(E_2, U_2)$. Then*
(1) $\mathbf{S}_{e\times f} = \Phi_{ef}$, *i.e.,* $\Phi_{ef}((H, E_1)) = (F_{(e\times f)(R_H)}, E_2)$,
(2) $\mathbf{S}^{-1}_{e\times f} = \Phi^{-1}_{ef}$, *i.e.,* $\Phi^{-1}_{ef}((G, E_2)) = (F_{(e\times f)^{-1}(R_G)}, E_2)$.

Proof. Let $H \in \mathbf{F}(E_1, U_1)$ and R_H be a corresponding relation, so $R_H[a] = H(a)$. Then, by Lemma 4.1 (1),

$$(e \times f)(R_H) = (e \times f)\big(\cup_{a\in E_1} (\{a\} \times R_H[a])\big) =$$

$$= \cup_{a\in E_1} (e \times f)\big(\{a\} \times R_H[a]\big) = \cup_{a\in E_1}\big(e(a) \times f(R_H[a])\big).$$

That means $(e \times f)(R_H)$ is a subset of $E_2 \times U_2$ and corresponding set valued mapping denoted by $G : E_2 \to U_2$ has its values given by $\big[(e \times f)(R_H)\big][p_2]$ for any $p_2 \in E_2$.

$$G(p_2) = \big[(e \times f)(R_H)\big][p_2] = \big[\cup_{a\in E_1} \big(e(a) \times f(R_H[a])\big)\big][p_2] =$$

$$= \cup_{a\in E_1} \big[e(a) \times f(R_H[a])\big][p_2] = \cup\{f(R_H[a]) : e(a) = p_2\} =$$

$$= \cup\{f(H(a)) : a \in e^{-1}(p_2)\}.$$

So, (G, E_2) is the image of (H, E_1) under Φ_{ef} as it was defined in [10]. So, $\Phi_{ef}((H, E_1)) = (G, E_2) = (F_{(e\times f)(R_H)}, E_2)$ (see Lemma 3.3 (3)) or $\Phi_{ef} = \mathbf{S}_{e\times f}$.
Similarly we can show (see [10]) that $\Phi^{-1}_{ef}((G, E_2)) = (D, E_1)$,
where $D(p_1) = f^{-1}(G(e(p_1))) = (e\times f)^{-1}(R_G)[p_1]$. So, $\Phi^{-1}_{ef}((G, E_2)) = (F_{(e\times f)^{-1}(R_G)}, E_1)$ or $\Phi^{-1}_{ef} = \mathbf{S}^{-1}_{e\times f}$.

□

Proposition 2.18 and 2.19 of [9] (Proposition 2.8 of [10]) follow from the properties of the image and the inverse image, which hold generally for any function.

In [10], for two soft topological spaces (E_1, U_1, τ) and (E_2, U_2, σ) a definition of a soft e-continuity of f was introduced by the following way.

6.7. Definition. Let (E_1, U_1, τ) and (E_2, U_2, σ) be two soft topological spaces and $x \in U_1$, $e : E_1 \to E_2$. A map $f : U_1 \to U_2$ is called soft e-continuous at the point x if for every $a \in E_1$ and every $e(a)$-soft open neighborhood (G, E_2) of $f(x)$ in (E_2, U_2, σ) there exists an a-soft open neighborhood (H, E_1) of x in (E_1, U_1, τ) such that $\Phi_{ef}((H, E_1))$ is a soft subset of (G, E_2). If the map f is soft e-continuous at any point $x \in E_1$, then we say that the map f is soft e-continuous.

Now we reformulate the definition above in the corresponding topological spaces $(E_1 \times U_1, \tau)$ and $(E_1 \times U_2, \sigma)$.

6.8. Definition. Let $(E_1 \times U_1, \tau)$ and $(E_1 \times U_2, \sigma)$ be two topological spaces, $x \in U_1$, $e : E_1 \to E_2$. A map $f : U_1 \to U_2$ is called soft e-continuous at the point x if for every $a \in E_1$ (i.e., for every $[a, x] \in c(E_1, x)$) and every open set $G \in \sigma$ containing $[e(a), f(x)]$ there exists an open set $H \in \tau$ containing $[a, x]$ such that $(e \times f)(H) \subset G$. If the map f is soft e-continuous at any point $x \in U_1$, then we say that the map f is soft e-continuous.

Since $\Phi_{ef}((H, E_1)) = (F_{(e\times f)(R_H)}, E_2)$ (see Theorem above), $\Phi_{ef}((H, E_1))$ is a soft subset of (G, E_2) iff $(e \times f)(R_H) \subset R_G$. That means the soft e-continuity of f at x means that the set of all continuity points (in the general topology sense) of $e \times f : (E_1 \times U_1, \tau) \to (E_2 \times U_2, \sigma)$ contains a full horizontal x-line $c(E_1, x)$. Since $\Phi_{ef} = \mathbf{S}_{e\times f}$ and $\Phi_{ef}^{-1} = \mathbf{S}_{e\times f}^{-1}$, the next theorem is clear and Propositions 2.18 and 2.19 of [10] follows from standard equivalent conditions of continuity.

6.9. Theorem. *The next conditions are equivalent*

(1) *A function f is soft e-continuous (in the sense of [10]),*
(2) *$e \times f : (E_1 \times U_1, \tau) \to (E_2 \times U_2, \sigma)$ is continuous (in the topological sense),*
(3) *$\Phi_{ef}^{-1}((G, E_2)) \in \tau$ for any $(G, E_2) \in \sigma$.*

Finally, we recall a notion of a soft set point mentioned in [22]. A soft point, denoted by e_F is a soft set for which $F(e) \neq \emptyset$ and $F(a) = \emptyset$ for all $a \in E\setminus\{e\}$ and $e_F \in (G, E)$ means $F(a) \subset G(a)$ for all $a \in E$. So, a soft point is in fact any vertical e-line $c(e, X) = \{e\} \times X$ on X, for $X \neq \emptyset$. By [22], Φ_{ef} $(= f_{pu})$ is soft continuous (soft pu-continuous see [22]) at a soft point e_F if for any soft open set (G, E_2) containing $\Phi_{ef}(e_F)$ there is a soft open set (H, E_1) containing e_F such that $\Phi_{ef}((H, E_1))$ is a soft subset of (G, E_2) and Φ_{ef} is soft continuous if it is so at any soft point. Since a point $P[e, x]$ is also a soft point (namely e_F where $F(a) = \emptyset$ for $a \neq e$ and $F(e) = \{x\}$), soft continuity of Φ_{ef} in the sense of [22] implies a topological continuity of $e \times f$ at any point $P[e, x] \in E \times U$. The opposite implication also holds, as we prove in the next theorem.

6.10. Theorem. *A function Φ_{ef} $(= f_{pu})$ is soft continuous (in the sense of [22]) if and only if $e \times f$ is continuous (in the topological sense). Consequently, the soft continuity is equivalent to the soft e-continuity.*

Proof. It is sufficient to prove "$\Leftarrow$". Let g_K (i.e., $K(g) \neq \emptyset$ and $K(e) = \emptyset$ for $e \neq g$) be a soft point and (G, E_2) be a soft open superset of $\Phi_{ef}(g_K)$. Since R_G is an open set in $E_2 \times U_2$, R_G is open neighborhood of a point $[e(g), f(x)]$ for any $x \in K(g)$. Since $e \times f$ is continuous at $[g, x]$, for any $x \in K(g)$ there is an open subset H_x of $E_1 \times U_1$ containing $[g, x]$ such that $(e \times f)(H_x) \subset R_G$. Put $H := \cup_{x\in K(g)} H_x \supset R_{g_K}$. Then H is open in $E_1 \times U_1$ and $(e \times f)(H) \subset R_G$, so $(F_{(e\times f)(H)}, E_2)$ is a soft subset of (G, E_2) (see Lemma 3.3 (2)). Since (F_H, E_1) is a soft open set containing g_K (see Lemma 3.3 (1)), $\Phi_{ef}(F_H, E_1) = (F_{(e\times f)(H)}, E_2)$ (see Theorem 6.6 (1)) is a soft subset of (G, E_2).

□

7. Conclusion

This paper deals with the study of the theory of soft topological spaces and the main result is to show a deep connection between a soft topology τ on $SS(E,U)$ and a topology $\tau_{E\times U} = \{R \subset E \times U : (F_R, E) \in \tau\}$ on the product $E \times U$. Any soft topological space (E, U, τ) can be considered as a topological space on $E \times U$ and any topological space (U, τ) can be considered as a soft topological space over U with respect to a singleton $E = \{e\}$. From this correspondence between (E, U, τ) and $(E \times U, \tau_{E\times U})$, it follows that many results from soft set theory are consequences of the topological results. In fact, (E, U, τ) and $(E \times U, \tau_{E\times U})$ are homeomorphic. A homeomorphism $h : (E \times U, \tau) \to (E, U, \tau)$ is given by $h([a, x]) = (F_{c(a,x)}, E)$, where $F_{c(a,x)}$ is given by $F_{c(a,x)}(e) = \{x\}$ for $e = a$ and $F_{c(a,x)}(e) = \emptyset$ for $e \neq a$.

Similarly, if $E = \{e\}$, then (E, U, τ) and (U, τ) are homeomorphic and a homeomorphism $h : (U, \tau) \to (E, U, \tau)$ is defined by $h(x) = (F_{c(e,x)}, E), x \in U$. This homeomorphism is a very good tool for finding a soft topological space which has a soft property P_1 (for example soft T_i) and it has not a soft property P_2 (soft T_j, $i < j$). Then, it is sufficient to find a topological space which has the property P_1 (T_i) and it has not the property P_2 (T_j, $i < j$, $i, j = 1, 2, 3, 4$).

Acknowledgement

The author acknowledges the financial support by the APVV-0269-11 project "Function spaces, bornologies, hyperspaces and topological structures".

References

[1] Aktaş, H. and Çağman, N. *Soft sets and soft groups*, Information Sciences **177** (2007), 2726–2735.

[2] Ali, M. I., Feng, F., Liu, X., Min, W. K. and Shabir, M. *On some new operations in soft set theory*, Comput. Math. Appl. **57** (2009), 1547–1553.

[3] Aygünoğlu, A. and Aygün, H. *Some notes on soft topological spaces*, Neural. Comput. Appl. **21** (2011), 113–119.

[4] Çağman, N. and Enginoglu, S. *Soft set theory and uni-int decision making*, European J. Oper. Res. **207** (2010), 848–855.

[5] Çağman, N. and Enginoglu, S. *Soft matrix theory and its decision making*, Comput. Math. Appl. **59** (2010), 3308–3314.

[6] Çağman, N., Karataş, S. and Enginoglu, S. *Soft topology*, Comput. Math. Appl. **62** (2011), 351–358.

[7] Chen, D., Tsang, E. C. C., Yeung, D. S. and Wang, X. *The parameterization reduction of soft sets and its applications*, Comput. Math. Appl. **49** (2005), 757–763.

[8] Feng, F., Jun, Y. B. and Zhao, X. *Soft semirings*, Comput.Math. Appl. **56**, no. 10 (2008), 2621–2628.

[9] Georgiou, D. N and Megaritis, A. C. *Soft set theory and topology*, Appl. Gen. Topol. **15**, no. 1 (2014), 93–109.

[10] Georgiou, D. N., Megaritis, A. C. and Petropoulos, V. I. *On Soft Topological Spaces*, Appl. Math. Inf. Sci. **7**, no. 5 (2013), 1889–1901.

[11] Hussain, S. and Ahmad, B. *Some properties of soft topological spaces*, Comput. Math. Appl. **62** (2011), 4058–4067.

[12] Kazanci, O., Yilmaz, S. and Yamak, S. *Soft Sets and Soft BCH-Algebras*, Hacettepe Journal of Mathematics and Statistics **39**, no. 2 (2010), 205–217.

[13] Kharal, A. and Ahmad, B. *Mappings of soft classes*, New Math. Nat. Comput. **7**, no. 3 (2011), 471–481.

[14] Maji, P. K., Roy, A. R. and Biswas, R. *An application of soft sets in a decision making problem*, Comput. Math. Appl. **44** (2002), 1077–1083.

[15] Maji, P. K., Biswas, R. and Roy, A. R. *Soft set theory*, Comput. Math. Appl. **45** (2003), 555–562.
[16] Maji, P. K., Biswas, R. and Roy, A. R. *Fuzzy soft sets*, J. Fuzzy Math. **9**, no. 3 (2001), 589–602.
[17] Min, W. K. *A note on soft topological spaces*, Comput. Math. Appl. **62** (2011), 3524–3528.
[18] Molodtsov, D. A. *Soft set theory-first results*, Comput. Math. Appl. **37** (1999), 19–31.
[19] Molodtsov, D. A. *The description of a dependence with the help of soft sets*, J. Comput. Sys. Sc. Int. **40**, no. 6 (2001), 977–984.
[20] Molodtsov, D. A. *The theory of soft sets* (in Russian), URSS Publishers, Moscow, 2004.
[21] Shabir, M. and Naz, M. *On soft topological spaces*, Comput. Math. Appl. **61** (2011), 1786–1799.
[22] Zorlutuna, I., Akdag, M., Min, W. K. and Atmaca, S. *Remarks on soft topological spaces*, Annals of Fuzzy Mathematics and Informatics **3**, no. 2 (2012), 171–185.
[23] Zou, Y. and Xiao, Z. *Data analysis approaches of soft sets under incomplete information*, Knowl. Base. Syst. **21** (2008), 941–945.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1101 – 1107

Some inclusion results of certain subclass of analytic functions associated with Poisson distribution series

G. Murugusundaramoorthy*†, K.Vijaya‡ and S.Porwal§

Abstract

The purpose of the present paper is to investigate some characterization for Poisson distribution series to be in the new subclasses $\mathcal{G}(\lambda, \alpha)$ and $\mathcal{K}(\lambda, \alpha)$ of analytic functions.

Keywords: Starlike functions, Convex functions, Hadamard product, Poisson distribution series.

2000 AMS Classification: 30C45

Received : 04.05.2015 *Accepted* : 09.09.2015 *Doi* : 10.15672/HJMS.20164513110

1. Introduction and Preliminaries

Let $\mathcal{A}$ be the class of functions f normalized by

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n \tag{1.1}$$

which are analytic in the open disk $\mathbb{U} = \{z : z \in \mathbb{C} \text{ and } |z| < 1\}$. As usual, we denote by $\mathcal{S}$ the subclass of $\mathcal{A}$ consisting of functions which are normalized by $f(0) = 0 = f'(0) - 1$ and also univalent in $\mathbb{U}$. Denote by $\mathcal{T}$ [19] the subclass of $\mathcal{A}$ consisting of functions of the form

$$f(z) = z - \sum_{n=2}^{\infty} a_n z^n, \quad a_n \geq 0, \quad n = 2, 3, \ldots. \tag{1.2}$$

*Department of Mathematics, School of Advanced Sciences, VIT University Vellore - 632014, INDIA, Email: gmsmoorthy@yahoo.com

†Corresponding Author.

‡Department of Mathematics, School of Advanced Sciences, VIT University Vellore - 632014, INDIA, Email: kvijaya@vit.ac.in

§Department of Mathematics, University Institute of Engineering and Technology,Chhatrapati Shahu Ji Maharaj University, Kanpur-208024, (U.P.), INDIA
Email: saurabhjcb@rediffmail.com

Also, for functions $f \in \mathcal{A}$ given by (1.1) and $g \in \mathcal{A}$ given by

$$g(z) = z + \sum_{n=2}^{\infty} b_n z^n,$$

we define the Hadamard product (or convolution) of f and g by

$$(1.3) \qquad (f * g)(z) = z + \sum_{n=2}^{\infty} a_n b_n z^n, \quad (z \in \mathbb{U}).$$

The class $\mathcal{S}^*(\alpha)$ of starlike functions of order α $(0 \leq \alpha < 1)$ may be defined as

$$\mathcal{S}^*(\alpha) = \left\{ f \in \mathcal{A} : \ \Re\left(\frac{zf'(z)}{f(z)}\right) > \alpha, \ z \in \mathbb{U} \right\}.$$

The class $\mathcal{S}^*(\alpha)$ and the class $\mathcal{K}(\alpha)$ of convex functions of order α $(0 \leq \alpha < 1)$

$$\begin{aligned} \mathcal{K}(\alpha) &= \left\{ f \in \mathcal{A} : \ \Re\left(1 + \frac{zf''(z)}{f'(z)}\right) > \alpha, \ z \in \mathbb{U} \right\} \\ &= \left\{ f \in \mathcal{A} : \ zf' \in \mathcal{S}^*(\alpha) \right\} \end{aligned}$$

were introduced by Robertson in [17]. We also write $\mathcal{S}^*(0) = \mathcal{S}^*$, where $\mathcal{S}^*$ denotes the class of functions $f \in \mathcal{A}$ that $f(\mathbb{U})$ is starlike with respect to the origin. Further, $\mathcal{K}(0) = \mathcal{K}$ is the well-known standard class of convex functions. It is an established fact that $f \in \mathcal{K}(\alpha) \Longleftrightarrow zf' \in \mathcal{S}^*(\alpha)$.

A function $f \in \mathcal{A}$ is said to be in the class $f \in \Re^{\tau}(A, B)$ if it satisfies the inequality

$$\left| \frac{f'(z) - 1}{(A - B)\tau - B[f'(z) - 1]} \right| < 1.$$

where $z \in \mathbb{U}, \tau \in \mathbb{C}\backslash\{0\}$, $-1 \leq B < A \leq 1$. The class $\Re^{\tau}(A, B)$ was introduced earlier by Dixit and Pal [6]. If we put

$$\tau = 1, \ A = \alpha \text{ and } \ B = -\alpha \ \ (0 < \alpha \leq 1),$$

we obtain the class of functions $f \in \mathcal{A}$ satisfying the inequality

$$\left| \frac{f'(z) - 1}{f'(z) + 1} \right| < \alpha \quad (z \in \mathbb{U}; 0 < \alpha \leq 1)$$

which was studied by (among others) Padmanabhan [12] and Caplinger and Causey [4].

Very recently, Porwal [13] introduce a power series whose coefficients are probabilities of Poisson distribution

$$K(m, z) = z + \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m} z^n, \qquad (z \in \mathbb{U})$$

and we note that, by ratio test the radius of convergence of above series is infinity. In [13], Porwal also defined the series

$$F(m, z) = 2z - K(m, z) = z - \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m} z^n, \qquad (z \in \mathbb{U}).$$

Now, we considered the linear operator

$$\mathcal{I}(m) : \mathcal{A} \to \mathcal{A}$$

defined by

$$(1.4) \qquad \mathcal{I}(m)f = K(m, z) * f(z) = z + \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m} a_n z^n.$$

Motivated by results on connections between various subclasses of analytic univalent functions by using generalized Bessel functions [1, 2, 3, 8, 11, 13, 15, 22], hypergeometric functions by Srivastava et al. [20] (see [5, 7, 9, 10, 18]) we obtain necessary and sufficient condition for functions $F(m, z)$ in $\mathcal{G}^*(\lambda, \alpha)$ and $\mathcal{K}^*(\lambda, \alpha)$. Further due to the works of Ramesha et al. [16], Padmanabhan [12],we estimate certain inclusion relations between the classes $\Re^\tau(A, B)$, and $\mathcal{G}^*(\lambda, \alpha)$ and $\mathcal{K}^*(\lambda, \alpha)$.

For $0 \leq \lambda < 1$ and $0 \leq \alpha < 1$, we let $\mathcal{G}(\lambda, \alpha)$ the subclass of functions $f \in \mathcal{A}$ which satisfy the condition

$$\Re\left(\frac{zf'(z) + \lambda z^2 f''(z)}{f(z)}\right) > \alpha, \ (z \in \mathbb{U}). \tag{1.5}$$

and also let $\mathcal{K}(\lambda, \alpha)$ the subclass of functions $f \in \mathcal{A}$ which satisfy the condition

$$\Re\left(\frac{z[zf'(z) + \lambda z^2 f''(z)]'}{zf'(z)}\right) > \alpha, \ (z \in \mathbb{U}). \tag{1.6}$$

Also denote $\mathcal{G}^*(\lambda, \alpha) = \mathcal{G}(\lambda, \alpha) \cap \mathcal{T}$ and $\mathcal{K}^*(\lambda, \alpha) = \mathcal{K}(\lambda, \alpha) \cap \mathcal{T}$.

1.1. Remark. It is of interest to note that for $\lambda = 0$, we have $\mathcal{G}(\lambda, \alpha) \equiv \mathcal{S}^*(\alpha)$ and $\mathcal{K}(\lambda, \alpha) \equiv \mathcal{K}(\alpha)$

To prove the main results, we need the following Lemmas.

1.2. Lemma. [21] *A function $f \in \mathcal{A}$ belongs to the class $\mathcal{G}(\lambda, \alpha)$ if*

$$\sum_{n=2}^{\infty}(n + \lambda n(n-1) - \alpha)|a_n| \leq 1 - \alpha.$$

1.3. Lemma. [21] *A function $f \in \mathcal{A}$ belongs to the class $\mathcal{K}(\lambda, \alpha)$ if*

$$\sum_{n=2}^{\infty} n(n + \lambda n(n-1) - \alpha)|a_n| \leq 1 - \alpha.$$

Further we can easily prove that the conditions are also necessary if $f \in \mathcal{T}$.

1.4. Lemma. *A function $f \in \mathcal{T}$ belongs to the class $\mathcal{G}^*(\lambda, \alpha)$ if and only if*

$$\sum_{n=2}^{\infty}(n + \lambda n(n-1) - \alpha)|a_n| \leq 1 - \alpha.$$

1.5. Lemma. *A function $f \in \mathcal{T}$ belongs to the class $\mathcal{K}^*(\lambda, \alpha)$ if and only if*

$$\sum_{n=2}^{\infty} n(n + \lambda n(n-1) - \alpha)|a_n| \leq 1 - \alpha.$$

2. Main Results

2.1. Theorem. *If $m > 0$, then $F(m, z)$ is in $\mathcal{G}^*(\lambda, \alpha)$ if and only if*

$$e^m \left[\lambda m^2 + (1 + 2\lambda)m\right] \leq 1 - \alpha. \tag{2.1}$$

Proof. Since $F(m, z) = z - \sum\limits_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m} z^n$ and by virtue of Lemma 1.4, it suffices to show that

$$\sum_{n=2}^{\infty}(n + \lambda n(n-1) - \alpha)\frac{m^{n-1}}{(n-1)!}e^{-m} \leq 1 - \alpha.$$

Let

$$\mathcal{L}_1(m,\lambda,\alpha)=\sum_{n=2}^{\infty}(n^2\lambda+n(1-\lambda)-\alpha)\frac{m^{n-1}}{(n-1)!}e^{-m}$$

Writing $n^2=(n-1)(n-2)+3(n-1)+1$ and $n=(n-1)+1$, and by simple computation, we get

$$\begin{aligned}\mathcal{L}_1(m,\lambda,\alpha)&=\sum_{n=2}^{\infty}\lambda(n-1)(n-2)\frac{m^{n-1}}{(n-1)!}e^{-m}\\&\quad+(1+2\lambda)\sum_{n=2}^{\infty}(n-1)\frac{m^{n-1}}{(n-1)!}e^{-m}+(1-\alpha)\sum_{n=2}^{\infty}\frac{m^{n-1}}{(n-1)!}e^{-m}\\&=\lambda\sum_{n=3}^{\infty}\frac{m^{n-1}}{(n-3)!}e^{-m}+(1+2\lambda)\sum_{n=2}^{\infty}\frac{m^{n-1}}{(n-2)!}e^{-m}\\&\quad+(1-\alpha)\sum_{n=2}^{\infty}\frac{m^{n-1}}{(n-1)!}e^{-m}\\&=e^{-m}\left[\lambda m^2e^m+(1+2\lambda)me^m+(1-\alpha)(e^m-1)\right]\\&=\lambda m^2+(1+2\lambda)m+(1-\alpha)(1-e^{-m}).\end{aligned}$$

But, this last expression is bounded above by $1-\alpha$ if and only if (2.1) is satisfied. □

2.2. Theorem. *If $m>0$, then $F(m,z)$ is in $\mathcal{K}^*(\lambda,\alpha)$ if and only if*

$$e^m\left[\lambda m^3+(1+5\lambda)m^2+(3+4\lambda-\alpha)m\right]\leq 1-\alpha. \tag{2.2}$$

Proof. Since $F(m,z)=z-\sum_{n=2}^{\infty}\frac{m^{n-1}}{(n-1)!}e^{-m}z^n$ and by virtue of Lemma 1.5, it suffices to show that

$$\sum_{n=2}^{\infty}(n^3\lambda+n^2(1-\lambda)-n\alpha)\frac{m^{n-1}}{(n-1)!}e^{-m}\leq 1-\alpha.$$

Let

$$\mathcal{L}_2(m,\lambda,\alpha)=\sum_{n=2}^{\infty}(n^3\lambda+n^2(1-\lambda)-n\alpha)\frac{m^{n-1}}{(n-1)!}e^{-m}$$

Writing $n^3=(n-1)(n-2)(n-3)+6(n-1)(n-2)+7(n-1)+1$,
$n^2=(n-1)(n-2)+3(n-1)+1$ and $n=(n-1)+1$, we can rewrite the above terms as

$$\begin{aligned}\mathcal{L}_2(m,\lambda,\alpha)&=\lambda\sum_{n=2}^{\infty}(n-1)(n-2)(n-3)\frac{m^{n-1}}{(n-1)!}e^{-m}\\&\quad+(1+5\lambda)\sum_{n=2}^{\infty}(n-1)(n-2)\frac{m^{n-1}}{(n-1)!}e^{-m}\\&\quad+(3+4\lambda-\alpha)\sum_{n=2}^{\infty}(n-1)\frac{m^{n-1}}{(n-1)!}e^{-m}\\&\quad+(1-\alpha)\sum_{n=2}^{\infty}\frac{m^{n-1}}{(n-1)!}e^{-m}\end{aligned}$$

$$= \lambda \sum_{n=4}^{\infty} \frac{m^{n-1}}{(n-4)!} e^{-m} + (1+5\lambda) \sum_{n=3}^{\infty} \frac{m^{n-1}}{(n-3)!} e^{-m}$$
$$+ (3+4\lambda-\alpha) \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-2)!} e^{-m} + (1-\alpha) \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m}$$
$$= e^{-m} \left[\lambda m^3 e^m + (1+5\lambda) m^2 e^m + (3+4\lambda-\alpha) m e^m + (1-\alpha)(e^m - 1)\right]$$
$$= \lambda m^3 + (1+5\lambda) m^2 + (3+4\lambda-\alpha) m + (1-\alpha)(1-e^{-m}).$$

But, this last expression is bounded above by $1-\alpha$ if and only if (2.2) is satisfied. □

By taking $\lambda = 0$,in Theorem 2.1and 2.2 we state the following corollaries:

2.3. Corollary. *If $m > 0$, then $F(m,z)$ is in $\mathcal{S}^*(\alpha)$ if*

(2.3) $$me^m \leq 1-\alpha.$$

2.4. Corollary. *If $m > 0$, then $F(m,z)$ is in $\in \mathcal{K}(\alpha)$ if*

(2.4) $$e^m m(m+3-\alpha) \leq 1-\alpha.$$

3. Inclusion Properties

Making use of the following lemma, we will study the action of the Poisson distribution series on the classes $\mathcal{K}(\lambda, \alpha)$.

3.1. Lemma. [6] *A function $f \in \mathfrak{R}^\tau(A,B)$ is of form (1.1), then*

(3.1) $$|a_n| \leq (A-B)\frac{|\tau|}{n}, \quad n \in \mathbb{N}\backslash\{1\}.$$

The bound given in (3.1) is sharp for

$$f(z) = \int_0^z \left(1 + \frac{(A-B)|\tau| z^{n-1}}{1+Bz^{n-1}}\right) dz \quad (n \geq 2;\ z \in \mathbb{U})$$

3.2. Theorem. *Let $m > 0$. If $f \in \mathfrak{R}^\tau(A,B)$,then $\mathcal{I}(m)f \in \mathcal{K}(\lambda,\alpha)$ if and only if*

(3.2) $$\frac{(A-B)|\tau| e^m \left[\lambda m^2 + (1+2\lambda)m\right]}{1-(A-B)|\tau|(1-e^{-m})} \leq 1-\alpha$$

where $\tau \in \mathbb{C}\backslash\{0\}, -1 \leq B < A \leq 1$.

Proof. Let f be of the form (1.1) belong to the class $\mathfrak{R}^\tau(A,B)$ then by virtue of Lemma 1.5, it suffices to show that

$$\sum_{n=2}^{\infty} n(n^2\lambda + n(1-\lambda) - \alpha)\frac{m^{n-1}}{(n-1)!} e^{-m} |a_n| \leq 1-\alpha.$$

Let

$$\mathcal{L}_3(m,\lambda,\alpha) = \sum_{n=2}^{\infty} n(n^2\lambda + n(1-\lambda) - \alpha)\frac{m^{n-1}}{(n-1)!} e^{-m} |a_n|.$$

Since $f \in \mathfrak{R}^\tau(A,B)$ by Lemma 3.1 we have $|a_n| \leq (A-B)\frac{|\tau|}{n}$, $n \in \mathbb{N}\backslash\{1\}$, hence we get

$$\mathcal{L}_3(m,\lambda,\alpha) \leq e^{-m} \sum_{n=2}^{\infty} (n^2\lambda + n(1-\lambda) - \alpha)\frac{m^{n-1}}{(n-1)!}(A-B)|\tau|$$
$$\leq (A-B)|\tau| e^{-m} \sum_{n=2}^{\infty} (n^2\lambda + n(1-\lambda) - \alpha)\frac{m^{n-1}}{(n-1)!}$$

Writing $n^2 = (n-1)(n-2) + 3(n-1) + 1$ and $n = (n-1) + 1$, and by using the similar arguments as in the proof of Theorem 2.1, we get

$$\mathcal{L}_3(m, \lambda, \alpha) \leq (A - B)|\tau| \left[\lambda m^2 + (1 + 2\lambda)m + (1 - \alpha)(1 - e^{-m})\right].$$

But,the last expression is bounded above by $1 - \alpha$ if and only if (3.2) is satisfied. Hence the proof is completed. □

3.3. Corollary. *Let $m > 0$ and $\lambda = 0$. If $f \in \mathfrak{R}^\tau(A, B)$, then $\mathcal{I}(m)f \in \mathcal{K}(\alpha)$ if and only if*

$$(A - B)|\tau|m \left[1 - (A - B)|\tau|(1 - e^{-m}\right]^{-1} \leq 1 - \alpha$$

where $\tau \in \mathbb{C}\backslash\{0\} - 1 \leq B < A \leq 1$.

3.4. Theorem. *Let $m > 0$, then*

$$G(m, z) = \int_0^z \frac{F(m, t)}{t})dt$$

is in $\mathcal{K}^*(\lambda, \alpha)$ if and only if

$$e^m \left[\lambda m^2 + (1 + 2\lambda)m\right] \leq 1 - \alpha. \tag{3.3}$$

Proof. Since

$$G(m, z) = z - \sum_{n=2}^{\infty} \frac{e^{-m}m^{n-1}}{(n-1)!} \frac{z^n}{n} = z - \sum_{n=2}^{\infty} \frac{e^{-m}m^{n-1}}{n!} z^n$$

by Lemma 1.5, we need only to show that

$$\sum_{n=2}^{\infty} n(n^2\lambda + n(1 - \lambda) - \alpha) \frac{m^{n-1}}{n!} e^{-m} \leq 1 - \alpha.$$

Now, let

$$\begin{aligned}\mathcal{L}_4(m, \lambda, \alpha) &= \sum_{n=2}^{\infty} n(n^2\lambda + n(1 - \lambda) - \alpha) \frac{m^{n-1}}{n!} e^{-m} \\ &= \sum_{n=2}^{\infty} (n^2\lambda + n(1 - \lambda) - \alpha) \frac{m^{n-1}}{(n-1)!} e^{-m}.\end{aligned}$$

Hence ,writing $n^2 = (n-1)(n-2) + 3(n-1) + 1$ and $n = (n-1) + 1$, and by using the similar arguments as in the proof of Theorem 2.1, we have

$$\mathcal{L}_4(m, \lambda, \alpha) \leq \lambda m^2 + (1 + 2\lambda)m + (1 - \alpha)(1 - e^{-m}),$$

which is bounded above by $1 - \alpha$ if and only if (3.3) holds. □

3.5. Theorem. *Let $m > 0$, then $G(m, z) = \int_0^z \frac{F(m,t)}{t})dt$ is in $\mathcal{G}^*(\lambda, \alpha)$ if and only if*

$$m\lambda + \left(1 - \frac{\alpha}{m}\right)\left(1 - e^{-m}\right) + \alpha e^{-m} \leq 1 - \alpha. \tag{3.4}$$

Proof. The proof of theorem is similar to that of Theorem 3.4, hence we omit the details involved. □

Acknowledgement: The authors thank the referee for his insightful suggestions to improve the paper in present form.

References

[1] Baricz, A.*Geometric properties of generalized Bessel functions*, Publ. Math. Debrecen, **73** (1-2), 155–178, 2008.

[2] Baricz, A. *Geometric properties of generalized Bessel functions of complex order*, Mathematica(Cluj) **48** (71) (1), 13–18, 2006.

[3] Baricz, A. *Generalized Bessel functions of the first kind, PhD thesis*, Babes-Bolyai University, Cluj-Napoca, 2008.

[4] Caplinger, T. R. and Causey, W. M. *A class of univalent functions*, Proc. Amer. Math. Soc. **39**, 357–361, 1973.

[5] Cho, N. E. Woo, S. Y. and. Owa, S. *Uniform convexity properties for hypergeometric functions*, Fract. Cal. Appl. Anal., **5** (3), 303–313, 2002.

[6] Dixit, K.K. and Pal, S.K. *On a class of univalent functions related to complex order*, Indian J. Pure. Appl. Math. **26** (9), 889–896, 1995.

[7] Merkes, E. and Scott, B. T. *Starlike hypergeometric functions*, Proc. Amer. Math. Soc. **12**, 885–888, 1961.

[8] Mondal, S. R. and Swaminathan, A. *Geometric properties of generalized Bessel functions*, Bull. Malaysian Math. Sci. Soc. **35** (1), 179–194, 2012.

[9] Mostafa, A. O. *A study on starlike and convex properties for hypergeometric functions*, J. Inequal. Pure Appl. Math. **10** (3), 1–16, 2009.

[10] Murugusundaramoorthy, G. and Magesh, N. *On certain subclasses of analytic functions associated with hypergeometric functions*, Appl. Math. Lett. **24**, 494–500, 2011.

[11] Murugusundaramoorthy, G. and Janani, T. *An application of generalized Bessel functions on certain subclasses of analytic functions*, Turkish Journal of Analysis and Number Theory **3** (1), 1-6, 2015.

[12] Padmanabhan,K. S. *On sufficient conditions for starlikeness*, Indian J. Pure Appl. Math. **32**, 543–550, 2001.

[13] Porwal, S. *An application of a Poisson distribution series on certain analytic functions* , J. Complex Anal. **2014**, Article ID 984135, 3 pages, 2014.

[14] Porwal, S. *Mapping properties of generalized Bessel functions on some subclasses of univalent functions*, Anal. Univ. Oradea Fasc. Matematica, **20** (2), 51-60, 2013.

[15] Porwal, S. and Dixit, K.K. *An application of generalized Bessel functions on certain analytic functions*, Acta Univ. Matthiae Belii, series Mathematics,**2013** 51–57, 2013 .

[16] Ramesha, C. Kumar, S. and Padmanabhan, K.S. *A sufficient condition for starlikeness* , Chinese J. Math. **23** (2), 167–171, 1995.

[17] Robertson, M. S. *On the theory of univalent functions* , Ann. Math. **37**, 374–408, 1936.

[18] Silverman, H. *Starlike and convexity properties for hypergeometric functions* , J. Math. Anal. Appl. **172** (3), 574–581, 1993.

[19] Silverman, H. *Univalent functions with negative coefficients*, Proc. Amer. Math. Soc. **51**, 109–116, 1975.

[20] Srivastava, H. M. Murugusundaramoorthy, G.and Sivasubramanian, S.*Hypergeometric functions in the parabolic starlike and uniformly convex domains*, Integr. Transf. Spec. Func. **18**, 511–520, 2007.

[21] Thulasiram, T. Suchithra, K. Sudharsan, T. V. and Murugusundaramoorthy, G. *Some inclusion results associated with certain subclass of analytic functions involving Hohlov operator*, Rev. R. Acad. Cienc. Exactas, Fis. Nat. Ser.- A Math.**108** (2), 711–720, 2014.

[22] Vijaya, K .Murugusundaramoorthy, G. and Kasthuri, M. *A note on subclasses of starlike and convex functions associated with Bessel functions*, J. Nonlinear Funct. Anal. **2014**, 2014:11.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1109 – 1124

Numerical solution of Burgers equation with nonlinear damping using non-polynomial tension spline

H.S.Shekarabi* and J.Rashidinia†‡

Abstract

Numerical solution of Burgers equation with nonlinear damping term has been investigated.We developed new approach based on non-polynomial cubic tension spline approximation.The proposed approach depends on the parameters involving in tension spline.By choosing suitable values of such parameters the optimal local truncation error of the scheme can be obtained.Convergence analysis of presented method has been discussed in details and we have shown under appropriate condition the method convergence.The method tested on two problems,numerical results have been compared with the exact solution to justify the usefulness and accurate nature of proposed method.

Keywords: Burgers equation,Non-polynomial tension spline,Convergence analysis.

2000 AMS Classification: 65M06,12,99.

Received : 24.05.2015 *Accepted* : 18.08.2015 *Doi* : 10.15672/HJMS.20164513111

*Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran
Email: hoda.shekarabi@kiau.ac.ir
†School of Mathematics,Iran University of science and Technology
Email: rashidinia@iust.ac.ir
‡Corresponding Author.

1. Introduction

The Burgers equation introduced by Burgers [1] provide fundamental pedagogical examples for many important equation in nonlinear Partial Differential equations such as traveling waves,shock formation,similarity solutions and singular perturbations [14,27,40,43], it appears in some of condensed matter and statistical Physics and non-physics problems such as vehicular traffic [7], The Kardar-Parisi-Zhang or KPZ equation[23,2],traffic flow,shallow water waves, gas dynamics, and fluids with the dissipative viscous behavior[28,29,39,30].
Furthermore,Burgers equation is studying in directed polymers[24,3] and has found interesting applications in cosmology, such as "Zel'dovich approximation"[46] and "adhesionmodel" [16]. Another application of Burgers equation is in the theory of turbulence and field[34,37,15,31].
Hofe[20] and Cole[8] have shown the Burgers equation can reduce to heat conduction equation.
In this paper, we investigate the solution of generalized one-dimensional Burgers equation with nonlinear damping term [38] of the form

$$u_t + auu_x - u_{xx} = g(u), \qquad x \in \Omega = [c, d], \quad t \geq 0, \tag{1.1}$$

with the initial condition

$$u(x, t_0) = \phi(x), \tag{1.2}$$

and the boundary conditions

$$\begin{aligned} &u(c, t) = p_0(t), \\ &u(d, t) = p_1(t), \end{aligned} \tag{1.3}$$

where $u(x, t)$ indicates the velocity for the space x and time t, a is parameter and $g(u)$ is damping term.

The Burgers equation in the first term is an unsteady term,the second and third term represents nonlinear convection and diffusion problem,with nonlinearity term,can be survived by many researcher.
Different numerical technique have been used for solving Burgers equation.Finite difference methods have been given by Biringen et al[4] and Kutluay et al.[25]and recently,by Inan et al [21].Finite elements methods have been given by Caldwell et al.[9]and Varglu et al.[42] and Ozis et al.[35].Spectral methods have been developed by Bar-yoseph et al.[5]and Mansell et al.[32].Pseudo-spectral method has been used by Darvishi et al.[12] and distributed approximation function approach has studied by Zang et al.[47] and Wei et al.[44].Boundary elements methods is given by Bahadir et al.[6].A wavelet collocation method has used by Garba [17],furthermore quasi wavelet based numerical method has been suggested by Wan et al.[45].Fast adaptive diffusion wavelet method have been survived by Goyal et al.[18].Least square quadratic B-spline finite elements has been given by Kutluay et al.[26].Various B-spline have been proposed by Dag et al.[13,41].B-spline and multi-quadratic quasi-interpolation have been described by Zhu et al and Chen et al.[48,10].
The present work attempts to use cubic non-polynomial spline[36,19,33].One of the important ability of this approximation is the tension parameters involving definition of non-polynomial cubic spline which can be chosen in such a way that the local truncation error of the proposed method can be optimal.Hence,it has been demonstrate that tension spline give better result.This paper is organized as follows:In section 2,derivation and formulation of the cubic non-polynomial tension spline along with consistency relation of second derivatives discussed in details.In section 3,the derivation of two level scheme

based on non-polynomial tension spline has been described.In section 4,convergence analysis of the present method has been discussed in detail and we have shown under appropriate condition the method converges.At the end,we illustrate the accuracy and efficiency of the proposed method by testing this approach on two test problems.comparison of the numerical result are given.

2. Non-polynomial tension spline

Following our earlier works,let $s(x)$ of class $c^2[c,d]$ be non-polynomial tension spline interpolating the function $u(x)$ at the grid point $x_l, l = 0, 1, 2, ..., n$.For each segment $[x_l, x_{l+1}], l = 0, 1, ..., n-1$,the non-polynomial $s(x)$ defined by

$$s(x) = a_l + b_l(x - x_l) + c_l(e^{\omega(x-x_l)} - e^{-\omega(x-x_l)} + d_l(e^{\omega(x-x_l)} + e^{-\omega(x-x_l)}), \tag{2.1}$$

where the a_l, b_l, c_l, d_l are unknown coefficients and ω is arbitrarily parameter. To determine the unknown coefficient in (2.1) we denote the following relations

$$\begin{aligned} s(x_l) &= u_l, & s(x_{l+1}) &= u_{l+1}, \\ s'(x_l) &= m_l, & s'(x_{l+1}) &= m_{l+1} \\ s''(x_l) &= M_l, & s''(x_{l+1}) &= M_{l+1}. \end{aligned} \tag{2.2}$$

The first and second derivatives of non-polynomial tension spline function $s(x)$ are

$$s' = b_l + \omega c_l(e^{\omega(x-x_l)} + e^{-\omega(x-x_l)}) + \omega d_l(e^{\omega(x-x_l)} - e^{-\omega(x-x_l)}), \quad l = 1, 2, ..., n \tag{2.3}$$

$$s'' = -\omega^2(c_l(e^{\omega(x-x_l)} - e^{-\omega(x-x_l)}) + d_l(e^{\omega(x-x_l)} + e^{-\omega(x-x_l)}), \quad l = 1, 2, ..., n \tag{2.4}$$

Now using (2.2)-(2.4) and after some algebraic manipulation,we can determine the unknown coefficients in (2.1) as

$$a_l = u_l - \frac{M_l}{\omega^2}, \qquad b_l = \frac{u_{l+1} - u_l}{h} + \frac{M_l - M_{l+1}}{\omega\theta},$$
$$c_l = \frac{2M_{l+1} - (e^{\theta} + e^{-\theta})M_l}{2\omega^2(e^{\theta} - e^{-\theta})}, \qquad d_l = \frac{M_l}{2\omega^2},$$

where $h = \frac{d-c}{n}$, $\theta = \omega h$.

Using the continuity of the first derivative at (x_l, u_l),that is $s'(x_l^-) = s'(x_l^+)$.We obtain the following equation for $l = 1, ..., n$.

$$\frac{u_{l+1} - 2u_l + u_{l-1}}{h^2} = \alpha M_{l+1} + 2\beta M_l + \alpha M_{l-1}, \tag{2.5}$$

where

$$\alpha = \frac{1}{\theta^2}(1 - \frac{2\theta}{e^{\theta} - e^{-\theta}}), \quad \beta = \frac{1}{\theta^2}(\frac{\theta(e^{\theta} + e^{-\theta})}{e^{\theta} - e^{-\theta}} - 1).$$

When $\omega \to 0$,that $\theta \to 0$,then $(\alpha, \beta) \to (\frac{1}{6}, \frac{1}{3})$, and the relations defined by (2.5) reduced into construction relation of conventional cubic spline.

Now by using the continuous of the first derivative,we have

$$s'(x_l^+) = u_{x_l} = \frac{u_{l+1} - u_l}{h} - h[\alpha M_{l+1} + \beta M_l] \tag{2.6}$$

$$s'(x_l^-) = u_{x_l} = \frac{u_l - u_{l-1}}{h} + h[\beta M_l + \alpha M_{l-1}] \tag{2.7}$$

combining (2.6) and (2.7),we obtain

$$m_l = s'(x_l) = u_{x_l} = \frac{u_{l+1} - u_{l-1}}{2h} - \frac{\alpha h}{2}[M_{l+1} + M_{l-1}] \tag{2.8}$$

Similarly ,we have

$$m_{l+1} = s'(x_{l+1}) = u_{x_{l+1}} = \frac{u_{l+1} - u_l}{h} + h[\beta M_{l+1} + \alpha M_l] \tag{2.9}$$

$$m_{l-1} = s'(x_{l-1}) = u_{x_{l-1}} = \frac{u_l - u_{l-1}}{h} - h[\beta M_{l-1} + \alpha M_l] \tag{2.10}$$

3. The method based on tension spline

The notation u_l^j is used for the discrete approximation value of $u(x_l, t_j)$,$l = 0, 1, ..., n$ and $j = 0, 1, ..., m$,in which n and m are integer and $x_l = c + lh$ and $t_j = t_0 + jk$,where k is the step size in t direction.

We consider the following finite difference approximation

$$\bar{u}_{t_l}^j = \frac{u_l^{j+1} - u_l^j}{k} = u_{t_l}^j + O(k) \tag{3.1}$$

$$\bar{u}_{t_{l+1}}^j = \frac{u_{l+1}^{j+1} - u_{l+1}^j}{k} = u_{t_{l+1}}^j + O(k) \tag{3.2}$$

$$\bar{u}_{t_{l-1}}^j = \frac{u_{l-1}^{j+1} - u_{l-1}^j}{k} = u_{t_{l-1}}^j + O(k) \tag{3.3}$$

$$\bar{u}_{x_l}^j = \frac{u_{l+1}^j - u_{l-1}^j}{2h} = u_{x_l}^j + O(h^2) \tag{3.4}$$

$$\bar{u}_{x_{l+1}}^j = \frac{3u_{l+1}^j - 4u_l^j + u_{l-1}^j}{2h} = u_{x_{l+1}}^j + O(h^2) \tag{3.5}$$

$$\bar{u}_{x_{l-1}}^j = \frac{-3u_{l-1}^j + 4u_l^j - u_{l+1}^j}{2h} = u_{x_{l-1}}^j + O(h^2), \tag{3.6}$$

By replacing space derivatives by non-polynomial tension spline

$$\bar{u}_{xx_l}^j = s''(x_l, t_j) = M_l^j + O(h^2) \tag{3.7}$$

$$\bar{u}_{x_l}^j = s'(x_l, t_j) = m_l^j + O(h^3) \tag{3.8}$$

By using the relations of (3.1),(3.7) and (3.8),we can obtain the new approximate solution of equation (1) as

$$(\frac{u_l^{j+1} - u_l^j}{k}) + au_l^j.m_l^j - g_l^j = M_l^j, \tag{3.9}$$

where m_l^j similar to (2.8) in jth time level and $g_l^j = g(u_l^j)$, $l = 1(1)n - 1$.
Furthermore,similar to (2.5) in jth time level,we get

$$\frac{u_{l+1}^j - 2u_l^j + u_{l-1}^j}{h^2} = \alpha M_{l+1}^j + 2\beta M_l^j + \alpha M_{l-1}^j. \tag{3.10}$$

We substitute (3.9) in (3.10) and by the help of (3.1)-(3.3),we obtain

$$\begin{aligned}\frac{u_{l+1}^j - 2u_l^j + u_{l-1}^j}{h^2} &= \alpha(\frac{u_{l+1}^{j+1} - u_{l+1}^j}{k}) + \alpha a u_{l+1}^j.m_{l+1}^j - \alpha g_{l+1}^j \\ &+ 2\beta(\frac{u_l^{j+1} - u_l^j}{k}) + 2\beta a u_l^j.m_l^j - 2\beta g_l^j \\ &+ \alpha(\frac{u_{l-1}^{j+1} - u_{l-1}^j}{k}) + \alpha a u_{l-1}^j.m_{l-1}^j - \alpha g_{l-1}^j.\end{aligned} \tag{3.11}$$

Now by using equations of (3.4)-(3.6) in (2.8)-(2.10),we have

$$
\begin{aligned}
m_l^j &= \frac{u_{l+1}^j - u_{l-1}^j}{2h} - \frac{\alpha h}{2}\left(\frac{u_{l+1}^{j+1} - u_{l+1}^j}{k}\right) - \frac{\alpha h a}{2} u_{l+1}^j\left(\frac{3u_{l+1}^j - 4u_l^j + u_{l-1}^j}{2h}\right) \\
&+ \frac{\alpha h}{2}\left(\frac{u_{l-1}^{j+1} - u_{l-1}^j}{k}\right) + \frac{\alpha h a}{2} u_{l-1}^j\left(\frac{-3u_{l-1}^j + 4u_l^j - u_{l+1}^j}{2h}\right) + \frac{\alpha h}{2} g_{l+1}^j - \frac{\alpha h}{2} g_{l-1}^j
\end{aligned} \tag{3.12}
$$

$$
\begin{aligned}
m_{l+1}^j &= \frac{u_{l+1}^j - u_l^j}{h} + \beta h\left(\frac{u_{l+1}^{j+1} - u_{l+1}^j}{k}\right) + \beta h a u_{l+1}^j\left(\frac{3u_{l+1}^j - 4u_l^j + u_{l-1}^j}{2h}\right) \\
&+ \alpha h\left(\frac{u_l^{j+1} - u_l^j}{k}\right) + \alpha h a u_l^j\left(\frac{u_{l+1}^j - u_{l-1}^j}{2h}\right) - \beta h g_{l+1}^j - \alpha h g_l^j
\end{aligned} \tag{3.13}
$$

$$
\begin{aligned}
m_{l-1}^j &= \frac{u_l^j - u_{l-1}^j}{h} - \beta h\left(\frac{u_{l-1}^{j+1} - u_{l-1}^j}{k}\right) - \beta h a u_{l-1}^j\left(\frac{-3u_{l-1}^j + 4u_l^j - u_{l+1}^j}{2h}\right) \\
&- \alpha h\left(\frac{u_l^{j+1} - u_l^j}{k}\right) - \alpha h a u_l^j\left(\frac{u_{l+1}^j - u_{l-1}^j}{2h}\right) + \beta h g_{l-1}^j + \alpha h g_l^j
\end{aligned} \tag{3.14}
$$

By substituting equations (3.12)-(3.14) in (3.11),we can obtain

$$
\begin{aligned}
&(b_0)_l^j u_{l+1}^{j+1} + (b_1)_l^j u_{l-1}^{j+1} + (b_2)_l^j u_l^{j+1} = b_3(u_{l+1}^j + u_{l-1}^j) + b_4 u_l^j \\
&+ b_5(g_{l+1}^j + g_{l-1}^j) + b_{10} g_l^j + b_6((u_{l+1}^j)^2 - (u_{l-1}^j)^2) - b_7 u_l^j(u_{l+1}^j - u_{l-1}^j) \\
&- b_8 u_l^j(B_l^j) - b_9(u_{l+1}^j(A_l^j) + u_{l-1}^j(C_l^j)), \\
&l = 1(1)n-1,
\end{aligned} \tag{3.15}
$$

where

$$
\begin{aligned}
&(b_0)_l^j = h^3\alpha\beta a(u_{l+1}^j - u_l^j) + h^2\alpha, \\
&(b_1)_l^j = h^3\alpha\beta a(u_l^j - u_{l-1}^j) + h^2\alpha, \\
&(b_2)_l^j = h^3\alpha^2 a(u_{l+1}^j - u_{l-1}^j) + 2\beta h^2, \\
&b_3 = h^2\alpha + k, \\
&b_4 = 2\beta h^2 - 2k, \\
&b_5 = kh^2\alpha, \\
&b_6 = h^3\alpha\beta a, \\
&b_7 = h^3\alpha a(\alpha - \beta), \\
&b_8 = 2\beta k h^2 a, \\
&b_9 = akh^2\alpha, \\
&b_{10} = 2\beta k h^2,
\end{aligned}
$$

and

$$A_l^j = \frac{u_{l+1}^j - u_l^j}{h} + \beta hau_{l+1}^j(\frac{3u_{l+1}^j - 4u_l^j + u_{l-1}^j}{2h})$$
$$+ \alpha hau_l^j(\frac{u_{l+1}^j - u_{l-1}^j}{2h}) - (h\beta g_{l+1}^j + \alpha hg_l^j),$$
$$B_l^j = \frac{u_{l+1}^j - u_{l-1}^j}{2h} - \frac{\alpha h}{2}au_{l+1}^j(\frac{3u_{l+1}^j - 4u_l^j + u_{l-1}^j}{2h})$$
$$+ \frac{\alpha h}{2}au_{l-1}^j(\frac{-3u_{l-1}^j + 4u_l^j - u_{l+1}^j}{2h}) + (\frac{\alpha h}{2}g_{l+1}^j - \frac{\alpha h}{2}g_{l-1}^j),$$
$$C_l^j = \frac{u_l^j - u_{l-1}^j}{h} - \beta hau_{l-1}^j(\frac{-3u_{l-1}^j + 4u_l^j - u_{l+1}^j}{2h})$$
$$- \alpha hau_l^j(\frac{u_{l+1}^j - u_{l-1}^j}{2h}) + (h\beta g_{l-1}^j + \alpha hg_l^j).$$

The above system can be associated with boundary conditions.By solving this system the approximate solution can be obtain.

3.1. The appropriate parameters. Using Taylor expansion about the grid point $u(x_l, t_j)$,finally we obtain the local truncation error

$$T_l^j = k^2h^2\{\alpha + \beta\}\frac{\partial^2 u}{\partial t^2} + h^4k\{\alpha - \frac{1}{12}\}\frac{\partial^4 u}{\partial x^4} + \frac{h^4k^2}{2}\alpha\frac{\partial^4 u}{\partial t^2 \partial x^2} + ...,$$

The consistency relation for (2.5) lead to the equation $2\alpha + 2\beta = 1$,by simplifying the above equation and choose $\alpha = \frac{1}{12}$ and $\beta = \frac{5}{12}$ obtain the scheme of $O(k^2 + h^4 + h^2k^2)$.

4. Convergence of the method

Here we analyze the convergence of the system,we can write system (3.15) in the matrix form

$$PU^{j+1} = QU^j + G(U^j) \tag{4.1}$$

P is tri-diagonal matrix with variable entries,Q is coefficient matrix of U^j with constant entries and $G(U^j)$ is nonlinear terms in this system.
to prove convergence,we suppose $a > 0$,$\rho = \max |u_l^j|, l = 1(1)n - 1, j = 0(1)m$.
In this paper $\|.\|$ means $\|.\|_\infty$.

4.1. Lemma. P is nonsingular.
proof. It is sufficient to solve that P is strictly diagonally dominant.Therefore we must prove

$$| (b_0)_l^j + (b_1)_l^j | \leq | (b_2)_l^j | \tag{4.2}$$

We have

$$| (b_0)_l^j + (b_1)_l^j | = | h^3\alpha\beta a(u_{l+1}^j - u_l^j) + h^3\alpha\beta a(u_l^j - u_{l-1}^j) + 2h^2\alpha |$$
$$= | h^3\alpha\beta a(u_{l+1}^j - u_{l-1}^j) | + 2h^2\alpha$$
$$\geq 2h^2\alpha - | h^3\alpha\beta a(u_{l+1}^j - u_{l-1}^j) | \tag{4.3}$$

By using inequality (4.3) in left hand side of equation (4.2),we obtain

$$2h^2\alpha - | h^3\alpha\beta a(u_{l+1}^j - u_{l-1}^j) | \leq | h^3\alpha^2 a(u_{l+1}^j - u_{l-1}^j) | + 2\beta h^2$$
$$2h^2(\alpha - \beta) \leq a(h^3\alpha\beta + h^3\alpha^2) | u_{l+1}^j - u_{l-1}^j | . \tag{4.4}$$

We know $\alpha - \beta = -\frac{1}{3}$,therefore inequality (4.4) is obvious and proof complete.

4.1. Theorem. *The discrete numerical scheme defined by* (3.15) *is convergent,provided that* $\|N\| \leq h^2(2 + h(\frac{a\rho}{24}))$.

proof. *We assume that* U^{j+1} *and* $\hat{U}^{j+1}$ *are exact and approximation solution of* (4.1), *respectively.The error in the solution is:*

$$U^{j+1} - \hat{U}^{j+1} = P^{-1}Q(U^j - \hat{U}^j) + P^{-1}[G(U^j) - G(\hat{U}^j)] \quad j = 0(1)m$$

$$E^{j+1} = P^{-1}QE^j + P^{-1}[G(U^j) - G(\hat{U}^j)] \tag{4.5}$$

where $E = (e_1, e_2, ..., e_n)^T$

Following [11] we have

$$G(U^j) - G(\hat{U}^j) = E^j N \tag{4.6}$$

N *is the coefficient matrix of the nonlinear term.*

Now by using of equation (4.6) *in* (4.5)*we obtain*

$$E^{j+1} = P^{-1}QE^j + P^{-1}E^j N \tag{4.7}$$

Using the infinity norm,we can write

$$\begin{aligned}
\| E^{j+1}\| &\leq \|P^{-1}Q + P^{-1}N\|.\|E^j\| \\
\| E^{j+1}\| &\leq \|P^{-1}(Q + N)\|.\|E^j\| \\
\| E^{j+1}\| &\leq \|P^{-1}\|.\|(Q + N)\|.\|E^j\| \\
&\leq (\|P^{-1}\|.\|(Q + N)\|)^2\|E^{j-1}\| \\
&\vdots \\
\| E^{j+1}\| &\leq (\|P^{-1}\|.\|(Q + N)\|)^{j+1}\|E^0\|
\end{aligned}$$

The method is convergent if

$$\|P^{-1}\|.\|(Q + N)\| \leq 1$$

$$\|(Q + N)\| \leq \frac{1}{\|P^{-1}\|}$$

$$|\|N\| - \|Q\|| \leq \|(Q + N)\| \leq \frac{1}{\|P^{-1}\|}. \tag{4.8}$$

Since $\|P\|\|P^{-1}\| \geq 1$,*we have*

$$\|N\| - \|Q\| \leq \frac{1}{\|P^{-1}\|} \leq \|P\|$$

$$\|N\| \leq \|Q\| + \|P\| \tag{4.9}$$

By simple calculation,we achieve

$$\|Q\| = h^2, \tag{4.10}$$

$$\|P\| \leq h^2 + h^3 a(\frac{\rho}{24}) \tag{4.11}$$

By substitute (4.10) *and* (4.11) *in* (4.9),*the proof complete.*

5. Numerical illustrations

To illustrate accuracy and ability of the proposed method,we considered two examples. Note that,the proposed non-linear tension spline is a two-level scheme therefore the starting level can be determined by the given initial condition.Finally we solve the arising

system.
Example1. *We consider equation*

$$u_t + auu_x = u_{xx}, \quad t \geq 0, \quad 0 \leq x \leq 1$$

with the following initial and boundary conditions

$$\begin{aligned} u(x,0) &= \sin(\pi x), \qquad 0 \leq x \leq 1 \\ u(0,t) &= 0, \\ u(1,t) &= 0, \quad t \geq 0. \end{aligned}$$

The exact solution of the above equation is taken

$$\begin{aligned} u(x,t) &= \frac{2\pi \sum_{n=1}^{\infty} a_n exp(-n^2\pi^2 t) n \sin(n\pi x)}{a_0 + \sum_{n=1}^{\infty} a_n exp(-n^2\pi^2 t)\cos(n\pi x)} \\ a_0 &= \int_0^1 exp\{(-2\pi)^{-1}[1-\cos(\pi x)]\}dx \\ a_n &= 2\int_0^1 exp\{(-2\pi)^{-1}[1-\cos(\pi x)]\}\cos(n\pi x)dx \qquad n = 1,2,3,... \end{aligned}$$

Example 1 is the Burgers equation without damping terms.The proposed scheme (3.15) *applied on example 1,with* $a = 0.1$ *and* 0.01*,*$k = 0.00001$ *and values of step size* $h = 0.02$ *and* $h = 0.01$ *for* $t_f = 0.1$*.The computed solution are compare with exact solution,the maximum absolute errors are tabulated in table 1.In table 2,we take* $h = 0.1$ *and* $k = 0.001$*,the results are computed for different time levels and different* a*.the maximum absolute error are tabulated in table 2. In Figures* $1-3$*,we show the graphs between exact and numerical solutions at* $t = 1, t = 3$ *and* $t = 5$ *in different* a*.*

Table 1. Maximum absolute error for example 1

x	$a = 0.1$		$a = 0.01$	
	$h = 0.02$	$h = 0.01$	$h = 0.02$	$h = 0.01$
0.10	8.45119(−4)	2.07841(−5)	8.37944(−7)	2.0917(−7)
0.20	1.44193(−4)	3.56293(−5)	1.43739(−6)	3.58991(−7)
0.30	1.83155(−4)	4.54565(−5)	1.83520(−6)	4.58581(−7)
0.40	2.04631(−4)	5.09991(−5)	2.06059(−6)	5.15176(−7)
0.50	2.10730(−4)	5.27339(−5)	2.13234(−6)	5.33425(−7)
0.60	2.02309(−4)	5.08367(−5)	2.05703(−6)	5.14941(−7)
0.70	1.78959(−4)	4.51683(−5)	1.82853(−6)	4.58146(−7)
0.80	1.39121(−4)	3.52934(−5)	1.42858(−6)	3.58431(−7)
0.90	6.59886(−4)	2.05162(−5)	8.28727(−7)	2.08597(−7)

Table 2. Maximum absolute error for example 1

x	a	t=*1*	t=*3*	t=*5*
0.1	0.1	3.05996(−5)	3.06791(−5)	3.07569(−5)
	0.01	3.07903(−7)	3.07924(−7)	3.79320(−7)
	0.001	3.08122(−9)	2.08135(−9)	3.08135(−9)
0.5	0.1	8.37130(−5)	8.39331(−5)	8.41479(−5)
	0.01	8.45321(−7)	8.45385(−7)	8.45406(−7)
	0.001	8.46212(−9)	8.46255(−9)	8.46255(−9)
0.9	0.1	2.70077(−5)	2.70793(−5)	2.71493(−5)
	0.01	2.73069(−7)	2.73089(−7)	2.73096(−7)
	0.001	2.73389(−9)	2.73402(−9)	2.73402(−9)

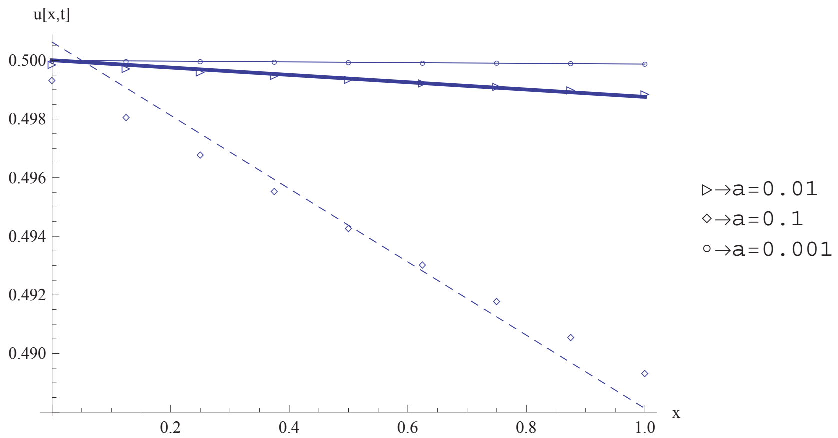


Figure 1:Approximate and exact solution for example 1 at $t = 1$,*with different a*

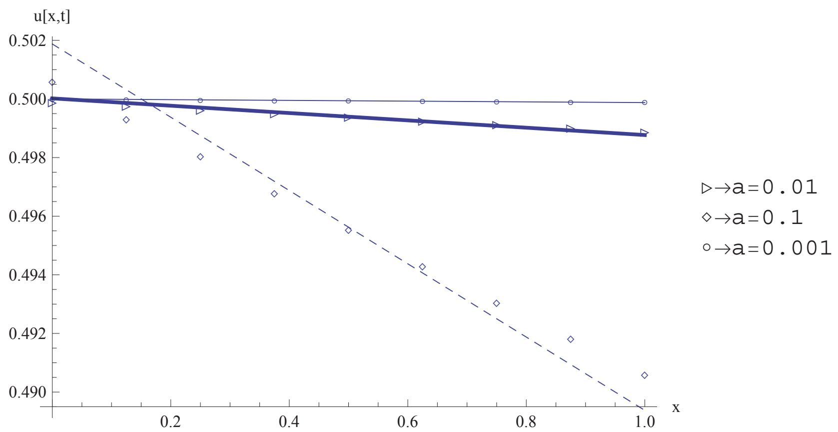


Figure 2:Approximate and exact solution for example 1 at $t = 3$,*with different a*

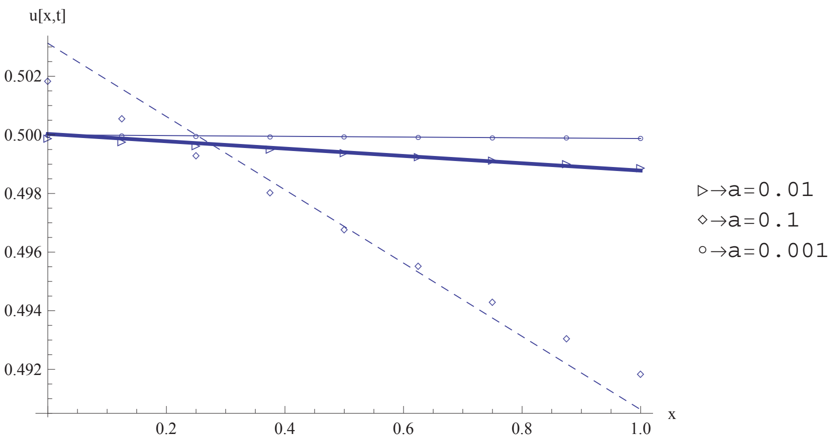


Figure 3:Approximate and exact solution for example 1 at $t = 5$,with different a

Example2. *We consider nonlinear damping equation*

$$u_t + auu_x = u_{xx} + bu(1-u), \quad t \geq 0, \quad 0 \leq x \leq 1$$

with the following initial and boundary conditions

$$u(x,0) = \frac{1}{2} + \frac{1}{2}\tanh(\frac{-a}{4}x), \qquad 0 \leq x \leq 1$$
$$u(0,t) = \frac{1}{2} + \frac{1}{2}\tanh[\frac{-a}{4}(-(\frac{a}{2} + \frac{2b}{a})t)],$$
$$u(1,t) = \frac{1}{2} + \frac{1}{2}\tanh[\frac{-a}{4}(1 - (\frac{a}{2} + \frac{2b}{a})t)], \quad t \geq 0.$$

The exact solution is

$$u(x,t) = \frac{1}{2} + \frac{1}{2}\tanh[\frac{-a}{4}(x - (\frac{a}{2} + \frac{2b}{a})t)] \quad t \geq 0.$$

In our computation,the computed solution are compare with exact solution.The maximum absolute error are reported in table 3-5.In Table 3,we take $a = b = 0.001$, $k = 0.00001$ *and* $h = 0.02$*.The results are computed for different time levels.In table 3,results have been compared with the results in references[22].The result show,our numerical results are more accurate in comparison to those given by Ismail et al.That result has been calculated by 5 terms in Adomian methods.In table 4,we take* $a = 0.001$ *and* $k = 0.0001$*,the results are computed for different step size and different* b*.The maximum absolute error for time* $t = 1$ *has been computed and tabulated in table 4.In table 5,we take* $h = 0.05$ *and* $k = 0.00001$*.The result are computed for different* a *and* b*.The maximum absolute error for two time level* $t = 0.5$ *and* $t = 1$ *have been computed and tabulated in table5. We show the graphs between exact and numerical solutions at* $t = 1$ *and* $a = 0.001$*,with different values of* h *and* b *in figures* $4 - 6$.

Table 3. Maximum absolute error for example 2

x	t	*[22]*	*present method*
0.1	0.005	9.68763(−6)	1.20184(−10)
	0.001	1.93753(−6)	3.04431(−11)
	0.01	1.93752(−5)	2.21709(−10)
0.5	0.005	9.68691(−6)	1.41649(−10)
	0.001	1.93738(−6)	3.05515(−11)
	0.01	1.93738(−5)	2.80496(−10)
0.9	0.005	9.68619(−6)	1.17931(−10)
	0.001	1.93724(−6)	3.02686(−11)
	0.01	1.93724(−5)	1.9840(−10)

Table 4. Maximum absolute error for example 2

x	$b = 0.001$		
	$h = 0.1$	$h = 0.05$	$h = 0.02$
0.1	3.08469(−8)	7.78978(−9)	1.24965(−9)
0.2	5.47325(−8)	1.30415(−8)	2.22143(−9)
0.3	7.16608(−8)	1.81554(−8)	2.91536(−9)
0.4	8.16355(−8)	2.07319(−8)	3.33144(−9)
0.5	8.46606(−8)	2.15716(−8)	3.46967(−9)
0.6	8.07397(−8)	2.06747(−8)	3.33005(−9)
0.7	6.98764(−8)	1.80417(−8)	2.91259(−9)
0.8	5.20745(−8)	1.367281(−8)	2.21728(−9)
0.9	2.73372(−8)	7.56845(−9)	1.2441(−9)
x	$b = 0.01$		
	$h = 0.1$	$h = 0.05$	$h = 0.02$
0.1	3.09723(−8)	7.82111(−9)	1.25449(−9)
0.2	5.49535(−8)	1.38968(−8)	2.22997(−9)
0.3	7.19487(−8)	1.82276(−8)	2.9265(−9)
0.4	8.19626(−8)	2.08142(−8)	3.34412(−9)
0.5	8.49993(−8)	2.16571(−8)	3.48284(−9)
0.6	8.10628(−8)	2.07567(−8)	3.34269(−9)
0.7	7.01566(−8)	1.81134(−8)	2.92366(−9)
0.8	5.22838(−8)	1.37274(−8)	2.22573(−9)
0.9	2.7447(−8)	7.5988(−9)	1.24887(−9)
x	$b = 1$		
	$h = 0.1$	$h = 0.05$	$h = 0.02$
0.1	1.53032(−7)	1.80708(−7)	1.86514(−7)
0.2	2.71202(−7)	3.19366(−7)	3.29698(−7)
0.3	3.54373(−7)	4.17367(−7)	4.03961(−7)
0.4	4.04031(−7)	4.75743(−7)	4.913151(−7)
0.5	4.20779(−7)	4.95119(−7)	5.11369(−7)
0.6	4.04834(−7)	4.75713(−7)	4.91322(−7)
0.7	3.56028(−7)	4.17331(−7)	4.30974(−7)
0.8	2.73812(−7)	3.19371(−7)	3.29714(−7)
0.9	1.57255(−7)	1.80824(−7)	1.86529(−7)

Table 5. Maximum absolute error for example 2

x	$a=b=0.001$		$a=b=0.0001$	
	$t=0.5$	$t=1$	$t=0.5$	$t=1$
0.10	7.76059(−8)	7.8118(−8)	7.73848(−10)	7.78809(−10)
0.20	1.3774(−7)	1.38711(−7)	1.37441(−9)	1.38383(−9)
0.30	1.80401(−7)	1.81813(−7)	1.802181(−9)	1.81513(−9)
0.40	2.05904(−7)	2.07467(−7)	2.05754(−9)	2.07274(−9)
0.50	2.14074(−7)	2.15714(−7)	2.14076(−9)	2.15672(−9)
0.60	2.05038(−7)	2.06597(−7)	2.05189(−9)	2.06705(−9)
0.70	1.78832(−7)	1.80157(−7)	1.79091(−9)	1.8038(−9)
0.80	1.35472(−7)	1.36436(−7)	1.35763(−9)	1.36699(−9)
0.90	7.4967(−8)	7.69406(−8)	7.51754(−10)	7.56673(−9)

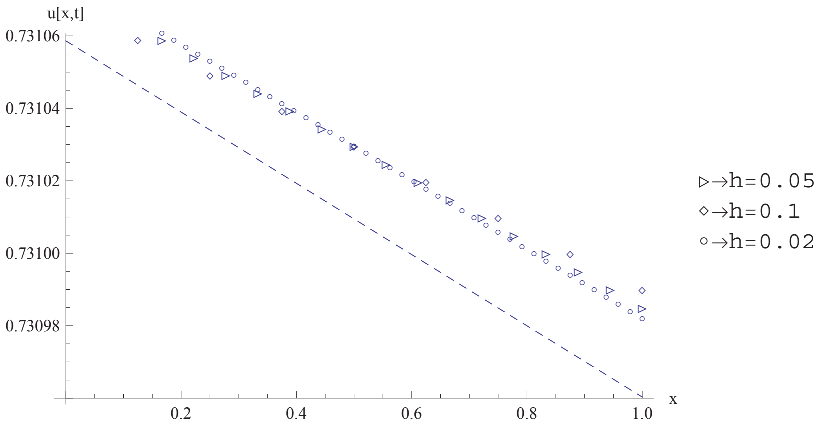


Figure 4:Approximate and exact solution for example 2 at $b=1$,with different h

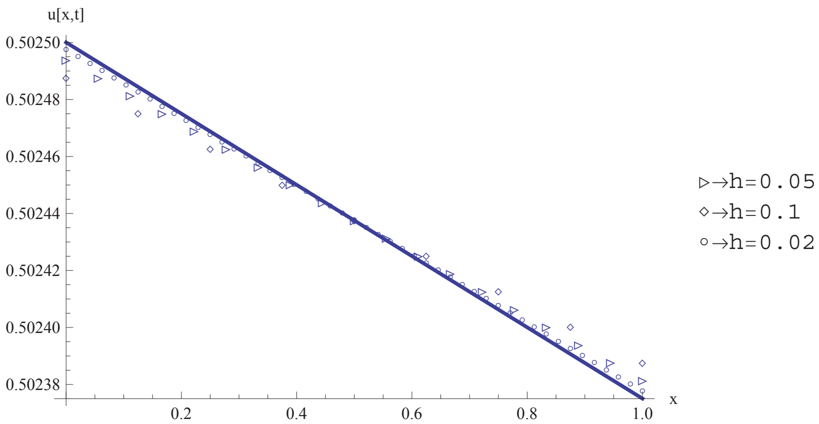


Figure 5:Approximate and exact solution for example 2 at $b=0.01$,with different h

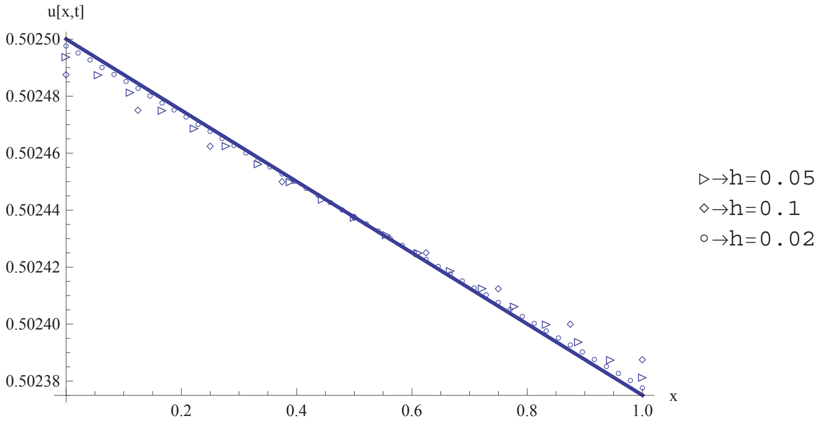


Figure 6:Approximate and exact solution for example 2 at $b = 0.001$,with different h

6. Conclusion

The basic goal of this work has been employed the non-polynomial tension spline as a reasonable basis for studying the approximate solutions for Burgers equations with nonlinear damping term.Finite difference approximation for time and tension spline for spatial are used.Presented scheme are of order $O(h^2 + k^2h^2 + h^4)$ and under appropriate condition the method convergence.The performance and accuracy of the method have been examined by applying in 2 examples.

References

[1] Burgers,J.M.*A mathematical model illustrating the theory of turbulence*,Adv. Appl. Mech**1**,163-170,1948.

[2] Barabasi,A.L. and Stanley,H.E.*Fractal Concepts in Surface Growth*(Cambridge University Press,Cambridge,1995)

[3] Bouchaud,J.P. and Mezard,M. and Parisi, G.*Scaling and intermittency in Burgers turbulence*, Phys. Rev.E**52**,3656-3674,1995.

[4] Biringen,S. and Satti,A.*Comaprison of several finite-difference methods*, J. Aircr**27**,90-92,1990.

[5] Bar-Yoseph,P. and Moses,E. and Zrahia,U. and Yarin,A.L.*Space-time spectral element methods for one dimensional non linear advection diffusion problems*,J.Comput.Phys**119**,62-75,1995.

[6] Bahadir,A.R. and Sagalam,M.*A mixed finite difference and boundary element approach to one-dimensional Burgers equation*, Appl.Math.Comput**160**,663-673,2005.

[7] Chowdhury,D and Santen,L. and Schadschneider,A.*Statistical physics of vehicular traffic and some related systems*, Phys. Rep**329**,199-329,2000.

[8] Cole,J.D.*On a quasi-linear paribolic equation occurring in aerodynamics*, Quart. Appl.Math**9**,225-236,1951.

[9] Caldwell,J. and Wanless,P. and Cook,E.*A finite element approach to Burgers equation*, Appl. Math. Model**5**,239-25,1981.

[10] Chen,R. and Wu,Z.*Applying multi-quadric quasi-interpolation to solve Burgers equation*, Appl. Math. Comput**172**,472-484,2006.

[11] Chawla,M.M.*A Sixth-order Tridiagonal Finite Difference Method for General Non-linear Two-point Boundary Value Problems*,J. Inst. Maths Applies**24**,35-42,1979.

[12] Darvishi,M.T. and Javidi,M.*A numerical solution of Burgers equation by pseudospectral method and Darvishis preconditioning*, Appl. Math. Comput **173**,421-429,2006.

[13] Dag,I. and Irk,D. and Saka,B.*A numerical solution of the Burgers equation using cubic B-splines*, Appl. Math. Comput**163**,199-211,2005.

[14] Evans,L.C.*Partial differential equations*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI**19**,363-367,1998.

[15] Frisch,U.*Turbulence,The Legacy of A.N. Kolmogorov*(Cambridge University Press, Cambridge,1995)

[16] Gurbatov,S.N. and Saichev A.I.*Probability distribution and spectra of potential hydrodynamic turbulence*,Radiophys.Quant.Electr**27**,303-313,1984.

[17] Garba,A.*A wavelet collocation method for numerical solution of Burgers equation*,(Tech. Rep.International Center for Theorytical Physics,1996)

[18] Goyal,K. and Mehra,M.*A fast adaptive diffusion wavelet method for Burgers equation*,Computers and Mathematics with Applications**68**,568-577,2014.

[19] Gopal,V. and Mohanty,R.K and Saha,L.M.*A new high accuracy non-polynomial tension spline method for the solution of one dimensional wave equation in polar coordinates*,Journal of the Egyptian Mathematical Society**22**,280-285,2014.

[20] Hopf,E.*The partial differential equation* $u_t + uu_x = u_{xx}$,Commun. Pure Appl. Math**3**,201-30,1950.

[21] Inan,B.and Bahadir,A.R*Numerical solution of the one-dimensional Burgers equation: Implicit and fully implicit exponential finite difference methods*, Indian Academy of Sciences**81**,547-556,2013.

[22] Ismail,H.N.A.and Raslan,K.and Abd Rabboh,A.*Adomian decomposition method for Burgers-Huxley an Burgers-Fisher equations*,Applied Mathematics and Computation**159**, 291-301,2004.

[23] Kardar,M.,Parisi,G.,Zhang,Y.C.*Dynamical scaling of growing interfaces*,Phys. Rev. Lett**56**,889-892,1986.

[24] Kardar M. and Zhang,Y.C.*Scaling of directed polymers in random media*,Phys. Rev. Lett**58**,2087-2090,1987.

[25] Kutluay,S.and Bahadir,A.R. and Ozdes,A.*Numerical solution of one-dimensional Burgers equation: explicit and exact-explicit finite difference methods*,J.Comput.Appl.Math**103**,251-261,1999.

[26] Kutluay,S. and Esen,A. and Dag,I.*Numerical solutions of the Burgers equation by the least-squares quadratic B-spline finite element method*,J.Comput.Appl.Math**167**,21-33,2004.

[27] LeVeque,R.J.*Numerical Methods for Conservation Laws*, Lectures in Mathematics ETH Zurich(Birkh auser Verlag,Basel,second ed,1992).

[28] Lighthill,M.J. and Whitham,G.B.*On kinematic waves.II.A theory of traffic flow on long crowded roads*, Proc. Roy. Soc.London.Ser.A**229**,317-345,1955.

[29] Logan,J.D.*An introduction to nonlinear partial differential equations*, Pure and Applied Mathematics(John Wiley and Sons Inc,New York,1994).

[30] Lighthill,M.J.*Viscosity effects in sound waves of finite amplitude*.Surveys in Mechanics,edited by Batchelor,G.K. and Davies,R.M.(Cambridge Univ.Press,London,1956).

[31] Lesieur,M.*Turbulence in Fluids, Fluid Mechanics and Its Applications*,Vol.**40**(Kluwer,1997)

[32] Mansell,G. and Merryfield,W. and Shizgal,B. and Weinert,U.*A comparison of differential quadrature methods for solution of partial differential equations*,Comput.Methods Appl. Mech. Engrg**104**,295-316,1993.

[33] Mohanty,R.K and Gopal,V. *A fourth-order finite difference method based on spline in tension approximation for the solution of one-space dimensional second-order quasi-linear hyperbolic equations*Advances in Difference Equations**70**,1-20,2013.

[34] Orszag,S.A.*Lectures on the Statistical Theory of Turbulence, in Fluid Dynamics*(New York,1973)

[35] Ozis,T. and Aksan,E.N. and Ozdes,A.*A finite element approach for solution of Burgers equation*,Appl.Math.Comput**139**,417-428,2003.

[36] Rashidinia,J and Mohammadi,R.*Tension spline approach for the numerical solution of nonlinear Klein-Gordon equation*,Computer Physics Communications **181**,78-91,2010.

[37] Rose,H.A. and Sulem,P.L.*Fully developed turbulence and statistical mechanics*,J. Phys. France**39**,441-484,1978.

[38] Rao,C.S. and Sachdev,P.L. and Ramaswamy,M.*Self-similar solutions of a generalized Burgers equation with nonlinear damping, Nonlinear Analysis:Real World Applications***4**,723-741,2003.

[39] Richards,P.I.*Shock waves on the highway*, Operations Res**4**,42-51,1956.

[40] Stoker,J.J.*Water waves The mathematical theory with applications*,Wiley Classics Library(John Wiley and Sons Inc.,New York,1992).

[41] Saka,B.and Dag,I.*Quartic B-spline collocation methods to the numerical solutions of the Burgers equation*, Chaos Solitons Fract**32**,1125-1137,2007.

[42] Varoglu,E. and Finn,W.*Space-time finite elements incorporating characterstics for burgers equation*, Int. J. Numer. Methods Eng**16**,171-184,1980.

[43] Whitham,G.B.*Linear and nonlinear waves, Pure and Applied Mathematics*(John Wiley and Sons Inc.,New York,1999).

[44] Wei,G.W.,Zhang,D.S.,Kouri,D.J.,Hoffman,D.K.*Distributed approximating functional approach to Burgers equation in one and two space dimensions*, Comput. Phys. Commun**111**,93-109,1998.

[45] Wan,D.C. and Wei,G.W.*The study of quasi wavelets based numerical method applied to Burgers equation*,Appl.Math.Mech**21**,1099-1110,2000.

[46] Zeldovich,Y.B.*Gravitational instability: an approximate theory for large density perturbations*, Astron. Astrophys**5**,84-89,1970.

[47] Zhang,D.S. and Wei,G.W. and Kouri,D.J.*Burgers equation with high Reynolds number*, Phys. Fluids**9**,1853-1855,1997.

[48] Zhu,C.G. and Wang,R.H.*Numerical solution of Burgers equation by cubic B-spline quasi-interpolation*, Appl. Math. Comput**208**,260-272,2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1125 – 1134

A classification of biharmonic hypersurfaces in the Minkowski spaces of arbitrary dimension

Nurettin Cenk Turgay *

Abstract

In this paper we study hypersurfaces with the mean curvature function H satisfying $\langle \nabla H, \nabla H \rangle = 0$ in a Minkowski space of arbitrary dimension. First, we obtain some conditions satisfied by connection forms of biconservative hypersurfaces with the mean curvature function whose gradient is light-like. Then, we use these results to get a classification of biharmonic hypersurfaces. In particular, we prove that if a hypersurface is biharmonic, then it must have at least 6 distinct principal curvatures under the hypothesis of having mean curvature function satisfying the condition above.

Keywords: biharmonic submanifolds, Lorentzian hypersurfaces, biconservative hypersurfaces, finite type submanifolds.

2000 AMS Classification: 53C40 (Primary), 53C42, 53C50.

Received : 19.01.2015 *Accepted* : 13.07.2015 *Doi* : 10.15672/HJMS.20164513115

1. Introduction

After Bang-Yen Chen conjectured that every biharmonic submanifold of a Euclidean space is minimal, biharmonic and biconservative submanifolds in semi-Euclidean spaces have been studied by many geometers (cf. [4, 5, 7, 8]). In particular, many results on biharmonic submanifolds in the Minkowski 4-space $\mathbb{E}^4_1$ and the semi-Euclidean space $\mathbb{E}^4_2$ have appeared since the middle of 1990s, [1, 2, 6, 9, 18].

On the other hand, several geometrical properties of biconservative submanifolds in Euclidean spaces have been obtained and some classification results of biconservative hypersurfaces have been given so far, [3, 12, 15, 17]. For example in [12], Hasanis and Vlachos obtained the complete classification of biconservative hypersurfaces in $\mathbb{E}^3$ and $\mathbb{E}^4$. Furthermore, Yu Fu have recently proved that the only biconservative surfaces in $\mathbb{E}^3_1$ are surfaces of revolution and null scrolls, [10]. Most recently, the complete classification of biconservative surfaces in 4-dimensional Lorentzian space forms is obtained in [11]

*Istanbul Technical University, Faculty of Science and Letters, Department of Mathematics, 34469 Maslak, Istanbul, Turkey, Email: turgayn@itu.edu.tr

Let M be a hypersurface in $\mathbb{E}_s^{n+1}$, $s = 0, 1$ with the shape operator S, mean curvature H and $x : M \to \mathbb{E}^m$ an isometric immersion. M is said to be biharmonic if the equation $\Delta^2 x = 0$ is satisfied or, equivalently, the system of differential equations

(BC) $\quad S(\nabla H) + \varepsilon \dfrac{nH}{2} \nabla H = 0,$

(BH1) $\quad \Delta H + H \mathrm{tr} S^2 = 0$

is satisfied, where N is the unit normal vector field (see [6, 13]) and $\varepsilon = \langle N, N \rangle$.

On the other hand, a hypersurface satisfying (BC) is said to be a biconservative hypersurface. From (BC), one can see that if a hypersurface M with non-constant mean curvature is biconservative, then ∇H is an eigenvector of its shape operator. Note that along with the increase of index, the difference between Euclidean space and Minkowski space is the appearance of light-like vectors. Thus, the hypersurfaces with light-like ∇H has no counterparts in Euclidean spaces and they are worth to be studied separately in terms of being biconservative or biharmonic.

1.1. Remark. For ease of elaboration, we want to abbreviate a hypersurface with mean curvature whose gradient is light-like to a MCGL-hypersurface.

In this work we study MCGL-hypersurfaces in the Minkowski space of arbitrary dimension. In Sect. 2, after we describe our notations, we give a summary of the basic facts and formulas that we will use. In Sect. 3, we focus on biconservative MCGL-hypersurfaces and obtain some necessary conditions. In Sect. 4, we prove the non-existence of biharmonic MCGL-hypersurfaces under some conditions.

2. Prelimineries

Let $\mathbb{E}_s^m$ denote the pseudo-Euclidean m-space with the canonical pseudo-Euclidean metric tensor g of index s given by

$$g = -\sum_{i=1}^{s} dx_i^2 + \sum_{j=s+1}^{m} dx_j^2,$$

where $(x_1, x_2, \ldots, x_m)$ is a rectangular coordinate system in $\mathbb{E}_s^m$. A non-zero vector $v \in \mathbb{E}_s^m$ is called space-like, time-like or light-like if $\langle v, v \rangle > 0$, $\langle v, v \rangle < 0$ or $\langle v, v \rangle = 0$, respectively.

Consider an oriented hypersurface M of the Minkowski space $\mathbb{E}_1^{n+1}$. We denote the Levi-Civita connections of $\mathbb{E}_1^{n+1}$ and M by $\widetilde{\nabla}$ and ∇, respectively. Then, the Gauss and Weingarten formulas are given, respectively, by

$$\widetilde{\nabla}_X Y = \nabla_X Y + h(X, Y)N, \tag{2.1}$$

$$\widetilde{\nabla}_X N = -S(X) \tag{2.2}$$

for all tangent vectors fields X, Y, where h, $\nabla^\perp$ and S are the second fundamental form, the normal connection and the shape operator of M, respectively, and N is the unit normal vector field associated with the orientation of M.

The Gauss and Codazzi equations are given, respectively, by

$$R(X, Y, Z, W) = \langle h(Y, Z), h(X, W) \rangle - \langle h(X, Z), h(Y, W) \rangle, \tag{2.3}$$

$$(\bar{\nabla}_X h)(Y, Z) = (\bar{\nabla}_Y h)(X, Z), \tag{2.4}$$

where R is the curvature tensor associated with the connection ∇ and $\bar{\nabla} h$ is defined by

$$(\bar{\nabla}_X h)(Y, Z) = \nabla_X^\perp h(Y, Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z).$$

M is said to be Lorentzian if its tangent space T_mM at every point $m \in M$ has two linearly independent null vectors. In this case, there exists a pseudo-orthonormal frame field $\{e_1, e_2, \ldots, e_n\}$ of the tangent bundle of M satisfying

$$\langle e_A, e_B \rangle = 1 - \delta_{AB}, \quad \langle e_A, e_a \rangle = 0, \quad \langle e_a, e_b \rangle = \delta_{ab}$$

for all $A, B = 1, 2$, $a, b = 3, 4, \ldots, n$. Then, the Levi-Civita connection ∇ of M becomes

$$\nabla_{e_i} e_1 = \phi_i e_1 + \sum_{b=3}^{n} \omega_{1b}(e_i) e_b, \tag{2.5a}$$

$$\nabla_{e_i} e_2 = -\phi_i e_2 + \sum_{b=3}^{n} \omega_{2b}(e_i) e_b, \tag{2.5b}$$

$$\nabla_{e_i} e_a = \omega_{2a}(e_i) e_1 + \omega_{1a}(e_i) e_2 + \sum_{b=3}^{n} \omega_{ab}(e_i) e_b, \tag{2.5c}$$

where $\phi_i = \phi(e_i) = \langle \nabla_{e_i} e_2, e_1 \rangle$ and $\omega_{jk}(e_i) = \langle \nabla_{e_i} e_j, e_k \rangle$, i.e., $\phi = -\omega_{12}$.

On the other hand, the shape operator S of an oriented Lorentzian hypersurface in $\mathbb{E}^{n+1}_1$ can be non-diagonalizable. If S is non-diagonalizable, then its characteristic polynomial may also have complex roots. However, in this case all eigenvectors of S are space-like.

Now, assume that M has non-diagonalizable shape operator S and consider the case that all of the eigenvalues of S are real at any point of M. In this case, locally, we may assume that the multiplicities of eigenvalues are constant at every point of M. Therefore, [14, Lemma 2.3 and Lemma 2.5] imply that there exists an appropriate pseudo-orthonormal frame field $\{e_1, e_2, \ldots, e_n\}$ of smooth vector fields such that the matric representation of S is in one of the following two forms.

$$\begin{aligned}
&\text{Case I. } S = \begin{pmatrix} k_1 & 1 & 0 & 0 & \ldots & 0 \\ 0 & k_1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & k_3 & 0 & \ldots & 0 \\ 0 & 0 & 0 & k_4 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & k_n \end{pmatrix}, \\
&\text{Case II. } S = \begin{pmatrix} k_1 & 0 & 1 & 0 & \ldots & 0 \\ 0 & k_1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & k_1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & k_4 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & k_n \end{pmatrix}
\end{aligned} \tag{2.6}$$

for some **smooth** functions $k_1, k_3, k_4, \ldots, k_n$, where the eigenvector e_1 of S is light-like, (see also [13, 16]). With the abuse of terminology, we call these vector fields $e_1, e_2, \ldots, e_n$ as principal directions and the functions $k_1, k_3, k_4, \ldots, k_n$ as principal curvatures. Moreover, we put

$$s_1 = 2k_1 + k_3 + \cdots + k_n = nH,$$

where H is the mean curvature of M.

3. Biconservative MCGL-hypersurfaces

In this section we focus on biconservative MCGL-hypersurfaces in the Minkowski space $\mathbb{E}_1^{n+1}$. As we described in the previous section, the shape operator S of a MCGL-hypersurfaces in the Minkowski space $\mathbb{E}_1^{n+1}$ is one of two forms given in (2.6). We study these two cases separately.

3.1. Case I. Consider a hypersurface M in the Minkowski space $\mathbb{E}_1^{n+1}$ with the shape operator S given by case I of (2.6). Then, we have

$$\begin{aligned}
&h(e_1,e_2)=-k_1, \quad h(e_2,e_2)=-1,\\
(3.1)\qquad &h(e_A,e_B)=\delta_{AB}k_A,\\
&h(e_1,e_1)=h(e_1,e_A)=h(e_2,e_A)=0, \qquad A,B=3,4,\dots,n.
\end{aligned}$$

Now, assume that M is a biconservative MCGL-hypersurface, i.e., ∇s_1 is light-like and (BC) is satisfied. Then, e_1 is proportional to ∇s_1 and we have

$$k_1=-\frac{s_1}{2}, \quad k_3+k_4+\cdots+k_n=2s_1, \tag{3.2a}$$

$$e_1(k_1)=e_3(k_1)=e_4(k_1)=\cdots=e_n(k_1)=0, \quad e_2(k_1)\neq 0. \tag{3.2b}$$

Let the distinct principal curvatures of M be $K_1,K_2,\dots,K_p$ with the multiplicities $\nu_1,\nu_2,\dots,\nu_p$, respectively, i.e., the characteristic polynomial of S is

$$\rho_S(t)=(t-K_1)^{\nu_1}(t-K_2)^{\nu_2}\dots(t-K_p)^{\nu_p} \tag{3.3}$$

with $K_1=k_1$ and $\nu_1\geq 2$. We also suppose that the functions $K_\alpha-K_\beta$ does not vanish on M, for all $\alpha\neq\beta\in\{1,2,\dots,p\}$. Then, (3.2a) becomes

$$K_1=-\frac{s_1}{2}, \quad \nu_2K_2+\nu_3K_3+\cdots+\nu_pK_p=(-2-\nu_1)K_1. \tag{3.4}$$

On the other hand, from the Codazzi equation (2.4) for $X=e_1$, $Y=Z=e_A$ we get

$$\psi_\alpha=\omega_{1A}(e_A)=\frac{e_1(K_A)}{K_1-K_A} \quad \text{if } k_A=K_\alpha,\ \alpha=2,3,\dots,p. \tag{3.5}$$

By rearranging the indices if necessary, we may assume that $\psi_2,\psi_3,\dots,\psi_r\neq 0$ and $\psi_{r+1}=\psi_{r+2}=\cdots=\psi_p=0$ for some $r\leq p$. Thus, from (3.5) we have

$$e_1(K_A)=0 \quad \text{if } k_A=K_\alpha,\ \alpha>r. \tag{3.6}$$

From Codazzi equation (2.4) for $X=e_1$, $Y=e_A$, $Z=e_B$ and $X=e_A$, $Y=e_B$, $Z=e_1$ we obtain

$$\omega_{1A}(e_B)(k_1-k_A)=\omega_{1B}(e_A)(k_1-k_B)=\omega_{AB}(e_1)(k_A-k_B), \quad A,B=2,3,\dots,n. \tag{3.7}$$

Moreover, from the equation $[e_A,e_B](k_1)=0$ we have

$$\omega_{1A}(e_B)=\omega_{1B}(e_A).$$

By combining the above equation with (3.7) one may obtain

$$\omega_{1B}(e_A)=0 \quad \text{if } k_A,k_B\neq K_1. \tag{3.8}$$

On the other hand, from the Codazzi equation $X=e_1$, $Y=e_1$, $Z=e_j$ and $X=e_2$, $Y=e_1$, $Z=e_j$ we have

$$\omega_{1j}(e_1)=0, \quad j=3,4,\dots,n. \tag{3.9}$$

In addition, by combining the Codazzi equation (2.4) for $X=e_A$, $Y=e_1$, $Z=e_a$ and $[e_a,e_A]\,(k_1)=0$, we obtain

$$\omega_{aA}(e_1)=\omega_{1A}(e_a)=\omega_{1a}(e_A)=0 \tag{3.10}$$

for all $a, A = 3, 4, \ldots, n$ such that $k_a = K_1 \neq k_A$. By summing up (3.8)-(3.10) we obtain

$$\begin{aligned} \nabla_{e_1} e_1 = \phi_1 e_1, \quad \nabla_{e_A} e_1 &= \phi_A e_1 + \omega_{1A}(e_A) e_A, \\ \omega_{1A}(e_x) &= 0, \quad x \neq 2, x \neq A \end{aligned} \tag{3.11}$$

for all $A = 3, 4, \ldots, n$ such that $K_1 \neq k_A$.

Hence, by combaining (3.11) and the Gauss equation $R(e_A, e_1, e_1, e_A) = 0$ we obtain

$$e_1(\omega_{1A}(e_A)) = \omega_{1A}(e_A)(\phi_1 - \omega_{1A}(e_A)) \quad \text{if } k_A \neq K_1$$

from which we get

$$e_1(\psi_\alpha) = \psi_\alpha(\phi_1 - \psi_\alpha), \quad \psi_\alpha = 2, 3, \ldots, r. \tag{3.12}$$

Next, we obtain the following lemma which we will use later.

3.1. Lemma. Let M be a biconservative MCGL-hypersurface in the Minkowski space $\mathbb{E}^{n+1}_1$ with the shape operator given by (3.1). Then the functions $\psi_3, \psi_4, \ldots, \psi_r$ defined above satisfy

$$W(\psi_2, \psi_3, \ldots, \psi_r) \begin{pmatrix} \nu_2(K_1 - K_2) \\ \nu_3(K_1 - K_3) \\ \vdots \\ \nu_r(K_1 - K_r) \end{pmatrix} = 0, \tag{3.13a}$$

where $W(\psi_2, \psi_3, \ldots, \psi_r)$ is an $r \times r$ matrix given by

$$W(\psi_2, \psi_3, \ldots, \psi_r) = \begin{pmatrix} \psi_2 & \psi_3 & \ldots & \psi_r \\ \psi_2^2 & \psi_3^2 & \ldots & \psi_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_2^r & \psi_3^r & \ldots & \psi_r^r \end{pmatrix}. \tag{3.13b}$$

Proof. By applying e_1 to the second equation in (3.4) and using (3.2b), we obtain

$$\nu_2 e_1(K_2) + \nu_3 e_1(K_3) + \cdots + \nu_p e_1(K_p) = 0. \tag{3.14}$$

Now, by induction we would like to show

$$\sum_{\alpha=2}^{r} (\psi_\alpha)^t \nu_\alpha (K_1 - K_\alpha) = 0, \quad t = 1, 2, \ldots. \tag{3.15}$$

Note that by combining (3.5) and (3.14) one can obtain (3.15) for $t = 1$. Suppose that (3.15) is satisfied for $t = l - 1$, i.e.,

$$\sum_{\alpha=2}^{r} (\psi_\alpha)^{l-1} \nu_\alpha (K_1 - K_\alpha) = 0, \quad n = 1, 2, \ldots. \tag{3.16}$$

By applying e_1 to this equation and using (3.2b), (3.5) and (3.12) we obtain

$$\sum_{\alpha=2}^{r} (l-1)(\psi_\alpha)^{l-1} \nu_\alpha (\phi_1 - \psi_\alpha)(K_1 - K_\alpha) = \sum_{\alpha=2}^{r} (\psi_\alpha)^l \nu_\alpha (K_1 - K_\alpha).$$

By combining this equation and (3.16) we obtain (3.15) for $t = l$. Thus, we have (3.15) for all t which implies (3.13). □

3.2. Case II. In this subsection, we consider the hypersurfaces with the shape operator given by case II of (2.6) in the Minkowski space $\mathbb{E}^{n+1}$. Now, assume that M is a biconservative MCGL-hypersurface. In this case, we have

$$\begin{aligned} &h(e_1,e_2)=-k_1, \quad h(e_1,e_1)=h(e_1,e_3)=h(e_2,e_2)=0,\\ &h(e_3,e_3)=k_1, \quad h(e_A,e_B)=\delta_{AB}k_A,\\ &h(e_1,e_1)=h(e_1,e_A)=h(e_2,e_A)=h(e_3,e_A)=0, \qquad A,B=4,5,\dots,n. \end{aligned} \tag{3.17}$$

Assume that the characteristic polynomial of S is as given by (3.3) with $K_1=k_1=-s_1/2$ and $\nu_1\geq 3$. Then, we have (3.4) and

$$e_1(K_1)=e_3(K_1)=e_4(K_1)=\cdots=e_n(K_1)=0, \quad e_2(K_1)\neq 0. \tag{3.18}$$

We again suppose that the functions $K_\alpha-K_\beta$ does not vanish on M.

Note that the Codazzi equation (2.4) for $X=e_1$, $Y=e_A$, $Z=e_A$ gives $e_1(k_A)=\omega_{1A}(e_A)(k_1-k_A)$ if $k_1\neq k_A$. Let $\psi_2,\psi_3,\dots,\psi_p$ be the functions defined by (3.5) such that $\psi_2,\psi_3,\dots,\psi_r\neq 0$ and $\psi_{r+1}=\psi_{r+2}=\cdots=\psi_p=0$ for some $r\leq p$.

(3.18) implies $[e_1,e_A](k_1)=0$. By computing the left-hand side of this equation we get $\omega_{1A}(e_1)=0, A=3,4,\dots,n$. In addition, the Codazzi equation (2.4) for $X=e_1$, $Y=e_2$, $Z=e_3$ gives $\phi_1=0$. Thus, we have $\nabla_{e_1}e_1=0$. Next, similar to previous subsection, we apply the Codazzi equation (2.4) for $X=e_i, Y=e_j, Z=e_k$ for each triplet (i, j, k) in the set $\{(1,2,a), (1,3,A), (3,A,1), (1,A,B), (A,B,1), (1,a,A)\}$ and combine equations obtained with $[e_A,e_B](k_1)=[e_A,e_a](k_1)=0$ to get $\nabla_{e_A}(e_1)\in\operatorname{span}\{e_1,e_A\}$ and $\omega_{1A}(e_x)=0,\ x\neq 2,A$, where $A,B,a=4,5,\dots,n$ with $A\neq B$, $k_A,k_B\neq K_1$, $k_a=K_1$. By combaining these equations with the Gauss equation $R(e_3,e_1,e_1,e_3)=0$ we obtain

$$e_1(\psi_\alpha)=-\psi_\alpha^2, \quad \alpha=1,2,\dots,r.$$

Therefore, similar to Lemma 3.1 we have

3.2. Lemma. Let M be a biconservative MCGL-hypersurface in the Minkowski space $\mathbb{E}^{n+1}_1$ with the shape operator given by (3.17). Then the functions $\psi_3,\psi_4,\dots,\psi_r$ defined above satisfy (3.13).

3.3. Biconservative hypersurfaces. In this subsection, we would like to obtain conditions satisfied by connection forms of biconservative MCGL-hypersurfaces (See [17, 10, 11] for implicit examples of biconservative hypersurfaces that have recently obtained).

Now we would like to obtain some necessary conditions for being biconservative of an MCGL-hypersurface by using Lemma 3.1 and Lemma 3.2.

3.3. Proposition. Let M be an MCGL-hypersurface in the Minkowski space $\mathbb{E}^{n+1}_1$ and $e_1,e_2,\dots,e_n$ its principal directions with corresponding principal curvatures $k_1,k_1,k_3,k_4,\dots,k_n$ such that e_1 is proportional to gradient of its mean curvature. If M is biconservative, then

(i) For any $3\leq i\leq n$ such that $k_i\neq k_1,\omega_{1i}(e_i)\neq 0$, there exists a $j\neq i$ such that $\omega_{1j}(e_j)=\omega_{1i}(e_i),\ k_j\neq k_i$.

(ii) Let $I_i=\{3\leq j\leq n|\omega_{1j}(e_j)=\omega_{1i}(e_i)\}$. Then,

$$\sum_{j\in I_i}(k_1-k_j)=0. \tag{3.19}$$

(iii) There exists a $j\in\{3,4,\dots,n\}$ such that $e_1(k_j)=\omega_{1l}(e_j)=0,\ k_1\neq k_l$.

Proof. Let $K_1,\dots,K_n$ and $\psi_2,\dots,\psi_r$ be the functions defined on the beginning of this section.

Assume that $\psi_2 \neq 0$ and $\psi_2 \neq \psi_j$, $2 < j \leq r$. Then, we have $\det W(\psi_2, \psi_3, \dots, \psi_r) = 0$ from (3.13) since the functions $K_1 - K_2$ is non-vanishing by the assumptions. Therefore, $\psi_3, \dots, \psi_r$ are not distinct and we may assume $\psi_{r-1} = \psi_r$. Thus (3.13) gives

$$W(\psi_2, \psi_3, \dots, \psi_{r-1}) \begin{pmatrix} \nu_2(K_1 - K_2) \\ \nu_3(K_1 - K_3) \\ \vdots \\ \nu_r(K_1 - K_r) + \nu_{r-1}(K_1 - K_{r-1}) \end{pmatrix} = 0.$$

Since $(K_1 - K_2)$ is non-vanishing, the above equation implies that $\psi_3, \dots, \psi_{r-1}$ are not distinct and we may assume either $\psi_{r-2} = \psi_{r-1}$ or $\psi_3 = \psi_4$. By repeating this procedure, one can get $\psi_3 = \dots = \psi_{r-1}$ and

$$\begin{aligned} \psi_2(K_1 - K_2) + \psi_3 \left(\sum_{\alpha=3}^{r} \nu_\alpha (K_1 - K_\alpha) \right) &= 0, \\ \psi_2^2(K_1 - K_2) + \psi_3^2 \left(\sum_{\alpha=3}^{r} \nu_\alpha (K_1 - K_\alpha) \right) &= 0 \end{aligned}$$

which gives $\psi_2 = \psi_3$ or $K_1 - K_2 = 0$ which yields a contradiction. Hence we have (i) of the proposition.

Let $l - 1$ of $\psi_2, \psi_3, \dots \psi_r$ be distinct and by rearranging indices if necessary, assume that they are $\psi_2, \psi_3, \dots \psi_l$. Note that we have $l - 1 \leq (r-1)/2$ because of (i) of the proposition. Moreover, we have $\det W(\psi_2, \psi_3, \dots, \psi_l) \neq 0$. Thus, (3.13) implies

$$W(\psi_2, \psi_3, \dots, \psi_l) \begin{pmatrix} \sum_{j \in I_2} \nu_j (K_1 - K_j) \\ \sum_{j \in I_3} \nu_j (K_1 - K_j) \\ \vdots \\ \sum_{j \in I_l} \nu_j (K_1 - K_j) \end{pmatrix} = 0$$

which gives (ii) of the proposition.

Now, assume that all of the functions $\omega_{1j}(e_j)$ are non-zero, i.e., $r = p$ and $\psi_2, \psi_3, \dots \psi_l$ are distinct. Note that we have $\bigcup_{j=2}^{l} I_j = \{2, 3, \dots, p\}$ and (ii) of the proposition implies $\sum_{j \in I_\alpha} \nu_j (K_1 - K_j) = 0$ or, equivalently,

$$\sum_{j \in I_\alpha} \nu_j K_j = \left(\sum_{j \in I_\alpha} \nu_j \right) K_1, \quad \alpha = 2, 3, \dots, l.$$

By summing these equations over α we get

$$\nu_2 K_2 + \nu_3 K_3 + \dots + \nu_p K_p = (\nu_2 + \nu_3 + \dots + \nu_p) K_1.$$

However, this equation and (3.4) give $K_1 \equiv 0$ on M which implies $\nabla s_1 = 0$. This is a contradiction because we have assummed that ∇s_1 is light-like. Hence, we have (iii) of the proposition. $\square$

4. Biharmonic MCGL-Hypersurfaces

In this section we study biharmonic MCGL-hypersurfaces with the shape operator given by (3.1) in the Minkowski space $\mathbb{E}_1^{n+1}$ and obtain some classification results.

Let M be a biharmonic MCGL- hypersurface with the shape operator given by (3.1). Then, we have (3.2a)-(3.13) obtained in the Sect. 3.1. In addition, from the Codazzi equation $X = e_2$, $Y = e_1$, $Z = e_2$ and $X = e_A$, $Y = e_2$, $Z = e_A$ we have

(4.1) $\qquad e_2(k_1) = 2\phi_1 = \omega_{1A}(e_A), \quad \text{if } k_A = K_1,\ A > 2.$

Moreover, since $e_1e_2(k_1) = [e_1, e_2](k_1)$, by using (3.2b) we get

(4.2) $\qquad e_1e_2(k_1) = -\phi_1 e_2(k_1).$

This equation and (4.1) imply

(4.3) $\qquad e_1(\phi_1) = -\phi_1^2.$

Now we would like to consider the biharmonic equation (BH1). By a direct calculation using (3.2b) and (4.2) we get

$$\left(e_1e_2 + e_2e_1 - \sum_{j=3}^{n} e_je_j - \nabla_{e_1}e_2 - \nabla_{e_2}e_1\right)(k_1) = 0$$

which gives

$$\begin{aligned}\Delta k_1 &= \sum_{j=3}^{n} \omega_{1j}(e_j)e_2(k_1) = \sum_{\alpha=1}^{r}\left(\sum_{k_A=K_\alpha} \omega_{1A}(e_A)e_2(k_1)\right)\\ &= (2\nu_1\phi_1 + \nu_2\psi_2 + \nu_3\psi_3 + \cdots + \nu_r\psi_r)\, e_2(k_1).\end{aligned}$$

By combaining the above equation and (4.1), we see that the biharmonic equation (BH1) becomes

(BH2) $\quad (4\nu_1\phi_1 + 2\nu_2\psi_2 + 2\nu_3\psi_3 + \cdots + 2\nu_r\psi_r)\phi_1 = -k_1(\nu_1K_1^2 + \nu_2K_2^2 + \cdots + \nu_pK_p^2).$

4.1. Theorem. *There exists no biharmonic MCGL-hypersurface with at most 5 distinct principal curvatures and the shape operator given by* (3.1) *in the Minkowski space* $\mathbb{E}_1^{n+1}$.

Proof. Let the distinct principal curvatures of M be K_1, K_2, K_3, K_4, K_5 with the multiplicities $\nu_1, \nu_2, \nu_3, \nu_4, \nu_5$, respectively, and consider the functions $\psi_2, \psi_3, \psi_4, \psi_5$ defined by (3.5). Now, toward contradiction we assume that M is a biharmonic MCGL-hypersurface, i.e., (BC) and (BH1) are satisfied.

Case I. $p < 4$. If the number of distinct principal curvatures is less then 4, the proof directly follows from Proposition 3.3.

Case II. $p = 4$. Next, we consider the case that M has exactly 4 distinct principal curvatures, i.e., $K_4 = K_5$. Then, because of (iii) of Proposition 3.3, we may assume $\psi_2 = 0$. Note that if $\psi_3 = 0$, then (i) of Proposition 3.3 implies $\psi_4 = 0$. In this subcase, we have $r = 1$ and (3.6) implies $e_1(K_\alpha) = 0,\ \alpha = 1, 2, 3, 4$. Thus (BH2) becomes

(4.4) $\qquad 4\nu_1\phi_1^2 = -k_1(\nu_1K_1^2 + \nu_2K_2^2 + \nu_3K_3^2 + \nu_4K_4^2).$

By applying e_1 to this equation and using (4.3) one can find $\nu_1\phi_1^3 = 0$. However, this equation and (4.4) implies $k_1 \equiv 0$. Thus, we have $\nabla s_1 = 0$ which contradicts with being light-like of ∇s_1. Hence, ψ_3 and ψ_4 are non-zero.

Therefore, (i) and (ii) of Proposition 3.3 imply

(4.5) $\qquad \psi_3 = \psi_4, \quad \nu_3(K_1 - K_3) + \nu_4(K_1 - K_4) = 0.$

Thus, (BH2) becomes

(4.6) $\qquad (a\phi_1 + b\psi_3)\phi_1 = -k_1(\nu_1K_1^2 + \nu_2K_2^2 + \nu_3K_3^2 + \nu_4K_4^2),$

where $a = 4\nu_1$ and $b = 2(\nu_3 + \nu_4)$ are some non-negative constants. Note that $\psi_2 = 0$ and (3.5) imply $e_1(K_2) = 0$.

Next, we apply e_1 to (4.6) and use (3.2b), (4.3), (3.12) to get

$$-(2a\phi_1^2 + b\psi_3^2)\phi_1 = -k_1 e_1(\nu_3 K_3^2 + \nu_4 K_4^2). \tag{4.7}$$

Then we use, (3.5) and (4.5) to compute the right-hand side of (4.7) and get

$$-(2a\phi_1^2 + b\psi_3^2)\phi_1 = -2k_1\psi_3(bK_1^2 - \nu_3 K_3^2 - \nu_4 K_4^2). \tag{4.8}$$

By applying e_1 to (4.8) again and using (3.2b), (4.3), (3.12) we get

$$\begin{aligned}(6a\phi_1^3 - b\phi_1\psi_3^2 + 2b\psi_3^3)\phi_1 = &-2k_1\psi_3(\phi_1 - \psi_3)(bK_1^2 - \nu_3 K_3^2 - \nu_4 K_4^2)\\ &+ 2k_1\psi_3 e_1(\nu_3 K_3^2 + \nu_4 K_4^2)\end{aligned} \tag{4.9}$$

By combining (4.7), (4.8) and (4.9) we get

$$\left(6a\phi_1^3 - b\phi_1\psi_3^2 + 2b\psi_3^3 + (\phi_1 - 3\psi_3)(2a\phi_1^2 + b\psi_3^2)\right)\phi_1 = 0. \tag{4.10}$$

Thus, we have $\psi_3 = c\phi_1$ for a constant c. However, in this case, from (4.3) and (3.12) we get $c = 2$, i.e., $\psi_3 = 2\phi_1$. However, this equation and (4.10) give $(a + 2b)\phi_1^4 = 0$ which is impossible to be satisfied because a, b are non-negative constants. Thus, the proof for this case is completed.

Case III. $p = 5$. Then, because of (iii) of Proposition 3.3, we may assume $\psi_2 = 0$. Note that, if $\psi_3 = 0$, then we have either $\psi_4 = \psi_5 \neq 0$ or $\psi_3 = \psi_4 = \psi_5 = 0$. However, these subcases and the other possible subcase $\psi_3 = \psi_4 = \psi_5$ are similar to case II. □

Acknowledgements

This work is supported by Scientific Research Agency of Istanbul Technical University (Project Number: ITU-BAP:37992). The author would like to express his sincere gratitude to the anonymous referee for his/her helpful comments that help to improve the quality of the manuscript.

References

[1] Arvanitoyeorgos, A., Defever, F., Kaimakamis, G. and Papantoniou, V., *Hypersurfaces of $\mathbb{E}_s^4$ with proper mean curvature vector* , J. Math. Soc. Japan, **59**, 797–809, 2007.

[2] Arvanitoyeorgos, A, Kaimakamis, G. and Magid, M., *Lorentz hypersurfaces in $\mathbb{E}_1^4$ satisfying $\Delta H = \alpha H$*. Illinois J. Math. **53**, 581–590, 2009.

[3] Caddeo, R., Montaldo, S., Oniciuc, C. and Piu, P.,*Surfaces in three-dimensional space forms with divergence-free stress-bienergy tensor* Annali di Matematica Pura ed Applicata **193**, 529–550, 2014.

[4] Chen, B. Y., *Some open problems and conjectures on submanifolds of finite type*. Soochow J. Math. **17**(2), 169-188, 1991.

[5] Chen, B.-Y., *A report on submanifolds of finite type*, Soochow J. Math., **22**, 117–337, 1996.

[6] Chen, B.-Y., *Total mean curvature and submanifolds of finite type, 2nd Edition*, World Scientific, Hackensack–NJ 2014.

[7] Chen, B.-Y. and Munteanu, M. I., *Biharmonic ideal hypersurfaces in Euclidean spaces*, Differential Geom. Appl. **31** (1), 1–16, 2013.

[8] Defever, F., *Hypersurfaces of $\mathbb{E}^4$ with harmonic mean curvature vector*, Math Nachr, **196**, 61–69, 1998.

[9] Defever, F., Kaimakamis, G. and Papantoniou, V., *Biharmonic hypersurfaces of the 4-dimensional semi-Euclidean space $\mathbb{E}_s^4$*, J. Math. Anal. Appl. **315**, 276–286, 2006.

[10] Fu, Y., *On bi-conservative surfaces in Minkowski 3-space*, J. Geom. Phys. **66**, 71–79, 2013.

[11] Fu, Y., *Explicit classification of biconservative surfaces in Lorentz 3-space forms*, Ann. Mat. Pura Appl. **194**, 805–822, 2015.

[12] Hasanis, T. and Vlachos, I., *Hypersurfaces in E^4 with harmonic mean curvature vector field*, Math Nachr, **172**, 145–169, 1995.

[13] Lucas, P. and Ramirez-Ospina H.F., *Hypersurfaces in the Lorentz-Minkowski space satisfying* $L_k\psi = A\psi + b$, Geom. Dedicata, **153**, 151–175, 2011

[14] Magid, M. A., *Lorentzian isoparametric hypersurfaces*, Pacific J. Math., **118** (1985), 165–197.

[15] Montaldo, S., Oniciuc, C. and Ratto, A., *Proper biconservative immersions into the Euclidean space*, arXiv:1312.3053 (to print) Ann. Mat. Pura Appl., DOI: 10.1007/s10231-014-0469-4.

[16] O'Neill, M. P., *Semi-Riemannian geometry with applications to relativity*. Academic Press, 1983.

[17] Turgay, N. C., *H-hypersurfaces with 3 distinct principal curvatures in the Euclidean spaces* (to print) Ann. Mat. Pura Appl., DOI: 10.1007/s10231-014-0445-z.

[18] Turgay, N. C., *Some classifications of biharmonic Lorentzian hypersurfaces in Minkowski 5-space* (to print) Mediterr. J. Math., DOI: 10.1007/s00009-014-0491-1.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1135 – 1142

The Zagreb coindices of a type of composite graph

Lihui Yang*, Xiaowei Ai† and Lin Zhang‡

Abstract

For a nontrivial graph G, its first and second Zagreb coindices are defined as the sum of degree sum of of nonadjacent vertex pairs and the sum of degree product of nonadjacent vertices pairs, respectively. Motivated by the work in [1], we study Zagreb coindices of a new kind of composite graph, namely, double graph. For any given nontrivial graph, explicit formulas are given for the Zagreb coindices of its double graph and k-iterated double graph, respectively.

Keywords: Degree; Zagreb indices; Zagreb coindices; double graph; k-iterated double graph.

2000 AMS Classification: 05C35; 05C12; 05C90.

Received : 30.12.2011 *Accepted* : 20.01.2013 *Doi* : 10.15672/HJMS.20164520079

1. Introduction

Let G be a simple graph with vertex set $V(G)$ and edge set $E(G)$. For a graph G, we let $d_G(v)$ be the degree of a vertex v in G, i.e., the number of the first neighbors of vertex v.

A topological index or graph invariant is a function defined on a (molecular) graph regardless of the labeling of its vertices. Till now, hundreds of different graph invariants have been employed in QSAR/QSPR studies, some of which have been proved to be successful (see [11]). Among those successful invariants, there are two topological indices,

*College of Mathematics and Computer Science, Hunan City University, Yiyang City, Hunan 413000, China
Email : llhhyang@yeah.net

†College of Mathematics and Information Science, Nanchang Hangkong University, Nanchang, Jiangxi 330063, P.R. China
Email : xiaoweiai2011@126.com

‡College of Mathematics and Information Science, Nanchang Hangkong University, Nanchang, Jiangxi 330063, P.R. China

relevant to our paper, called the *first Zagreb index* and the *second Zagreb index* (see [3, 4, 8, 10, 14]), defined as

$$M_1(G) = \sum_{u \in V(G)} [d_G(u)]^2 \text{ and } M_2(G) = \sum_{uv \in E(G)} d_G(u) d_G(v),$$

respectively.

Equivalently, we can rewrite the first Zagreb index as

$$M_1(G) = \sum_{uv \in E(G)} [d_G(u) + d_G(v)].$$

More recently, the authors [2] proposed two new Zagreb-type indices, namely, the *first Zagreb coindex* and *second Zagreb coindex* as

$$\overline{M}_1(G) = \sum_{uv \notin E(G)} [d_G(u) + d_G(v)] \text{ and } \overline{M}_2(G) = \sum_{uv \notin E(G)} d_G(u) d_G(v),$$

respectively.

It is well-known that one can construct many graphs from simpler graphs via various graph operations. Thus, it is important to understand how certain invariants of such composite graphs are related to the corresponding invariants of the original graphs.

More recently, Ashrafi et al. [1] investigated Zagreb coindices and presented explicit formulas for these new graph invariants under several graph operations, including union, join, Cartesian product, disjunction product, etc. Ashrafi et al. [2] determined the extremal values of Zagreb coindices over some special classes of graphs. Hua and Zhang [5] revealed some relations between Zagreb coindices and some other distance-based topological indices.

The *double graph* (see [9]) G^* of a given graph G is constructed by making two copies of (including the initial edge set of each) and adding edges u_1v_2 and u_2v_1 for every edge uv of G. For a nontrivial graph G, its *k-iterated double graph* G^{k*}, is defined as

$$G^{1*} = G^* \text{ and } G^{k*} = (G^{(k-1)*})^* \text{ for } k \geq 2.$$

In particular, it is generally assumed that $G^{0*} = G$ for the sake of consistence.

For results on double graphs, see [6, 7, 12, 13] and the references cited therein.

Motivated by the work in [1], we study Zagreb coindices of double graph. For any given nontrivial graph, explicit formulas are given for the Zagreb coindices of its double graph and k-iterated double graph, respectively.

2. Main results

We begin with some notation and terminology used in the proof of our results.

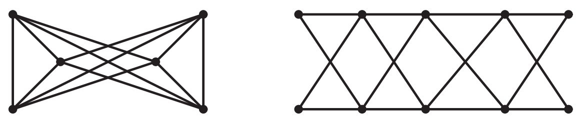


Fig. 1. The double graphs of C_3 and P_5.

For each vertex u in a nontrivial graph G, we call the corresponding vertices u_1 and u_2, in G^*, the *clone vertices* of u. As examples, we depicted the double graphs C_3^* (Fig. 1(a)) and P_5^* (Fig. 1(b)) of C_3 (Fig. 1(a)) and P_5, respectively.

For a given vertex v in G, if we let $\overline{D}_G^1(v) = \sum\limits_{uv \notin E(G)} [d_G(u) + d_G(v)]$ and $\overline{D}_G^2(v) = \sum\limits_{uv \notin E(G)} d_G(u)d_G(v)$, then we can rewrite the expressions of $\overline{M}_1(G)$ and $\overline{M}_2(G)$ as

$$\overline{M}_1(G) = \frac{1}{2} \sum_{v \in V(G)} \overline{D}_G^1(v) \tag{2.1}$$

and

$$\overline{M}_2(G) = \frac{1}{2} \sum_{v \in V(G)} \overline{D}_G^2(v), \tag{2.2}$$

respectively.

Similarly, if we denote $D_G^2(v) = \sum\limits_{uv \in E(G)} d_G(u)d_G(v)$, then the second Zagreb index of G can be rewritten as

$$M_2(G) = \frac{1}{2} \sum_{v \in V(G)} D_G^2(v). \tag{2.3}$$

In the following, we shall state and prove our main results of this paper.

2.1. Theorem. *Let G be a nontrivial graph of order n and size m. Then*
(i) $\overline{M}_1(G^*) = 8\overline{M}_1(G) + 8m$;
(ii) $\overline{M}_2(G^*) = 8\overline{M}_2(G) - 8M_2(G) + 16m^2$.

Proof. For the sake of convenience, we label all vertices in G as $\{v_1, \ldots, v_n\}$. Suppose that x_i and y_i are the corresponding clone vertices, in G^*, of v_i for each $i = 1, \ldots, n$. For any given vertex v_i in G and its clone vertices x_i and y_i, there exists $d_{G^*}(x_i) = d_{G^*}(y_i) = 2d_G(v_i)$ by the definition of double graph.

For $v_i, v_j \in V(G)$, if $v_iv_j \notin E(G)$, then $x_ix_j \notin E(G)$, $y_iy_j \notin E(G)$, $x_iy_j \notin E(G)$ and $y_ix_j \notin E(G)$.

So we need only to consider total contribution of the following three types of nonadjacent vertex pairs both to $\overline{M}_1(G^*)$ and to $\overline{M}_2(G^*)$.

- Type 1: The nonadjacent vertex pairs $\{x_i, x_j\}$ and $\{y_i, y_j\}$, where $v_iv_j \notin E(G)$.
- Type 2: The nonadjacent vertex pairs $\{x_i, y_i\}$ for each $i = 1, \ldots, n$.
- Type 3: The nonadjacent vertex pairs $\{x_i, y_j\}$ and $\{y_i, x_j\}$, where $v_iv_j \notin E(G)$.

The total contribution of nonadjacent vertex pairs of type 1 to $\overline{M}_1(G^*)$ and $\overline{M}_2(G^*)$ are, respectively, given by

$$\begin{aligned}
\sum_{y_iy_j \notin E(G^*)} [d_{G^*}(y_i) + d_{G^*}(y_j)] &= \sum_{x_ix_j \notin E(G^*)} [d_{G^*}(x_i) + d_{G^*}(x_j)] \\
&= \sum_{v_iv_j \notin E(G)} [2d_G(v_i) + 2d_G(v_j)] \\
&= 2 \sum_{v_iv_j \notin E(G)} [d_G(v_i) + d_G(v_j)] \\
&= 2\overline{M}_1(G)
\end{aligned}$$

and

$$\begin{aligned}\sum_{y_iy_j \notin E(G^*)} d_{G^*}(y_i)d_{G^*}(y_j) &= \sum_{x_ix_j \notin E(G^*)} d_{G^*}(x_i)d_{G^*}(x_j) \\ &= \sum_{v_iv_j \notin E(G)} [2d_G(v_i)] \cdot [2d_G(v_j)] \\ &= 4 \sum_{v_iv_j \notin E(G)} d_G(v_i)d_G(v_j) \\ &= 4\overline{M}_2(G).\end{aligned}$$

The total contribution of nonadjacent vertex pairs of type 2 to $\overline{M}_1(G^*)$ and $\overline{M}_2(G^*)$ are, respectively, given by

$$\begin{aligned}\sum_{i=1}^{n}[d_{G^*}(x_i) + d_{G^*}(y_i)] &= \sum_{i=1}^{n}[2d_G(v_i) + 2d_G(v_i)] \\ &= 4\sum_{i=1}^{n} d_G(v_i) \\ &= 8m\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{n} d_{G^*}(x_i)d_{G^*}(y_i) &= \sum_{i=1}^{n}[2d_G(v_i)][2d_G(v_i)] \\ &= 4\sum_{i=1}^{n}[d_G(v_i)]^2 \\ &= 4\left[\sum_{i=1}^{n} d_G(v_i)\right]^2 - 8\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} d_G(v_i)d_G(v_j) \\ &= 16m^2 - 8M_2(G) - 8\overline{M}_2(G).\end{aligned}$$

Now, we consider the total contribution of nonadjacent vertex pairs of type 3 to $\overline{M}_1(G^*)$ and $\overline{M}_2(G^*)$, respectively.

For each x_i, there exist $n - 1 - d_G(v_i)$ vertices in the set $\{y_1, \ldots, y_n\}$, among which every vertex together with x_i compose a nonadjacent vertex pairs of G^*. The total contribution of these $n-1-d_G(v_i)$ nonadjacent vertex pairs to $\overline{M}_1(G^*)$ is $\sum_{x_iy_j \notin E(G^*)} [d_{G^*}(x_i) + d_{G^*}(y_j)] = \sum_{v_iv_j \notin E(G)} [2d_G(v_i) + 2d_G(v_j)] = 2\overline{D}_G^1(v_i)$ and to $\overline{M}_2(G^*)$ is $\sum_{x_iy_j \notin E(G^*)} [d_{G^*}(x_i)d_{G^*}(y_j)] = 4 \sum_{v_iv_j \notin E(G)} d_G(v_i)d_G(v_j) = 4\overline{D}_G^2(v_i)$.

So we have

$$\begin{aligned}\sum_{i\neq j; x_iy_j \notin E(G^*)} [d_{G^*}(x_i) + d_{G^*}(y_j)] &= \sum_{i=1}^{n} 2\overline{D}_G^1(v_i) \\ &= 2\sum_{i=1}^{n} \overline{D}_G^1(v_i) \\ &= 4\overline{M}_1(G) \qquad \text{(by Eq. (1))}\end{aligned}$$

and

$$\begin{aligned}\sum_{i\neq j;x_iy_j\notin E(G^*)}[d_{G^*}(x_i)d_{G^*}(y_j)] &= \sum_{i=1}^{n}4\overline{D}_G^2(v_i)\\ &= 8\overline{M}_2(G) \qquad (\text{ by Eq. (2)}).\end{aligned}$$

Therefore,

$$\begin{aligned}\overline{M}_1(G^*) &= \sum_{x_ix_j\notin E(G^*)}[d_{G^*}(x_i)+d_{G^*}(x_j)] + \sum_{y_iy_j\notin E(G^*)}[d_{G^*}(y_i)+d_{G^*}(y_j)] + \\ &\quad \sum_{i=1}^{n}[d_{G^*}(x_i)+d_{G^*}(y_i)] + \sum_{i\neq j;x_iy_j\notin E(G^*)}[d_{G^*}(x_i)+d_{G^*}(y_j)]\\ &= 8\overline{M}_1(G)+8m\end{aligned}$$

and

$$\begin{aligned}\overline{M}_2(G^*) &= \sum_{x_ix_j\notin E(G^*)}d_{G^*}(x_i)d_{G^*}(x_j) + \sum_{y_iy_j\notin E(G^*)}d_{G^*}(y_i)d_{G^*}(y_j) + \\ &\quad \sum_{i=1}^{n}d_{G^*}(x_i)d_{G^*}(y_i) + \sum_{i\neq j;x_iy_j\notin E(G^*)}d_{G^*}(x_i)d_{G^*}(y_j)\\ &= 8\overline{M}_2(G)+(16m^2-8M_2(G)-8\overline{M}_2(G))+8\overline{M}_2(G)\\ &= 8\overline{M}_2(G)-8M_2(G)+16m^2.\end{aligned}$$

This completes the proof. □

Now, we give two examples as applications of Theorem 1.

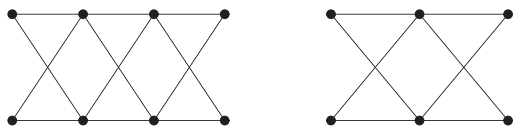


Fig. 2. The graphs G_{2n}.

2.2. Example. Consider Zagreb coindices of the graph G_{2n}, as shown in Fig. 2.

It can be easily seen that G_{2n} is just the double graph of the n-vertex path P_n. By an elementary calculation, we obtained $\overline{M}_1(P_n) = 2(n-2)^2$, $\overline{M}_2(P_n) = 2n^2-10n+13$ and $M_2(P_n) = 4n-8$. It then follows from Theorem 1 that $\overline{M}_1(G_{2n}) = 8\times 2(n-2)^2 + 8(n-1) = 16n^2-56n+56$ and $\overline{M}_2(G_{2n}) = 8(2n^2-10n+13)-8(4n-8)+16(n-1)^2 = 32n^2-144n+184$.

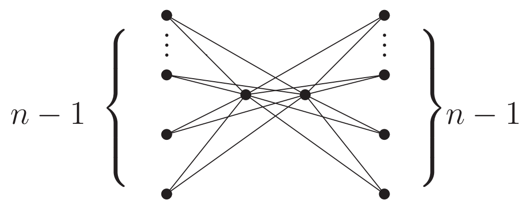


Fig. 3. The graphs H_{2n}.

2.3. Example. Consider Zagreb coindices of the graph H_{2n}, as shown in Fig. 3.

It can be easily seen that H_{2n} is just the double graph of the n-vertex star S_n. By an elementary calculation, we obtained $\overline{M}_1(S_n) = n^2 - 3n + 2$, $\overline{M}_2(S_n) = \frac{1}{2}(n^2 - 3n + 2)$ and $M_2(S_n) = (n-1)^2$. It then follows from Theorem 1 that $\overline{M}_1(H_{2n}) = 8(n^2 - 3n + 2) + 8(n-1) = 8n^2 - 16n + 8$ and $\overline{M}_2(H_{2n}) = 8[\frac{1}{2}(n^2 - 3n + 2)] - 8(n-1)^2 + 16(n-1)^2 = 12n^2 - 28n + 16$.

Now, we give formulas for Zagreb coindices of k-iterated double graphs.

2.4. Theorem. *Let G be a nontrivial graph of order n and size m, and let G^{k*} be its k-th iterated double graph. Then*
(i) $\overline{M}_1(G^{k*}) = 8^k \overline{M}_1(G) + 2^{2k+1}(2^k - 1)m$;
(ii) $\overline{M}_2(G^{k*}) = 8^k \overline{M}_1(G) - [8^k(2^k - 1)]M_2(G) + 2[8^k(2^k - 1)]m^2$.

Proof. For any nontrivial graph G with n vertices and m edges, the number of vertices in G^* is $2n$ and the number of edges in G^* equals to $2m$ plus those edges between the sets $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$, that is, $2m + \sum_{i=1}^{n} d_G(v_i) = 4m$.

Now, we can deduce that G^{k*} has $2^k n$ vertices and $4^k m$ edges.

By Theorem 1 and the definition of k-th iterated double graph, for $k \geq 1$, we have

$$\overline{M}_1(G^{k*}) = 8\overline{M}_1(G^{(k-1)*}) + 8 \cdot (4^{k-1}m) = 8\overline{M}_1(G^{(k-1)*}) + 2^{2k+1}m. \tag{2.4}$$

By the recursive relations (4), we have

$$\begin{aligned}
\overline{M}_1(G^{k*}) &= 8\overline{M}_1(G^{(k-1)*}) + 2^{2k+1}m \\
&= 8[8\overline{M}_1(G^{(k-2)*}) + 2^{2(k-1)+1}m] + 2^{2k+1}m \\
&= 8^2\overline{M}_1(G^{(k-2)*}) + 2^{2k+2}m + 2^{2k+1}m \\
&= 8^3\overline{M}_1(G^{(k-3)*}) + 2^{2k+3}m + 2^{2k+2}m + 2^{2k+1}m \\
&= \ldots \\
&= 8^k\overline{M}_1(G^{(0)*}) + 2^{3k}m + \ldots + 2^{2k+3}m + 2^{2k+2}m + 2^{2k+1}m \\
&= 8^k\overline{M}_1(G) + 2^{2k+1}(2^k - 1)m.
\end{aligned}$$

Let us proceed to (ii). By Theorem 1 and the definition of k-th iterated double graph, for $k \geq 1$, we have

$$\overline{M}_2(G^{k*}) = 8\overline{M}_2(G^{(k-1)*}) - 8M_2(G^{(k-1)*}) + 16(4^{k-1}m)^2,$$

that is,

$$\overline{M}_2(G^{k*}) = 8\overline{M}_2(G^{(k-1)*}) - 8M_2(G^{(k-1)*}) + 16^k m^2. \tag{2.5}$$

By a similar argument to that employed in Theorem 1 to treat the second Zagreb coindex and using Eq. (3) at the same time, we obtain $\sum_{x_i y_j \in E(G^*)} d_{G^*}(x_i) d_{G^*}(y_j) = 4\sum_{i=1}^{n} D_G^2(v_i) = 8M_2(G)$.

In view of this equality, we obtain

$$\begin{aligned}
M_2(G^*) &= \sum_{x_ix_j \in E(G^*)} d_{G^*}(x_i)d_{G^*}(x_j) + \sum_{y_iy_j \in E(G^*)} d_{G^*}(y_i)d_{G^*}(y_j) + \\
&\quad \sum_{x_iy_j \in E(G^*)} d_{G^*}(x_i)d_{G^*}(y_j) \\
&= 4\sum_{v_iv_j \in E(G)} d_G(v_i)d_G(v_j) + 4\sum_{v_iv_j \in E(G)} d_G(v_i)d_G(v_j) + \\
&\quad \sum_{x_iy_j \in E(G^*)} d_{G^*}(x_i)d_{G^*}(y_j) \\
&= 8M_2(G) + 8M_2(G) \\
&= 16M_2(G).
\end{aligned}$$

So we have the recursive relation $M_2(G^{k*}) = 16M_2(G^{(k-1)*})$ for each $k \geq 1$, and then

$$M_2(G^{k*}) = 16^kM_2(G^{(0)*}) = 16^kM_2(G). \tag{2.6}$$

By Eqs. (5) and (6), we obtain

$$\begin{aligned}
\overline{M}_2(G^{k*}) &= 8\overline{M}_2(G^{(k-1)*}) - 8\cdot 16^{k-1}M_2(G) + 16^km^2 \\
&= 8[8\overline{M}_2(G^{(k-2)*}) - 8\cdot 16^{k-2}M_2(G) + 16^{k-1}m^2] - 8\cdot 16^{k-1}M_2(G) + 16^km^2 \\
&= 8^2\overline{M}_2(G^{(k-2)*}) - 8^2\cdot 16^{k-2}M_2(G) - 8\cdot 16^{k-1}M_2(G) + 8\cdot 16^{k-1}m^2 + 16^km^2 \\
&= \ldots \\
&= 8^k\overline{M}_1(G^{(0)*}) - [8\cdot 16^{k-1} + 8^2\cdot 16^{k-2} + \ldots + 8^k\cdot 16^0)]M_2(G) + \\
&\quad [16^k + 8\cdot 16^{k-1} + \ldots + 8^{k-1}\cdot 16^1]m^2 \\
&= 8^k\overline{M}_1(G) - [8\cdot 16^{k-1} + 8^2\cdot 16^{k-2} + \ldots + 8^k\cdot 16^0)]M_2(G) + \\
&\quad [16^k + 8\cdot 16^{k-1} + \ldots + 8^{k-1}\cdot 16^1]m^2 \\
&= 8^k\overline{M}_1(G) - [8^k(2^k-1)]M_2(G) + 2[8^k(2^k-1)]m^2,
\end{aligned}$$

as claimed. □

References

[1] A.R. Ashrafi, T. Došlić, A. Hamzeh, The Zagreb coindices of graph operations, *Discrete Appl. Math.*, 158 (2010) 1571-1578.

[2] A.R. Ashrafi, T. Došlić, A. Hamzeh, Extremal graphs with respect to the Zagreb coindices, *MATCH Commun. Math. Comput. Chem.*, 65 (2011) 85-92.

[3] A. Behtoei, M. Jannesari, B. Taeri, Maximum Zagreb index, minimum hyper-Wiener index and graph connectivity, *Appl. Math. Lett.*, 22 (2009) 1571-1576.

[4] H. Hua, Zagreb M_1 index, indenpedence number and connectivity in graphs, *MATCH Commun. Math. Comput. Chem.*, 60 (2008) 45-56.

[5] H. Hua, S. Zhang, Relations between Zagreb coindices and some distance-based topological indices, *MATCH Commun. Math. Comput. Chem.*, 68 (2012) 199-208.

[6] H. Hua, On the degree distance of some composite graphs, *Bull. Aust. Math. Soc.*, 85 (2012) 164-171.

[7] H. Hua, S. Zhang, K. Xu, Further results on the eccentric distance sum, *Discrete Appl. Math.*, 160 (2012) 170-180.

[8] M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, The first and second Zagreb indices of graph operations, *Discrete Appl. Math.*, 157 (2009) 804-811.

[9] E. Munarini, C. Perelli Cippo, A. Scagliola, N. Zagaglia Salvi, Double graphs, *Discrete Math.*, 308 (2008) 242-254.

[10] P.S. Ranjini, V. Lokesha, I.N. Cangül, On the Zagreb indices of the line graphs of the subdivision graphs, *Appl. Math. Comput.*, 218 (2011) 699-702.

[11] R. Todeschini, V. Consoni, Handbook of Molecular Descriptors, Wiley-VCH, New York, 2002.

[12] J. Wang, Q. Ma, Some results on edge cover coloring of double graphs, *Appl. Math.*, 3 (2012) 264-266.

[13] L. Xu, Y. Hou, Equienergetic bipartite graphs, *MATCH Commun. Math. Comput. Chem.*, 57 (2007) 363-370.

[14] K. Xu, The Zagreb indices of graphs with a given clique number, *Appl. Math. Lett.*, 24 (2011) 1026-1030.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1143 – 1159

The gamma half-Cauchy distribution: properties and applications

Ayman Alzaatreh*, M. Mansoor†, M. H. Tahir‡, §, M. Zubair¶ and Shakir Ali Ghazali‖

Abstract

A new distribution, namely, the *Gamma-Half-Cauchy* distribution is proposed. Various properties of the *Gamma-Half-Cauchy* distribution are studied in detail such as limiting behavior, moments, mean deviations and Shannon entropy. The model parameters are estimated by the method of maximum likelihood and the observed information matrix is obtained. Two data sets are used to illustrate the applications of *Gamma-Half-Cauchy* distribution.

Keywords: Folded Cauchy distribution, half-Cauchy distribution, gamma distribution, Shannon entropy.

2000 AMS Classification: 60E05; 62E10; 62N05.

Received : 08.12.2014 *Accepted* : 05.06.2015 *Doi* : 10.15672/HJMS.20157011058

*Department of Mathematics, Nazarbayev University, Astana-010000, Kazakhstan, Email: ayman.alzaatreh@nu.edu.kz

†Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mansoor.abbasi143@gmail.com

‡Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mtahir.stat@gmail.com, mht@iub.edu.pk

§Corresponding Author.

¶Department of Statistics, Government Degree College Khairpur Tamewali, Pakistan, Email: zubair.stat@yahoo.com

‖Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: shakir.ghazali@iub.edu.pk

1. Introduction

Half-Cauchy distribution is the folded standard Cauchy distribution around the origin so that positive values are observed. Modeling with half (or folded) distributions has been proposed and five folded distributions have been reported so far in literature, namely, the students' t, normal, normal-slash, logistic and Cauchy. These folded distributions have been used in Bayesian paradigm when a proper prior is necessary. Although some applications of the half Cauchy distribution exist in the literature, but the fact that the finite moments of order greater than or equal to one do not exist, the central limit theorem does not hold. This fact reduces the applicability of this distribution in modeling real life scenarios.

A random variable X has the half-Cauchy (HC) distribution with scale parameter $\sigma > 0$, if its cumulative distribution function (cdf) is given by

$$(1.1) \qquad F(x) = \frac{2}{\pi} \tan^{-1}(x/\sigma), \quad x > 0.$$

The probability density function (pdf) corresponding to (1.1) is

$$(1.2) \qquad f(x) = \frac{2}{\pi\,\sigma} \left[1 + (x/\sigma)^2\right]^{-1}.$$

Henceforth, we denote by $X \sim$HC(σ), the random variable having the HC density in (1.2) with parameters σ. As a heavy tailed distribution, the HC distribution has been used as an alternative to exponential distribution to model dispersal distances [18] as the former predicts more frequent long distance dispersed events than the later. Paradis *et al.* [16] used the HC distribution to model ringing data on two species of tits (Parus caeruleus and Parus major) in Britain and Ireland.

Few generalizations of the HC distribution exist in the literature, namely, beta-half-Cauchy (BHC) by Cordeiro and Lemonte [9], Kumaraswamy-half-Cauchy (KHC) by Ghosh [11] and Marshall-Olkin half-Cauchy (MOHC) by Jacob and Kayakumar [13]. In this paper, we propose a new generalization of the HC distribution using the technique defined by Alzaatreh *et al.* [7].

Let $r(t)$ be the probability density function (pdf) of a random variable $T \in [a, b]$ for $-\infty \leq a < b \leq \infty$ and let $F(x)$ be the cumulative distribution distribution function (cdf) of a random variable X such that the link function $W(\cdot) : [0, 1] \longrightarrow [a, b]$ satisfies the following conditions:

$$(1.3) \qquad \begin{cases} (i) & W(\cdot) \text{ is differentiable and monotonically non-decreasing, and} \\ (ii) & W(0) \to a \text{ and } W(1) \to b. \end{cases}$$

The *T-X family* of distributions defined by Alzaatreh *et al.* [7] as

$$(1.4) \qquad G(x) = \int_a^{W[F(x)]} r(t)\, dt.$$

If $T \in (0, \infty)$, X is a continuous random variable and $W[F(x)] = -\log[1 - F(x)]$, then the pdf corresponding to (1.4) is given by

$$(1.5) \qquad g(x) = \frac{f(x)}{1 - F(x)} r\Big(-\log\left[1 - F(x)\right]\Big) = h_f(x) r\Big(H_f(x)\Big),$$

where $h_f(x)$ and $H_f(x)$ are, respectively, the hazard and cumulative hazard function corresponding to $f(x)$. For more details about the *T-X family*, one is refer to Alzaatreh *et al.* [3, 6], Alzaatreh and Ghosh [5] and Lee *et al.* [14].

If a random variable T follows the gamma distribution with parameters α and β, $r(t) = \left(\beta^\alpha\, \Gamma(\alpha)\right)^{-1} t^{\alpha-1}\, \mathrm{e}^{-t/\beta}, \quad t \geq 0$. Then from (1.5),the pdf of *Gamma-X* family of distributions is given by

$$g(x) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)}\, f(x) \Big(-\log\big[1 - F(x)\big]\Big)^{\alpha-1} \big[1 - F(x)\big]^{\frac{1}{\beta}-1}. \tag{1.6}$$

The cdf corresponding to (1.6) is

$$G(x) = \frac{1}{\Gamma(\alpha)}\, \gamma\Big(\alpha, -\beta^{-1} \log\big[1 - F(x)\big]\Big), \tag{1.7}$$

where $\gamma(\alpha, t) = \int_0^t u^{\alpha-1}\, \mathrm{e}^{u}\, du$ is the incomplete gamma function. Several properties of gamma-X family have been studied in literature. For more details see Alzaatreh *et al.* [3, **?**, 6, 4, 8].

The paper is unfolded as follows. In Section 2, we define a new generalization of the HC distribution, namely, *Gamma-half-Cauchy* (GHC) distribution. In Section 3, some properties of the GHC are investigated. The density of the order statistics is obtained in Section 4. In Section 5, the model parameters are estimated by the method of maximum likelihood and the observed information matrix is determined. In Section 6, we explore the usefulness of the proposed distribution by means of two real data sets. Finally, Section 7 offers some concluding remarks.

2. The gamma-half Cauchy (GHC) distribution

From (1.1), (1.2), (1.6) and (1.7), it follows that the pdf and cdf of the GHC are given by

$$\begin{aligned} g(x) \quad &= \quad \frac{2}{\pi\,\sigma\,\Gamma(\alpha)\,\beta^{\alpha}} \big[1 + (x/\sigma)^2\big]^{-1} \Big(-\log\big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)^{\alpha-1} \\ &\quad \times \big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]^{\frac{1}{\beta}-1} \end{aligned} \tag{2.1}$$

and

$$G(x) = \frac{1}{\Gamma(\alpha)}\, \gamma\Big(\alpha, -\beta^{-1}\, \log\big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big), \tag{2.2}$$

respectively. Henceforth, a random variable having pdf in (2.2) is denoted by $X \sim \mathrm{GHC}(\alpha, \beta, \sigma)$.

<u>Special cases of GHC distribution:</u>

(i) If $\alpha = \beta = 1$ in (2.2), the GHC distribution reduces to the HC distribution with parameter σ.

(ii) If $\alpha = 1$ in (2.2), the GHC distribution reduces to the exponentiated HC distribution with parameters β and σ.

(iii) If $\alpha = n + 1$ and $\beta = 1$ in (2.2), the density of GHC reduces to the density of the nth upper record of the HC distribution.

Note that the special case in (ii) does not exist in the literature and it is considered another generalization of the HC distribution.

The survival function (sf), $S(x)$, hazard rate function (hrf), $h(x)$, and cumulative hazard rate function (chrf), $H(x)$, of X are, respectively, given by

$$S(x) = 1 - \frac{1}{\Gamma(\alpha)}\, \gamma\Big(\alpha, -\beta^{-1}\, \log\big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big),$$

$$h(x) = \frac{2\Big(-\log\big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)^{\alpha-1} \big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]^{\frac{1}{\beta}-1}}{\pi\,\sigma\,\beta^{\alpha}\,(1 + (x/\sigma)^2)\,\Big\{\Gamma(\alpha) - \gamma\Big(\alpha, -\beta^{-1}\,\log\big[1 - 2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)\Big\}}$$

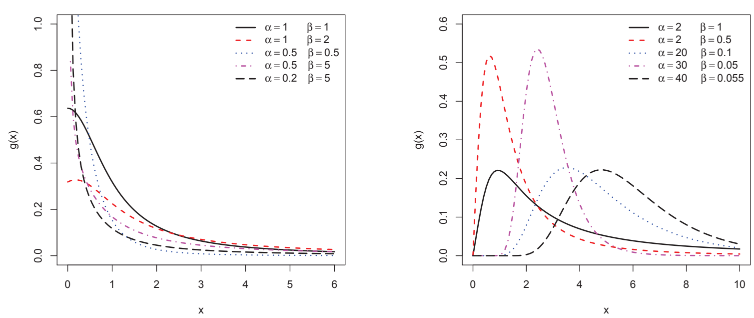
and

$$H(x) = -\log\left[1 - \frac{1}{\Gamma(\alpha)}\,\gamma\Big(\alpha, -\beta^{-1}\log\left[1 - 2\,\pi^{-1}\tan^{-1}(x/\sigma)\right]\Big)\right].$$

2.1. Asymptotic behavior of the pdf. The limit of the pdf of X as $x \to \infty$ is 0. Further, the limits of the pdf of X as $x \to 0^+$ are given by

$$\lim_{x\to 0^+} g(x) = \begin{cases} \infty, & \text{if } \alpha < 1 \\ \frac{2}{\pi\,\sigma\,\beta}, & \text{if } \alpha = 1, \\ 0, & \text{if } \alpha > 1. \end{cases}$$

In Figures 1 and 2, various graphs of the density when $\sigma = 1$ and for different values of α and β are displayed. Figure 1 indicates that the GHC distribution is well-suited for right-skewed data. For fixed $\alpha \le 1$, the density is always reversed-J shaped. For fixed $\alpha > 1$, the peakedness increases as β decreases. Also, Figure 2 shows that the hazard function of the GHC distribution has DFR (decreasing failure) or UBT (upside down bathtub) properties.



(a) (b)

Figure 1. Plots of the GHC densities for various values of α and β.

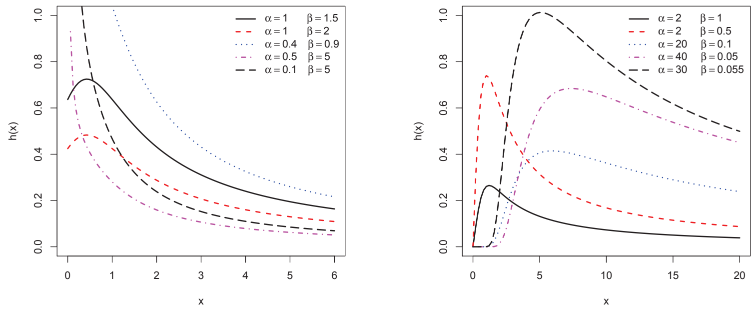
3. Properties of the GHC distribution

In this section, we provide some properties of the GHC distribution. Some proofs are omitted in case of trivial results.

The following Lemma gives the relation between GHC and gamma distributions.

3.1. Transformation.

3.1. Lemma. *If a random variable Y follows the gamma distribution with parameters α and β, then $X = \sigma \cot\left(\frac{\pi}{2}\,\mathrm{e}^{-Y}\right) \sim \mathrm{GHC}(\alpha, \beta, \sigma)$.*



(a) (b)

Figure 2. Plots of GHC hazard rates for various values of α and β.

3.2. Mode.

3.2. Lemma. *The mode of GHC distribution is the solution of* $k(x) = 0$*, where*

$$\begin{aligned} k(x) &= -\frac{x}{\sigma} + \pi^{-1}\Big[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\Big]^{-1} \\ &\times \left\{\frac{\alpha - 1}{\log\left[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\right]^{-1}} - \frac{1}{\beta} + 1\right\}. \end{aligned} \tag{3.1}$$

Proof. Setting $g'(x)$ is equivalent to,

$$\begin{aligned} g'(x) &= \frac{4\,[1 + (x/\sigma)]^{-2}}{\pi\sigma^2\Gamma(\alpha)\beta^\alpha}\Big(-\log\big[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)^{\alpha-1} \\ &\times\Big[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\Big]^{\frac{1}{\beta}-1} \times k(x), \end{aligned} \tag{3.2}$$

where

$$\begin{aligned} k(x) &= -(x/\sigma) + \pi^{-1}\left[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\right]^{-1} \\ &\times\left\{\frac{\alpha - 1}{\log\left[1 - 2\pi^{-1}\,\tan^{-1}(x/\sigma)\right]^{-1}} - \frac{1}{\beta} + 1\right\}. \end{aligned} \tag{3.3}$$

Hence the critical values of $g(x)$ is the solution of $k(x) = 0$. □

Note that equation implies the following; when $\alpha = \beta = 1$, the mode of GHC is at $x = 0$ which is the mode of HC distribution. When $\alpha < 1$, implies that $x < 0$ and as $x \to 0^+$, $k(x) \to \infty$. Also, when $\alpha = 1$, $x = 0$ is a modal point and as $x \to 0^+$, $k(x) \to \frac{2}{\pi\,\sigma\,\beta}$. Hence, when $\alpha \leq 1$, GHC has a unique mode at $x = 0$.

3.3. Quantile function. The following Lemma gives the quantile function for the GHC distribution.

3.3. Lemma. *The mode of GHC distribution is given by*

$$Q(\lambda) = \sigma\cot\Big(0.5\,\pi\,\mathrm{e}^{-\beta\,\gamma^{-1}\left(\alpha,\,\lambda\,\Gamma(\alpha)\right)}\Big). \tag{3.4}$$

Proof. Follows by inverting equation 2.1. □

3.4. Shannon entropy.

3.4. Theorem. *The Shannon entropy for the GHC distribution is given by*

$$\eta_X = 3\log(0.5\,\pi) + \alpha(\beta-1) + \log\big(\beta\,\Gamma(\alpha)\big) + (1-\alpha)\,\psi(\alpha) - 2\sum_{k=1}^{\infty} w_k\,(1+2k\beta)^{-\alpha}, \tag{3.5}$$

where $\psi(\cdot)$ *is the digamma function and* $w_k = \frac{(-1)^k\,(\pi)^{2k}\,B_{2k}}{2k\,(2k)!}$.

Proof. Based on Alzaatreh *et al.* [8], the Shannon entropy for the gamma-X family is given by

$$\eta_X = -\mathbb{E}\left\{\log f\left(F^{-1}\left(1-\mathrm{e}^{-T}\right)\right)\right\} + \alpha(1-\beta) + \log\big(\beta\,\Gamma(\alpha)\big) + (1-\alpha)\,\psi(\alpha), \tag{3.6}$$

where $T \sim Gamma(\alpha,\beta)$.

We first need to find $-\mathbb{E}\left\{\log f\left(F^{-1}\left(1-\mathrm{e}^{-T}\right)\right)\right\}$, where $f(x)$ and $F(x)$ are the pdf and cdf of HC distribution. It follows immediately that $\log f\left(F^{-1}\left(1-\mathrm{e}^{-T}\right)\right) = \log(0.5\,\pi) + 2\log\big(\sin(0.5\,\pi\,\mathrm{e}^{-T})\big)$ and hence by using the series expansion for $\log\big(\sin(0.5\,\pi\,\mathrm{e}^{-T})\big)$ (see [12]) as

$$\log\big(\sin(0.5\,\pi\,\mathrm{e}^{-T})\big) = \log(0.5\,\pi) - T + \sum_{k=1}^{\infty} \underbrace{\frac{(-1)^k\,(\pi)^{2k}\,B_{2k}}{2k\,(2k)!}}_{w_k}\,\mathrm{e}^{-2kT}, \tag{3.7}$$

where B_{2k} is the Bernoulli number.

Therefore,

$$-\mathbb{E}\left\{\log f\left(F^{-1}\left(1-\mathrm{e}^{-T}\right)\right)\right\} = 3\,\log(0.5\,\pi) - 2\mathbb{E}(T) + 2\sum_{k=1}^{\infty} w_k\,\mathbb{E}\big(\mathrm{e}^{-2kT}\big). \tag{3.8}$$

The results in (3.5) followed by substituting (3.8) in (3.6) and noting that $\mathbb{E}(T) = \alpha\beta$ and $\mathbb{E}\big(\mathrm{e}^{-2kT}\big) = (1+2k\beta)^{-\alpha}$. □

3.5. Moments. By using the Lemma 3.1, the rth moments of GHC distribution can be written as

$$\mathbb{E}(X^r) = \frac{\sigma^r}{\beta^\alpha\,\Gamma(\alpha)} \int_0^\infty \Big(\cot(0.5\,\pi\,\mathrm{e}^{-u})\Big)^r\, u^{\alpha-1}\,\mathrm{e}^{-u/\beta}\,du. \tag{3.9}$$

A series expansion for $\cot(0.5\,\pi\,\mathrm{e}^{-u})$ can be obtained from [12] as follows

$$\cot(0.5\,\pi\,\mathrm{e}^{-u}) = \sum_{k=0}^{\infty} v_k\,\mathrm{e}^{-(2k-1)u}, \tag{3.10}$$

where $v_k = \frac{2\,(-1)^k\,(\pi)^{2k-1}\,B_{2k}}{(2k)!}$.

Hence,

$$\Big(\cot(0.5\,\pi\,\mathrm{e}^{-u})\Big)^r = \sum_{k_1,\dots,k_r=0}^{\infty} v_{k_1,\dots,k_r}\,\mathrm{e}^{-(2s_r-1)u},$$

where $v_{k_1,\dots,k^r} = v_{k=1}\,v_{k=2}\,\dots\,v_{k=r}$ and $s_r = k_1 + k_2 + \cdots + k_r$.

Therefore, from (3.9) we get

$$\mathbb{E}(X^r) = \sigma^r \sum_{k_1,\dots,k_r=0}^{\infty} v_{k_1,\dots,k_r}\,\big(2\beta\,s_r - \beta + 1\big)^{-\alpha}. \tag{3.11}$$

3.5. Theorem. *Let* $X \sim GHC(\alpha,\beta,\sigma)$, *then* $\mathbb{E}(X^r)$ *exists iff* $\beta < r^{-1}$.

Proof. The rth moment of GHC can be obtained from

$$\mathbb{E}(X^r) = \int_0^1 x^r\, g(x)\, dx + \int_1^\infty x^r\, g(x)\, dx, \tag{3.12}$$

where $g(x)$ is defined in (2.2).

Without loss of generality assume $\sigma = 1$. From (3.12), the existence of $\mathbb{E}(X^r)$ equivalent to the existence of $\int_1^\infty x^r\, g(x)\, dx$. Now,

$$\int_1^\infty x^r\, g(x)\, dx = \frac{1}{\pi\, \beta^\alpha\, \Gamma(\alpha)}\ \mathbb{I}, \tag{3.13}$$

where

$$\begin{aligned}\mathbb{I} \;&=\; \int_1^\infty \frac{x^r}{1+x^2} \left\{-\log\left[1-0.5\pi^{-1}\ \tan^{-1}(x)\right]\right\}^{\alpha-1} \\ &\qquad\qquad \times \left[1-0.5\pi^{-1}\ \tan^{-1}(x)\right]^{\frac{1}{\beta}-1} dx.\end{aligned} \tag{3.14}$$

Consider the following inequality (Abramowitz and Stegun [1])

$$x < -\log(1-x) < \frac{x}{1-x}, \quad x < 1, x \neq 0. \tag{3.15}$$

Now, for $\alpha \geq 1$, one can use the right hand-side of the inequality in (3.15) to show that

$$I < \int_1^\infty \underbrace{\frac{x^r}{1+x^2} \left[0.5\pi^{-1}\ \tan^{-1}(x)\right]^{\alpha-1} \left[1-0.5\pi^{-1}\ \tan^{-1}(x)\right]^{\frac{1}{\beta}-\alpha} dx}_{\tau_1(x)}. \tag{3.16}$$

Let $\tau_2(x) = x^{-\frac{1}{\beta}+\alpha+r-2}$, then $\lim_{x\to\infty} \frac{\tau_1(x)}{\tau_2(x)} = \left(0.5\pi^{-1}\right)^{\frac{1}{\beta}-\alpha}$. Therefore, $\int_1^\infty \tau_1(x)$ exists iff $\int_1^\infty \tau_2(x)$ exists iff $\frac{1}{\beta} > \alpha + r - 1$. Since $\alpha \geq 1$, this implies that $\frac{1}{\beta} > r$. If $\alpha < 1$, the left hand side of the inequality in (3.15) implies that

$$I < \int_1^\infty \frac{x^r}{1+x^2} \left[0.5\pi^{-1}\ \tan^{-1}(x)\right]^{\alpha-1} \left[1-0.5\pi^{-1}\ \tan^{-1}(x)\right]^{\frac{1}{\beta}-1} dx. \tag{3.17}$$

Similarly, one can show the right hand side of the integrand in exists iff $\frac{1}{\beta} > r$. This ends the proof. $\square$

3.6. Mean deviations. The mean deviations about the mean $\big(\delta_1(X) = \mathbb{E}(|X - \mu_1'|)\big)$ and about the median $\big(\delta_2(X) = \mathbb{E}(|X - M|)\big)$ of X can be expressed as

$$\delta_1(X) = 2\mu_1' F(\mu_1') - 2m_1(\mu_1') \quad \text{and} \quad \delta_2(X) = 2\mu_1' - 2m_1(M), \tag{3.18}$$

respectively, where $\mu_1' = \mathbb{E}(x)$ can be obtained from (3.11) by setting $r = 1$ and M is the median of the GHC which can be calculated from Lemma 3.3 as

$$M = \sigma \cot\left(0.5\, \pi\, \mathrm{e}^{-\beta\, \gamma^{-1}\left(\alpha,\, 0.5\, \Gamma(\alpha)\right)}\right). \tag{3.19}$$

Further, $F(\mu_1')$ can easily be computed from the (2.1) and $m_1(z) = \int_0^z x\, f(x)\, dx$ (the first incomplete moment of X) can be computed from

$$m_1(z) = \int_0^z \cot(0.5\, \pi\, \mathrm{e}^{-u})\, u^{\alpha-1}\, \mathrm{e}^{-u/\beta}\, du. \tag{3.20}$$

The result immediately follows from (3.10) as

$$m_1(z) = \frac{\sigma}{\Gamma(\alpha)} \sum_{k=0}^\infty v_k\, (1 + 2\beta\, k - \beta)^{-\alpha}\ \gamma\Big(\alpha, \frac{z}{\beta}(1 + 2\beta\, k - \beta)\Big). \tag{3.21}$$

3.7. Mean residual life function. Let X be a random variable with cdf F such that $\mathbb{E}(X) < \infty$. The mean residual life (MRL) function $\boldsymbol{\xi}(x)$ of X is defined by $\boldsymbol{\xi}(x) = \mathbb{E}(X - x|X > x)$. It plays a major rule in many fields such as industrial reliability, life insurance and biomedical science. The following theorem provides an expansion for the MRL for the GHC distribution.

3.6. Theorem. *Let X be a random variable which follows the $GHC(\alpha, \beta, \sigma)$ such that $\beta < 1$, then the MRL function is given by*

$$\boldsymbol{\xi}(x) = \frac{\sigma}{\Gamma(\alpha)\, S(x)} \sum_{k=0}^{\infty} v_k \, \frac{\Gamma(\alpha, (2k + \beta^{-1} - 1)\, x)}{(2\beta\, k - \beta + 1)^{\alpha}} - x \tag{3.22}$$

where $\Gamma(x, a) = \int_x^{\infty} t^{a-1}\, \mathrm{e}^{-t}\, dt$ is the upper incomplete gamma function and v_k is defined in (3.10) and $S(x)$ is survival function of GHC defined in section 2.

Proof. From Lemma (3.1)

$$\mathbb{E}(X|X > x) = \frac{\sigma}{\beta^{\alpha}\, \Gamma(\alpha) S(x)} \int_y^{\infty} \cot\left(0.5\pi \mathrm{e}^{-y}\right)\, y^{\alpha-1}\, \mathrm{e}^{-y/\beta}\, dy.$$

On using the expansion in (3.10), one can get the result in (3.22). □

3.8. Reliability estimation. The reliability parameter R is defined as $R = P(X > Y)$, where X and Y are independent random variables. Many applications of the reliability parameter have appeared in the literature such as the area of classical stress-strength model and the breakdown of a system having two components. If X and Y are two continuous random variables with cdfs $F_1(x)$ and $F_2(y)$ and their pdfs $f_1(x)$ and $f_2(y)$ respectively. Then, the reliability parameter R can be written as

$$R = P(X > Y) = \int_{-\infty}^{\infty} F_2(x)\, f_1(x)\, dx. \tag{3.23}$$

3.7. Theorem. *Suppose that X and Y are two independent GHC random variables with parameters α_1, β_1 and α_2, β_2, and fixed scale parameter σ. Then*

$$R = \frac{1}{\Gamma(\alpha_1)\, \Gamma(\alpha_2)} \sum_{k=0}^{\infty} \left(\frac{\beta_1}{\beta_2}\right)^{\alpha_1 + k} \frac{(-1)^k\, \Gamma(\alpha_1 + \alpha_2 + k)}{k!\, \Gamma(\alpha_2 + k)}. \tag{3.24}$$

Proof. On using the following series expansion from [1]

$$\gamma(\alpha, x) = \sum_{k=0}^{\infty} \frac{(-1)^k\, x^{k+\alpha}}{k!\ (k + \alpha)}, \tag{3.25}$$

and then substituting $u = -\log\left[1 - 2\pi^{-1}\, \tan^{-1}(x/\sigma)\right]$, (3.23) reduces to

$$R = \frac{1}{\beta_1^{\alpha_1}\, \Gamma(\alpha_1)\, \Gamma(\alpha_2)} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\, (\alpha_2 + k)\, \beta_2^{\alpha_2 + k}} \int_0^{\infty} u^{\alpha_1 + \alpha_2 + k - 1}\, \mathrm{e}^{-u/\beta_1}\, du. \tag{3.26}$$

The result in (3.24) follows immediately from (3.26). □

3.9. Mixture representation of GHC density.

3.8. Theorem. *The GHC distribution is the linear combination of infinite exponentiated-HC densities*

$$g(x) = \sum_{k=0}^{\infty} w_{i,j}\, h_{(\alpha+k+i,\, \sigma)}(x), \tag{3.27}$$

where $h_{(\alpha+k+i,\,\sigma)}(x)$ *is the exponentiated-HC density with power parameter* $\alpha+k+i$ *and*

$$w_{i,j}=\sum_{j=0}^{k}\sum_{i=0}^{\infty}\binom{k+1-\alpha}{k}\binom{k}{j}\binom{\frac{1}{\beta}-1}{i}\frac{(-1)^{j+k+i}\;p_{j,k}}{(\alpha-j-1)(\alpha+k+i)\,\Gamma(\alpha)\,\beta^{\alpha}}.$$

Proof. Based on the formula given at
`http:// functions.wolfram.com/ ElementaryFunctions/Log/06/01/04/03/`, we can write

$$\Big(-\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)^{\alpha-1}=$$

$$(\alpha-1)\sum_{k=0}^{\infty}\binom{k+1-\alpha}{k}\sum_{j=0}^{k}\frac{(-1)^{j+k}\binom{k}{j}\,p_{j,k}}{(\alpha-j-1)}\Big[\frac{2}{\pi}\,\tan^{-1}(x/\sigma)\Big]^{\alpha+k-1}. \tag{3.28}$$

Here, the constants $p_{j,k}$ (for $j\geq 0$ and $k\geq 1$) can be determined recursively by

$$p_{j,k}=k^{-1}\sum_{m=1}^{\infty}[k-m(j+1)]\,c_m\;p_{j,k-m},$$

where $p_{j,0}=1$ and $c_k=(-1)^{k+1}\,(k+1)^{-1}$.
Now, using the generalized binomial series expansion

$$\Big[1-2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\Big]^{\frac{1}{\beta}-1}=\sum_{i=0}^{\infty}(-1)^{i}\binom{\frac{1}{\beta}-1}{i}\Big[\frac{2}{\pi}\,\tan^{-1}(x/\sigma)\Big]^{i}, \tag{3.29}$$

where $\binom{a}{i}=a(a-1)\cdots(a-i+1)/i!$.
The result (3.27) follows immediately by substituting (3.28) and (3.29) in (2.2). □

Note that the second summation in $w_{i,j}$ is finite whenever β^{-1} is a natural number.

4. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1,\ldots,X_n$ is a random sample from the GHC distribution. Let $X_{i:n}$ denote the ith order statistic. Then, the pdf of $X_{i:n}$ can be expressed as

$$\begin{aligned}
f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!}\,f(x)\,F(x)^{i-1}\,\{1-F(x)\}^{n-i}\\
&= \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}(-1)^{j}\binom{n-i}{j}\,f(x)\,F(x)^{j+i-1}.
\end{aligned}$$

Inserting (2.1) and (2.2) in the last equation and after some algebra, we obtain

$$\begin{aligned}
f_{i:n}(x) \;=\; \sum_{j=0}^{n-i}\frac{(-1)^{j}\,\Gamma(n+1)\,(i+j)^{-1}}{\Gamma i\;\Gamma(j+1)\Gamma(n-i-j+1)}&\Bigg\{\frac{2}{\pi\,\sigma\,\Gamma(\alpha)\,\beta^{\alpha}}\,\big[1+(x/\sigma)^2\big]^{-1}\\
&\times\Big(-\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)^{\alpha-1}\Big[1-2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\Big]^{\frac{1}{\beta}-1}\\
&\times\left[\frac{\gamma\Big(\alpha,-\beta^{-1}\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x/\sigma)\big]\Big)}{\Gamma(\alpha)}\right]^{j+i-1}\Bigg\}.
\end{aligned}$$

Hence,

$$f_{i:n}(x)=\sum_{j=0}^{n-i}\eta_j\;f_{\alpha,\beta,(j+i)}(x), \tag{4.1}$$

where

$$\eta_j = \frac{(-1)^j\,\Gamma(n+1)}{(i+j)\,\Gamma(i)\,\Gamma(j+1)\,\Gamma(n-i-j+1)}$$

and $f_{\alpha,\beta,(j+i)}(x)$ is the exponentiated-GHC density with parameters $\big(\alpha,\beta,(i+j)\big)$.

Equation (4.1) is the main result of this section. It reveals that the pdf of the GHC order statistics is a linear combination of exponentiated-GHC densities. So, several mathematical quantities of these order statistics like ordinary and incomplete moments, factorial moments, mgf, mean deviations and several others can be derived from those quantities of the GHC distribution.

5. Estimation and information matrix

In this section, the method of maximum likelihood estimation is used to estimate the GHC distribution parameters. The maximum likelihood estimates (MLEs) enjoy desirable properties that can be used when constructing confidence intervals and regions and deliver simple approximations that work well in finite samples. The resulting approximation for the MLEs in distribution theory is easily handled either analytically or numerically. Let $x_1,\dots,x_n$ be a sample of size n from the GHC distribution given by (2.2). The log-likelihood function for the vector of parameters $\Theta = (\alpha,\beta,\sigma)^\top$ can be expressed as

$$\begin{aligned}\ell &= n\,\log\left[\frac{2}{\pi\,\sigma\,\Gamma(\alpha)\,\beta^{\alpha}}\right] - \sum_{i=1}^{n}\log\left[1+(x_i/\sigma)^2\right]\\ &\quad +(\alpha-1)\sum_{i=1}^{n}\log\Big(-\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]\Big)\\ &\quad +\Big(\tfrac{1}{\beta}-1\Big)\sum_{i=1}^{n}\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]\end{aligned}$$

The components of the score vector $J(\Theta)$ are given by

$$\begin{aligned}J_\alpha &= -n\,\psi(\alpha) - n\,\log\beta + \sum_{i=1}^{n}\log\Big(-\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]\Big),\\ J_\beta &= -n\,\alpha\,\beta^{-1} - \beta^{-2}\sum_{i=1}^{n}\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right],\\ J_\sigma &= -n\,\sigma^{-1} + 2\sigma^{-3}\sum_{i=1}^{n}x_i^2\,\left[1+(x_i/\sigma)^2\right]^{-1}\\ &\quad -2(\alpha-1)\,\pi^{-1}\,\sigma^{-2}\sum_{i=1}^{n}\left\{\frac{x_i\,\tan^{-1\prime}(x_i/\sigma)\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]^{-1}}{-\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]}\right\}\\ &\quad +2\left(\tfrac{1}{\beta}-1\right)\,\pi^{-1}\,\sigma^{-2}\sum_{i=1}^{n}\left\{\frac{x_i\,\tan^{-1\prime}(x_i/\sigma)}{\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]}\right\}.\end{aligned}$$

Setting these equations to zero and solving them simultaneously yield the maximum likelihood estimates (MLEs) of the model parameters. Numerical methods can be used to obtain the MLE $\widehat{\Theta}$. For example, the Newton-Raphson iterative technique could be applied to solve the likelihood equations and obtain $\widehat{\Theta}$ numerically. For interval estimation of the parameters, we require the 3×3 observed information matrix $J(\Theta) = \{-J_{rs}\}$ (for $r,s = \alpha,\beta,\sigma$) given in Appendix A. The observed information matrix can be determined numerically from standard maximization routines, which provide the observed information matrix as part of their output; e.g., one can use the R functions `optim` or `nlm`, the Ox function `MaxBFGS`, the SAS procedure `NLMixed`, among others, to compute $J(\Theta)$ numerically.

Under standard regularity conditions, the multivariate normal $N_3(0, J(\widehat{\Theta})^{-1})$ distribution can be used to construct approximate confidence intervals for the model parameters. Here, $J(\widehat{\Theta})$ is the total observed information matrix evaluated at $\widehat{\Theta}$. Then, the $100(1-\gamma)\%$ confidence intervals for α, β and σ are given by $\hat{\alpha} \pm z_{\gamma^*/2} \times \sqrt{var(\hat{\alpha})}$, $\hat{\beta} \pm z_{\gamma^*/2} \times \sqrt{var(\hat{\beta})}$ and $\hat{\sigma} \pm z_{\gamma^*/2} \times \sqrt{var(\hat{\sigma})}$, respectively, where the $var(\cdot)$'s denote the diagonal elements of $J(\widehat{\Theta})^{-1}$ corresponding to the model parameters, and $z_{\gamma^*/2}$ is the quantile $(1-\gamma^*/2)$ of the standard normal distribution.

The likelihood ratio (LR) statistic can be used to check if the GHC distribution is strictly "superior" to the HC distribution for a given data set. The test of $H_0 : \alpha = \beta = 1$ versus $H_1 : H_0$ *is not true* is equivalent to compare the GHC and HC distributions and the statistic $w = -2\log\lambda = 2\{\ell(\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}) - \ell(1, 1, \tilde{\sigma})\}$, where $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}$ are the MLEs under H_1 and $\tilde{\sigma}$ is the MLE under H_0, is asymptotically follows chi-square distribution with 2 degrees of freedom. Similarly, the test of $H_0 : \alpha = 1$ versus $H_1 : \alpha \neq 1$ is equivalent to compare the GHC and exponentiated HC distributions with the statistic $w = 2\{\ell(\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}) - \ell(1, \tilde{\beta}, \tilde{\sigma})\}$, where $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}$ are the MLEs under H_1 and $\tilde{\beta}$ and $\tilde{\sigma}$ are the MLEs under H_0. In this case w is asymptotically follows chi-square distribution with 1 degrees of freedom.

5.1. Simulation study. We evaluate the performance of the maximum likelihood method for estimating the GHC parameters using Monte Carlo simulation for a total of twenty four parameter combinations and the process is repeated 200 times. Two different sample sizes $n = 100$ and 300 are considered. The MLEs and the standard deviations of the parameter estimates are listed in Table 1. The MLEs of α, β and σ are determined by solving the nonlinear equations $U(\Theta) = \mathbf{0}$. From Table 1, we note that the ML method performs well for estimating the model parameters. Also, as the sample size increases, the biases and the standard deviations of the MLEs decrease as expected.

Table 1: MLEs and standard deviations for various parameter values.

Sample size	Actual values			Estimated values			Standard deviations		
n	α	β	σ	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\sigma}$	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\sigma}$
100	0.5	0.5	1	0.5267	0.4094	3.6791	0.0060	0.0272	0.6534
	0.5	1.0	2	0.5212	0.9324	2.9044	0.0080	0.0308	0.4838
	0.5	1.5	1	0.5315	1.4004	1.1285	0.0085	0.0329	0.0556
	0.5	2.0	2	0.5168	1.9218	2.36179	0.0100	0.0426	0.1342
	1.0	0.5	1	1.0416	0.4409	2.2191	0.0164	0.0176	0.4728
	1.0	1	2	1.0741	0.9578	2.1989	0.0605	0.0186	0.0939
	1.0	1.5	1	1.3303	1.4166	1.0224	0.1236	0.0274	0.0513
	1.0	2.0	2	1.4304	1.8939	1.9073	0.1399	0.0424	0.1084
	1.5	0.5	1	1.7037	0.4683	1.3396	0.0992	0.0111	0.3024
	1.5	1.0	2	2.2656	0.9118	2.0189	0.2288	0.0194	0.1082
	1.5	1.5	1	2.1739	1.3711	0.9861	0.1726	0.0315	0.0570
	1.5	2.0	2	2.1626	1.8253	2.1688	0.1758	0.0455	0.2187
300	0.5	0.5	1	0.5070	0.4529	1.7515	0.0020	0.0100	0.1402
	0.5	1	2	0.5040	0.9787	2.1165	0.0022	0.0095	0.0311
	0.5	1.5	1	0.5075	1.4764	1.0328	0.0027	0.0124	0.0153
	0.5	2.0	2	0.5014	1.9610	2.1106	0.0026	0.0138	0.0291
	1.0	0.5	1	1.0140	0.5001	1.0231	0.0052	0.0047	0.0159
	1.0	1.0	2	1.0120	0.9854	2.0763	0.0069	0.0061	0.0299
	1.0	1.5	1	1.0196	1.4891	1.0263	0.0077	0.0075	0.0148
	1.0	2.0	2	1.0281	1.9801	2.0308	0.0084	0.0107	0.0314
	1.5	0.5	1	1.5326	0.4970	1.0183	0.0104	0.0036	0.0157
	1.5	1.0	2	1.6108	0.9887	1.9955	0.035	0.0059	0.0381
	1.5	1.5	1	1.7063	1.4497	0.9605	0.0479	0.0109	0.0193
	1.5	2.0	2	1.6754	1.9499	1.9397	0.0335	0.0160	0.0475

6. Applications

In this section, we provide two applications to real data to illustrate the importance of the GHC distribution. The model parameters are estimated by the method of maximum likelihood and three well-recognized goodness-of-fit statistics are calculated to compare the GHC distribution with other competing models.

The first data set represents the annual food discharge rates for the 39 years (1935-1973) at Floyd River located in James, Iowa, USA. The Floyd River data were reported by Mudholkar and Hutson [15] and Akinsete *et al.* [2]. The second data set consists of the waiting times between 65 consecutive eruptions of the Kiama Blowhole (da Silva *et al.* [10]; Pinho *et al.* [17]). The Kiama Blowhole is a tourist attraction located nearly 120km to the south of Sydney. The swelling of the ocean pushes the water through a hole bellow a cliff. The water then erupts through an exit usually drenching whoever is nearby. The times between eruptions of a 1340 hours period starting from July 12th of 1998 were recorded using a digital watch. Both data sets are reported in Appendix B.

We fitted the GHC model to the three data sets and compared it with other models: the BHC, KHC, EHC and HC. The measures of goodness-of-fit statistics including the log-likelihood function evaluated at the MLEs $(-\log \hat{\ell})$, Akaike information criterion (AIC) and Kolmogrov-Smirnov (K-S) are computed to compare the fitted models. In general, the smaller the values of these statistics, the better the fit to the data. The required computations are carried out using the R-software.

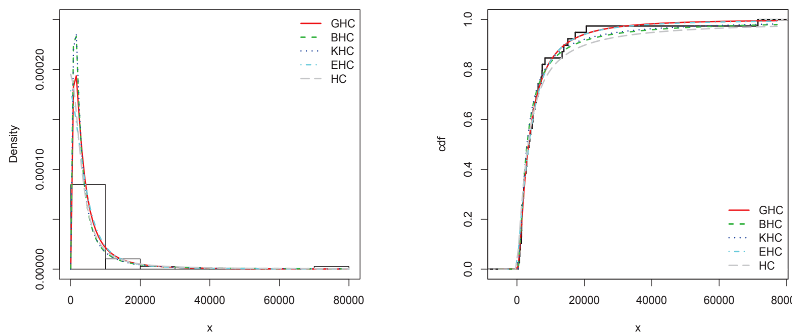
Table 2: The statistics $-\log \hat{\ell}$, AIC and K-S for the data sets 1, 2 and 3.

Distribution	α	β	σ	$-\log \hat{\ell}$	AIC	K-S	K-S p-value
			Data set 1				
GHC	45.9778	0.1554	4.4487	**376.3683**	**7758.7366**	**0.0648**	**0.9932**
	(101.8881)	(0.1758)	(34.8257)				
BHC	61.8037	1.1002	59.2418	377.9875	761.9750	0.0856	0.9141
	(158.0451)	(0.2233)	(146.9170)				
KHC	73.2921	1.1512	50.0133	377.8883	761.7766	0.0833	0.9287
	(322.4964)	(0.2615)	(215.1563)				
EHC	-	0.5947	5992.2102	378.5194	761.0387	0.1154	0.6351
	-	(0.2178)	(2563.2997)				
HC	-	-	3262.2630	379.6545	761.3090	0.1388	0.4029
	-	-	(661.1149)				
			Data set 2				
GHC	26.0412	0.1670	0.5759	**293.8255**	**593.6509**	**0.0962**	**0.5938**
	(40.6471)	(0.1346)	(1.9159)				
BHC	41.9366	1.6173	1.1937	294.9065	595.8130	0.1020	0.5189
	(91.5773)	(0.2709)	(2.4699)				
KHC	38.3343	1.7084	1.1251	294.8059	595.6118	0.1030	0.5059
	(87.5243)	(0.3251)	(2.4761)				
EHC	-	0.1283	211.1940	299.3473	602.6947	0.1576	0.0833
	-	(0.1585)	(251.4769)				
HC	-	-	28.3486	306.4299	614.8597	0.1595	0.0771
	-	-	(4.5011)				

Table 2 lists the MLEs and their corresponding standard errors (in parentheses) of the model parameters for data sets 1 and 2. The numerical values of the model selection statistics $-\log\hat{\ell}$, AIC and K-S, and p-values are listed in Table 2. In general, the results from Table 2 indicate that the GHC distribution provides the best fit among the BHC, KHC, EHC and HC models. The histogram of the data sets 1 and 2, and the estimated pdfs and cdfs of the GHC distribution and its competitive models are displayed in Figures 3 and 4. These Figures support the results in Table 2. To compare the GHC distribution with its sub-models, EHC and HC distributions, the LR test is used for both data sets 1 and 2. When comparing the fits between GHC and EHC (HC) for data 1, $w = 4.2902$ ($w = 6.5608$) with p-value=0.0383 (p-value=0.0376). For data 2, $w = 11.0436$ ($w = 12.6044$) with p-value=0.0009 (p-value=0.0018). These values suggest that GHC performs significantly better for both data sets when comparing it with the sub-models EHC and HC distributions.

7. Concluding remarks

In this paper, we propose a generalization of half-Cauchy distribution called the *gamma-half-Cauchy* distribution. We study some properties of gamma-half Cauchy distribution including quantile function, moments, mean deviations and Shannon entropy. The maximum likelihood method is used for estimating the model parameters and the observed information matrix is analytically derived. We fit the gamma-half-Cauchy to two real data sets to demonstrate its usefulness. The new model provides consistently better fit than other competing models.



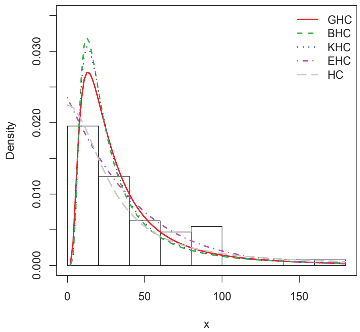
(a) Estimated pdfs (b) Estimated cdfs

Figure 3. Plots of the estimated pdfs and cdfs of the GHC, BHC, KHC and EHC models for data set 1.

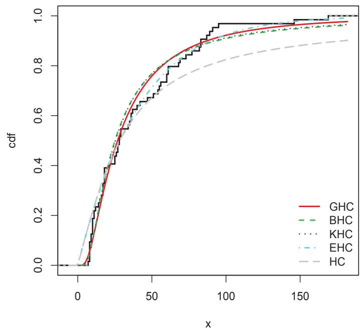
Appendix A

The observed information matrix for the parameter vector $\Theta = (\alpha, \beta, \sigma)^\top$ is given by

$$J(\Theta) = -\frac{\partial^2 \ell(\Theta)}{\partial\,\Theta\,\partial\,\Theta^\top} = -\begin{pmatrix} J_{\alpha\alpha} & J_{\alpha\beta} & J_{\alpha\sigma} \\ . & J_{\beta\beta} & J_{\beta\sigma} \\ . & . & J_{\sigma\sigma} \end{pmatrix},$$



(a) Estimated pdfs



(b) Estimated cdfs

Figure 4. Plots of the estimated pdfs and cdfs of the GHC, BHC, KHC and EHC models for data set 2.

whose elements are

$$
\begin{aligned}
J_{\alpha\alpha} &= -n\,\psi'(\alpha)\ , \\
J_{\alpha\beta} &= -\frac{n}{\beta}\ , \\
J_{\alpha\sigma} &= \frac{2}{\pi\,\sigma^2}\sum_{i=1}^{n}\left\{\frac{x_i\,\tan^{-1\prime}(x_i/\sigma)\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]^{-1}}{-\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]}\right\}\ , \\
J_{\beta\beta} &= \frac{n\alpha}{\beta}+\frac{2}{\beta^3}\sum_{i=1}^{n}\log\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]\ , \\
J_{\beta\sigma} &= -\frac{2}{\pi\,\sigma^2\beta^2}\ \sum_{i=1}^{n}\left\{\frac{x_i\,\tan^{-1\prime}(x_i/\sigma)}{\left[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\right]}\right\}
\end{aligned}
$$

$$
\begin{aligned}
J_{\sigma\sigma} \quad = \quad & \frac{2}{\sigma^2} + \sum_{i=1}^{n} \left\{ \frac{4\,x_i^4}{\sigma^6\,[1+(x_i/\sigma)^2]^2} - \frac{6\,x_i^2}{\sigma^4\,[1+(x_i/\sigma)^2]} \right\} \\
& - \left(\tfrac{1}{\beta} - 1\right) \sum_{i=1}^{n} \left\{ \frac{4\,x_i\,\tan^{-1\prime}(x_i/\sigma)}{\pi\,\sigma^3\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]} \right. \\
& \left. + \frac{4\,x_i^2\,\tan^{-1\prime}(x_i/\sigma)^2}{\pi^2\,\sigma^4\left\{\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\right\}^2} + \frac{2\,x_i^2\,\tan^{-1\prime\prime}(x_i/\sigma)}{\pi\,\sigma^4\left\{\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\right\}} \right\} \\
& - (\alpha - 1) \\
& \sum_{i=1}^{n} \left\{ \frac{4\,x_i\,\tan^{-1\prime}(x_i/\sigma)}{\pi\,\sigma^3\,\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\Big(1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\Big)} \right. \\
& + \frac{4\,x_i^2\,\tan^{-1\prime}(x_i/\sigma)^2}{\pi^2\,\sigma^4\,\log\left\{\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\right\}^2\left\{\Big(1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\Big)\right\}^2} \\
& + \frac{4\,x_i^2\,\tan^{-1\prime}(x_i/\sigma)^2}{\pi^2\,\sigma^4\,\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\left\{\Big(1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\Big)\right\}^2} \\
& \left. + \frac{2\,x_i^2\,\tan^{-1\prime\prime}(x_i/\sigma)}{\pi\,\sigma^4\,\log\big[1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\big]\Big(1-2\,\pi^{-1}\,\tan^{-1}(x_i/\sigma)\Big)} \right\},
\end{aligned}
$$

where $\psi(\alpha) = \frac{\partial \log \Gamma(\alpha)}{\partial\alpha} = \frac{\Gamma(\alpha)'}{\Gamma(\alpha)}$ is the polygamma function and $\psi'(\alpha) = \frac{\partial^2 \log \Gamma(\alpha)}{(\partial\alpha)^2} = \frac{\partial\,\psi(\alpha)}{\partial\,\alpha}$ is the trigamma function.

Appendix B

The first data set are: 1460, 4050, 3570, 2060, 1300, 1390, 1720, 6280, 1360, 7440, 5320, 1400, 3240, 2710, 4520, 4840, 8320, 13900, 71500, 6250, 2260, 318, 1330, 970, 1920, 15100, 2870, 20600, 3810, 726, 7500, 7170, 2000, 829, 17300, 4740, 13400, 2940, 5660.

The second data set were reported by professor Jim Irish and can be obtained at http://www.statsci.org/data/oz/kiama.html. The data are: 83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 29, 54, 91, 8, 17, 55, 10, 35,47, 77, 36, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82, 29, 8, 60, 61, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9, 12.

Acknowledgments

The authors would like to thank the Editor and the two referees for careful reading and for comments which greatly improved the paper.

References

[1] Abramowitz, M. and Stegun, A.I. *Handbook of Mathematical Functions* (Dover, New York 1968)

[2] Akinsete, A., Famoye, F. and Lee, C. *The beta-Pareto distribution*, Statistics **42**, 547–563, 2008.

[3] Alzaatreh, A., Famoye, F. and Lee, C. *Gamma-Pareto distribution and its applications*, Journal Modern Applied Statistical Methods **11**, 78–94, 2012.

[4] Alzaatreh, A., Famoye, F. and Lee, C. *The gamma-normal family of distribution: Properties and applications*, Computational Statistics and Data Analysis **69**,67–80, 2014.

[5] Alzaatreh, A. and Ghosh, I. *A study of the gamma-Pareto (IV) distribution and its applications*, Communications in Statistics–Theory and Methods **45**, 636–654, 2016.

[6] Alzaatreh, A., Ghosh, I. and Said, H. *On the gamma-logistic distribution*, Journal Modern Applied Statistical Methods **13**, 55–70, 2014.

[7] Alzaatreh, A., Lee, C. and Famoye, F. *A new method for generating families of distributions*, Metron **71**, 63–79, 2013.

[8] Alzaatreh, A., Lee, C. and Famoye, F. *T-normal family of distribution: A new approach to generalize the normal distribution*, Journal of Statistical Distribution and Applications **1**, Article 16, 2014.

[9] Cordeiro, G.M. and Lemonte, A.J. *The beta-half Cauchy distribution*, Journal of probability and Statistics Article ID 904705, 18 pages, 2011.

[10] da Silva, R.V., de Andrade, T.A.N., Maciel, D.B.M., Campos, R.P.S. and Cordeiro, G.M. *A new lifetime model: The gamma extended Fréchet distribution*, Journal of Statistical Theory and Applications **12**, 39–54, 2013.

[11] Ghosh, I. *The Kumaraswamy-half Cauchy distribution: Properties and applications*, Journal of Statistical Theory and Applications **13**, 122–134, 2014.

[12] Gradshteyn, I.S. and Ryzhik, I.M. *Table of Integrals, Series and Products*, Sixth edition (Academic Press, San Diego, 2000).

[13] Jacob, E. and Jayakumar, K. *On half Cauchy distribution and process*, International Journal of Statistics and Mathematics **3**, 77–81, 2013.

[14] Lee, C., Famoye, F. and Alzaatreh, A. *Methods for generating families of continuous distribution in the recent decades*, Wiley Inerdisciplinary Reviews: Computational Statistics **5**, 219–238, 2013.

[15] Mudholkar, G.S. and Hutson, A.D. *The exponentiated Weibull family: Some properties and a flood data application*, Communications in Statistics–Theory and Methods **25**, 3059–3083, 1996.

[16] Paradis, E., Baillie, S.R. and Sutherland, W.J. *Modeling large-scale dispersal distances*, Ecological Modeling **151**, 279–292, 2002.

[17] Pinho, L.G., Cordeiro, G.M. and Nobre, J. *The Harris extended exponential distribution*, Communications in Statistics–Theory and Methods **44**, 3486–3502, 2015.

[18] Shaw, M.W. *Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance*, Proceedings Royal Socociety of London **B25**, 243–248, 1995.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1161 – 1173

Kaplan-Meier estimator in competing risk contexts

Pablo Martínez-Camblor* †‡

Abstract

Survival analysis has become in a common procedure in biomedical researches. Conventionally, the well-known nonparametric Kaplan-Meier (KM) estimator is used in order to approximate the real survivor curve. However, in competing risk contexts where more than one failure cause compete to occur and only one of them is of interest, the direct use of the Kaplan-Meier statistic does not perform correctly and, in order to obtain a good estimation, it must be adapted. In this work, via Monte Carlo simulations, the author explores the behavior of the Kaplan-Meier estimator in a competing risk context. In addition, differences between KM and multiple decrement methods are pointed out. Finally, a real-data problem is used in order to illustrate the situation.

Keywords: Competing risks, Kaplan-Meier estimator, Multiple decrement, Survival Analysis.

2000 AMS Classification: 62N01

Received : 11.02.2015 *Accepted* : 13.06.2015 *Doi* : 10.15672/HJMS.20157111086

1. Introduction

Conventionally, survival analysis is devoted to the study of data where the response of interest is the time required for certain (studied) event, which inevitably happens, to occur. Main particularities of these studies are: *i)* on one hand, the distribution of time is often strongly asymmetric and usual parametric models based on the normal law do not perform adequately and, *ii)* the researcher frequently does not have a complete knowledge on the time to event for each subject included in the study; he/she knows that the event does not occur in a period of time but he/she does not know how long the event is needing to occur. These situations are frequently repeated in the nature; perhaps the bio-sanitary (the study of time to death in patients with some particular disease) is one of the most known fields. Of course, there exists a vast literature about

*Hospital Universitario de Asturias (HUCA), Email: pablomc@ficyt.es

†Universidad Autonoma de Chile, Chile

‡Pablo Martínez-Camblor. Dpto. Estadística e IO y DM, Facultad de Ciencias, c/ Calvo Sotelo s/n, 33005 Oviedo, Asturies, Spain.

statistical survival methods, among all, we want to remark the indispensable monograph of Kalbfleish and Prentice (2002).

Let T be the non-negative random variable representing the failure time of interest, as is well-known, in survival analysis, there are mainly three different ways to specify its distribution (see, for instance, Kalbfleish and Prentice (2002)): the survivor function, the probability density function, and the hazard function. The survivor function stands for the probability that the event occurs after a fixed value of time, t, that is,

$$S(t) = \mathcal{P}\{T > t\}, \qquad 0 \leq t < \infty. \tag{1.1}$$

Note that if F denotes the standard cumulative distribution function (CDF) for the random variable T, $S(t) = 1 - F(t)$ $(0 \leq t < \infty)$. Directly, when T is an absolutely continuous variable, the probability density function (PDF) is defined in the standard form,

$$f(t) = d[1 - S(t)]/dt = dF(t)/dt, \qquad 0 \leq t < \infty. \tag{1.2}$$

Obviously, it holds $S(t) = \int_t^\infty f(u)du$. Finally, the hazard function stands for the rate of that the event occurs instantaneously after the time t when it is known that it does not happen before t; that is,

$$\begin{aligned} \lambda(t) &= \lim_{h \to 0^+} \mathcal{P}\{T < t + h | T \geq t\}/h \\ &= f(t)/S(t) = -d\log(S(t))/dt, \qquad 0 \leq t < \infty. \end{aligned} \tag{1.3}$$

Integrating with respect to t and taking into account that $S(0) = 1$, it holds the equality

$$S(t) = \exp\left\{-\int_0^t \lambda(u)du\right\} = \exp\{-\Lambda(t)\}, \qquad 0 \leq t < \infty. \tag{1.4}$$

where $\Lambda(t) = \int_0^t \lambda(u)du$ is known as the *cumulative hazard* function. Standard analysis of survival data usually includes the non-parametric Kaplan-Meier (KM) estimator (Kaplan and Meier (1958) for the survivor curve estimation and the semi-parametric proportional hazard Cox regression (Cox (1972)) in order to explore possible covariate effects.

Under the usual assumption of independence between time to event and censoring time, the KM estimator has really good properties (in the Section 2, some properties of the KM estimator are pointed out); in addition, it has a direct and simple probabilistic interpretation. However, when the studied event not necessarily happens; i.e., there exists one (or more) event which is incompatible with the studied one, the KM estimator overestimated the probability that the event happens. In practice, these situations are really frequent; for instance, when the studied variable is the time to recurrence of some disease; obviously, death without recurrence makes not possible the disease relapses or, when the researcher is interested in the time to death by a particular cause; the death for other cause is, logically, not compatible with the considered event. In this work, the author explores the survival curve estimation in the *competing risk* setting. Particularly, the advantages of using the multiple decrement (MD) estimator (Aalen (1978)) are investigated via Monte Carlo simulations (Section 4). From a real problem dataset, in Section 5, the differences between the KM and the MD estimators are pointed out; particularly, the distribution of the time-free of leukemia in patients with myelodysplasia is analyzed. Finally, in Section 6, the author presents his conclusions.

2. The Kaplan-Meier estimator

The well-known Kaplan-Meier or product-limit estimator was proposed in 1958 in one of the most (or the most, depending on the consulted source) cited and popular statistical paper (Kaplan and Meier (1958)). In that work, the authors proposed a non-parametric method for the estimation of the cumulative distribution function from incomplete observations. The standard mathematical formulation is as follows: let $\boldsymbol{T} = \{T_1, \ldots, T_N\}$ be the times to event and let $\boldsymbol{C} = \{C_1, \ldots, C_N\}$ be the censor times, let F and G be the CDFs for the time to event and the censor time, respectively. The observed times are $\boldsymbol{Z} = \{Z_1, \ldots, Z_N\}$ where $Z_j = \min\{T_j, C_j\}$ $(1 \leq i \leq N)$. In addition, it is also known what time is really observed; i.e., the final available information are the pairs $\{(Z_1, \delta_1), \ldots, (Z_N, \delta_N)\}$, where $\delta_j = I_{T_j}(Z_j)$ (takes the value 1 if the time to event is observed and 0 otherwise). Then, the KM estimator for the survivor function is defined by

$$\hat{S}_N(t) = \prod_{j=1}^{N} \left\{ 1 - \frac{\delta_{(j)} \cdot I_{(-\infty,t]}(Z_{(j)})}{N - j + 1} \right\}, \tag{2.1}$$

where for $j \in 1 \ldots, N$, the pairs $(Z_{(j)}, \delta_{(j)})$ satisfy that $Z_{(1)} \leq Z_{(2)} \leq \cdots \leq Z_{(N)}$. In this context, the Kaplan-Meier is the maximum-likelihood estimator. In addition, their properties have been deeply studied; asymptotic normality can be derived from the work of Csörgő (1996) in which, under usual and mild assumptions, the so-called *Hungarian embeddings* (Komlós, Major and Tusnády (1975)) and the law of the iterated logarithm are generalized to the random right censorship case. Although some alternative methods have been proposed (see, for instance Peto et al. (1975) or Simon and Lee (1982)), the variance of the KM estimator is usually approximated from the Greenwood's formula (Greenwood (1926)),

$$\mathbb{V}[\hat{S}_N(t)] = \sum_{j=1}^{N} \frac{\delta_{(j)} \cdot I_{(-\infty,t]}(Z_{(j)})}{N - \sum_{j=1}^{N} I_{(-\infty,t]}(Z_{(j)})}.$$

On the other hand, Bitouzé et al. (1999) provided a Dvoretzky-Kiefer-Wolfowitz type inequality for the Kaplan-Meier estimator; in particular, they established that there exists an absolute constant K such that,

$$\mathcal{P}\left\{ \sup_{t \in \mathbb{R}} |(1 - G(t)) \cdot (\hat{S}_N(t) - S(t))| > \lambda/\sqrt{n} \right\} \leq 2.5 \cdot e^{-2\lambda^2 + K\lambda},$$

for any positive value λ. Figure 1 depicts the Kaplan-Meier estimation joint with a 95% confidence band (computed using the Greenwood's formula), for the time to death (at left) and the time-free of leukemia (right) for the *Myelodysplastic* dataset. This dataset is from a retrospective study that included high-risk patients reported to the Spanish Group of Myelodysplastic Syndromes Registry (RESMD) between years 2000 and 2013. This data will be used in order to illustrate the considered problem (see Section 6). Anyway, interested readers are referred to Bernal et al. (2015) for additional information about this study. The dataset includes a total of 968 patients (1,273.7 persons-year), 616 of them died during the follow-up. Two-hundred sixty eight patients (27.7%) developed leukemia during the follow-up and 403 died without leukemia. In spite that, of course, these 403 patients are not going to develop leukemia anymore, they are considered as censored for the KM estimator; i.e., their weights are spread among the subjects who are still at risk.

The Kaplan-Meier estimator, like the traditional empirical estimator for the CDF, initially assigns to each sample point a weight of $1/N$ (N stands for the sample size). The main particularity is that, at the time that one subject is censored, KM assumes that its (future) behavior will be similar to the behavior of subjects who are still at risk; therefore, these subjects inherit the weight of the censored subject. Suppose that the minimum time $Z_{(1)}$ corresponds with an event, at this time KM produces a jump of $1/N$, the second time $Z_{(2)}$ is a censored subject; then, subjects who are still in the study

Figure 1. For the Myelodysplastic dataset: at left, Kaplan-Meier estimation for the time to death, at right, Kaplan-Meier estimation for the leukemia-free time. In both panels, 95% confidence bands are included (in gray).

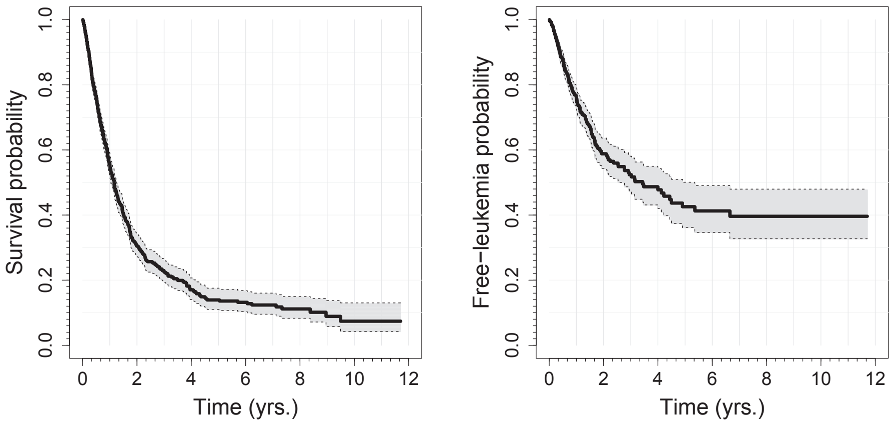


Table 1. Kaplan-Meier contruction for the case described in the manuscript: considered sorted sample is: $\{(Z_{(1)}, 1), (Z_{(2)}, 0), (Z_{(3)}, 1), \cdots, (Z_{(N)}, \delta_{(N)})\}$.

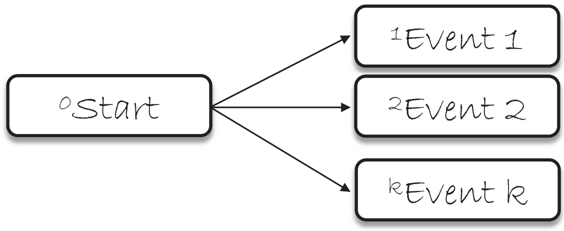
Time	**at risk**	$\boldsymbol{\delta}$	**survival**
0	N		1
$Z_{(1)}$	N	1	$1 - [1/N]$
$Z_{(2)}$	$N-1$	0	$1 - [1/N]$
$Z_{(3)}$	$N-2$	1	$1 - [1/N] - [(1/N) + 1/N(N-2)]$
			....

$(N-2)$ inherit its weight $(1/N)$; therefore the new weight of these $N-2$ subjects will be $1/N + 1/N(N-2)$. Hence, if the third observed time, $Z_{(3)}$, is again an event, KM will produce, at time $Z_{(3)}$, a jump of $1/N + 1/N(N-2)$. Table 1 depicts schematically the KM construction.

3. The competing risk context

There are many real situations in which the event of interest does not always occur; i.e., there exist other events, incompatible with the studied one, which can happen before. The study of the time to death for some particular cause; death for other causes makes not possible the studied event (see, for instance, Verduijn et al. (2011)), the study of the time-free of one particular disease; death for other causes makes impossible the relapse of the considered disease (Boo et al. (2015)), or the study of the transplanted organ

Figure 2. Usual competing risk schema. Transitions from the state ^{0}Start to k the different events are the quantities of interest.



survivor; death of patient does not permit the study of the organ failure (Martínez-Camblor et al. (2015)) are just a few examples of the so-called *competing risk* context. Of course, there exists a vast literature on this topic; see, for instance, Tiatsis (1998) and references therein and Andersen et al. (2002) for the multi-state models approach to competing risk, but our purpose is not to make a revision. Rather, we discuss the problem of the Kaplan-Meier estimator on this context. Figure 2 depicts the standard schema for the competing risk setting; $P_{0,i} = \mathcal{P}_{0,i}(t) = \mathcal{P}\{T_i \leq t\}$ where T_i is the time required to achieve the jth event, with $j \in 1, \ldots, k$ are the main quantities of interest.

In the competing risk contests, the sample must provide information about the observed time and on what event has been really observed. Therefore, $Z_j = \min\{C_j, T_{1,j}, \ldots, T_{k,j}\}$ ($T_{i,j}$ $(1 \leq j \leq N)$ is the time that the subject j would need to achieve the event (state) i) and $\delta_j = i$ with $i \in 0, 1, \ldots, k$ stands for the observed event for the subject j (0 when no event has still happened, i.e., at the final of the study, the subject is still at risk; censored subjects). In order to study the distribution of the time to one particular event (for instance, the ith one, with i taking any value in $1, \ldots, k$), a frequent -and wrong- practice is to consider the rest of the events as censored and then, to estimate the distribution of interest from the Kaplan-Meier estimator. The main issues of this procedure are:

i) Although the independence assumption between the times to event and the time to censoring is plausible, usually, the times to the different events involved in a competing risk setting are strongly dependent. Notice that a patient died before having a relapse, is not going to relapse anymore; the censorship provides information about the considered even. This effect is known as informative censorship.
ii) Due to patients which experiencing a competing event, different to the studied one, are not going to achieve, directly, the event of interest anymore (they are not going to do the transition from ^{0}Start to the studied event), subjects which are still at risk; i.e., those which can still experimenting the event of interest, must not inherit their weights.
iii) In the standard survival analysis, the probability of survival and the probability of event are equivalent quantities ($1 = \mathcal{P}\{T > t\} + \mathcal{P}\{T \leq t\}$). In the competing risk context, there are more involved events and the fact that a subject does not suffer the studied event does not imply that this subject is free of events.

In this context, it holds the equality,

$$\begin{aligned} 1 =& \mathcal{P}\{T > t\} + \mathcal{P}\{T \leq t\} \\ =& \mathcal{P}\{T > t\} + \mathcal{P}\{T \leq t \wedge \delta = 1\} + \cdots + \mathcal{P}\{T \leq t \wedge \delta = k\}. \end{aligned} \tag{3.1}$$

The quantities $\mathcal{P}\{T \leq t \wedge \delta = i\}$ $(1 \leq i \leq k)$ are the *cumulative incidence functions*. However, Andersen, Abildstrom and Rosthoj (2002) claimed that: 'this is, in fact, a rather unfortunate name for this quantity as it may give the incorrect impression that it is a cumulative intensity'. Alternative proposed names are marginal or crude failure probabilities.

4. Multiple decrements method

The nonparametric Kaplan-Meier estimator can be adapted for the competing risk setting in the so-called multiple decrement (MD) method. The considered estimator for the *general transition probabilities* was proposed by Aalen (1978). However, and in spite that different papers have tried to popularize this procedure (see, for instance, Martínez-Camblor et al. (2009) and references therein) it is still little used by practitioners and it is unknown by the physicians. The MD procedure assumes that the probability that two different events occur simultaneously is zero (i.e., $\mathcal{P}\{T_i = T_l\} = 0$ for $1 \leq i \neq l \leq k$). From this proviso, $P_{0,l} = \mathcal{P}\{T \leq t,\, \delta = l\}$ (transition probability between the states 0 and l, $1 \leq l \leq k$) is equivalent to the probability that all the involved times were greater or equal to t and the studied one was exactly t, that is

$$P_{0,l} = \int_0^t S(u)\lambda_l(u)dt = \int_0^t S(u)d\Lambda_l(u), \tag{4.1}$$

where $\lambda_l(u)$ is the hazard function referred to event l. A direct plug-in method using the KM estimator for estimate $S(u)$, and the Nelson-Aalen estimator to estimate the cumulative incidence function, let us to obtain the MD estimator by

$$\mathrm{MD}_l(t) = \sum_{t_j \leq t} \frac{r_{l,j}}{N_j} \prod_{t_i \leq t} \left(1 - \frac{\sum_{l=1}^k r_{l,i}}{N_i}\right), \tag{4.2}$$

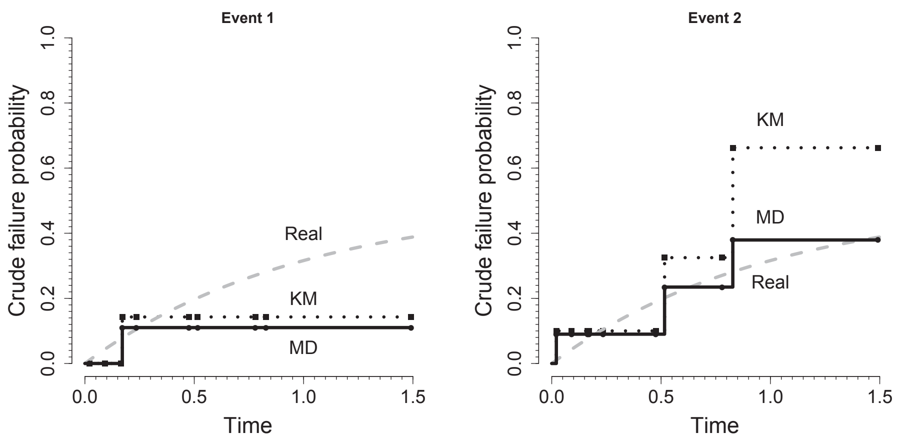
where $r_{l,j}$ $(1 \leq l \leq N)$ and N_j are the number of subjects which have suffered the event l and which were at risk just before of moment t_j $(1 \leq j \leq N)$, respectively. Of course, theoretical properties of the $\mathrm{MD}_l(\cdot)$ estimator have been deeply studied. In Aalen (1978) is proved its uniform consistency (with rate $\log(N) \cdot N^{-1/2}$) and its weak convergence to an adequate Gaussian process (with the usual rate $N^{-1/2}$). Recently, Njamen-Njomen and Ngatchou-Wandji (2014) developed adapted stochastic processes to the Nelson-Aalen and Kaplan-Meier estimators.

In order to illustrate the problem we simulate a three independent times from an exponential law (with mean 1): $T_{1,j}$, $T_{2,j}$ and C_j, in ten subjects $(1 \leq j \leq 10)$. We compute $Z_j = \min\{T_{1,j}, T_{2,j}, C_j\}$ and define $\delta_j = i$, where $i = 1$ if $Z_j = T_{1,j}$, $i = 2$ if $Z_j = T_{2,j}$ and $\delta_j = 0$ if $Z_j = C_j$ $(1 \leq j \leq 10)$. Table 2 depicts the computed estimations by using the KM and the MD methods for the events 1 and 2. Real values (for both events, they are the CDF of an exponential distribution with mean 1) are also reported. Note that in this case, all involved subdistributions are the same. Figure 3 depicts the curves. Since the KM considers censored all events different to the studied one, its 'jumps' are frequently bigger than the MD ones. Obviously, for a fixed point of time t, the MD estimator considers at risk only those subjects which at this time, have not suffer any event.

Table 2. Results for one simulation example of competing risk setting. Sample size was 10 and two different events were simulated ($\delta = 1, 2$; $\delta = 0$ stands for censored data). Direct Kaplan-Meier (KM) and its modification for the multiple decrement (MD). Real values are the same for both considered events.

	Subjects			KM		MD	
Time	**at risk**	δ	**Real**	**1**	**2**	**1**	**2**
0.021	10	2	0.010	0.000	0.100	0.000	0.090
0.091	9	0	0.043	0.000	0.100	0.000	0.090
0.164	8	0	0.076	0.000	0.100	0.000	0.090
0.171	7	1	0.079	0.143	0.100	0.110	0.090
0.235	6	0	0.105	0.143	0.100	0.110	0.090
0.476	5	0	0.189	0.143	0.100	0.110	0.090
0.516	4	2	0.202	0.143	0.325	0.110	0.234
0.779	3	0	0.271	0.143	0.325	0.110	0.234
0.828	2	2	0.281	0.143	0.662	0.110	0.379
1.492	1	0	0.388	0.143	0.662	0.110	0.379

Figure 3. Referred to the data shown in Table 2. At left, real (gray), KM and MD estimations for the event 1. At right, real (gray), KM and MD estimations for the event 2.



5. Monte Carlo simulation study

In order to study the behavior of the direct Kaplan-Meier (KM) and the Multiple decrement (MD) estimators on the competing risk setting, a Monte Carlo simulation study was carried out. The time of studied event, $T_1 = \exp\{D_1\}$, where D_1 was drawn from a normal distribution with mean μ (values of $-1/2$ and $1/2$ were considered) and variance one; the time to the competing risk event, $T_2 = \exp\{D_2\}$, with D_2 generated

Table 3. Mean $\pm$ standard deviation of the 1,000 Monte Carlo iterations for the quantity $100 \cdot \tau^{-1} \cdot \int_0^\tau |\hat{S}(t) - S(t)|dt$ where $S(t)$ is the real subdistribution function and $\hat{S}(t)$ its estimation based on KM and on MD estimators and τ is the maximum observed time for $\mu = -1/2$.

		$\rho = 0.0$		$\rho = 0.25$	
N	c	KM	MD	KM	MD
50	10	7.05 ± 2.39	2.10 ± 1.39	5.90 ± 2.26	1.79 ± 1.19
	-1/4	7.88 ± 3.72	3.85 ± 2.20	6.85 ± 3.40	3.55 ± 1.99
	-1/2	7.99 ± 3.87	4.19 ± 2.37	7.15 ± 3.67	4.11 ± 2.28
250	10	5.55 ± 1.09	0.62 ± 0.41	4.35 ± 0.96	0.48 ± 0.33
	-1/4	6.67 ± 1.52	1.23 ± 0.74	5.67 ± 1.42	1.15 ± 0.65
	-1/2	6.86 ± 1.85	1.46 ± 0.90	5.76 ± 1.73	1.37 ± 0.79
1000	10	4.60 ± 0.81	0.23 ± 0.17	3.47 ± 0.72	0.16 ± 0.11
	-1/4	6.12 ± 0.91	0.48 ± 0.29	5.01 ± 0.81	0.45 ± 0.26
	-1/2	6.27 ± 0.97	0.57 ± 0.33	5.24 ± 0.90	0.54 ± 0.31
		$\rho = 0.75$		$\rho = -0.50$	
50	10	3.17 ± 1.44	1.17 ± 0.81	9.84 ± 2.90	3.12 ± 1.85
	-1/4	4.69 ± 2.67	3.17 ± 1.81	10.03 ± 3.99	4.43 ± 2.48
	-1/2	5.44 ± 3.17	3.72 ± 1.98	10.16 ± 4.47	5.03 ± 2.71
250	10	2.11 ± 0.62	0.29 ± 0.21	8.79 ± 1.14	1.05 ± 0.68
	-1/4	3.24 ± 1.08	1.00 ± 0.59	9.23 ± 1.70	1.53 ± 0.91
	-1/2	3.52 ± 1.31	1.24 ± 0.66	9.36 ± 2.04	1.82 ± 1.03
1000	10	1.55 ± 0.38	0.10 ± 0.07	8.12 ± 0.78	0.42 ± 0.27
	-1/4	2.91 ± 0.59	0.39 ± 0.22	8.99 ± 0.94	0.66 ± 0.38
	-1/2	2.96 ± 0.69	0.49 ± 0.26	8.98 ± 0.99	0.76 ± 0.44

from a standard normal distribution and $\mathbb{E}[D_1 \cdot D_2] = \rho$ (values of 0, 1/4, 1/2 and 3/4 were considered). Finally, the censoring time, $C = exp\{N\}$, where N was drawn, independently, from a normal distribution with mean c (values of 10, -1/4 and -1/2 were considered) and variance one. The (simulated) observed data were the pairs (Z, δ) where $Z = \min\{C, T_1, T_2\}$ and $\delta = i$ ($i = 0$ if $Z = C$, $i = 1$ if $Z = T_1$, and $i = 2$ if $Z = T_2$). Mean $\pm$ standard deviation of the average error, $100 \cdot \tau^{-1} \int_0^\tau |\hat{S}(t) - S(t)|dt$ with $\tau = \max_{1 \leq j \leq N} Z_j$ based on 1,000 Monte Carlo iterations are reported (N stands for the sample size, $S(t)$ denotes the real subdistribution function and $\hat{S}(t)$ its estimation).

Table 3 depicts the observed results when $\mu = -1/2$. In this case, the probability that the considered event happens is: $\mathcal{P}\{T_1 < T_2\} = 0.638$, 0.658, 0.761 and 0.611 for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively. The expected censorship percentages were 0% ($c = 10$); 32.6%, 34.5%, 40.1% and 28.3% ($c = -1/4$) for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively; and 39.7%, 41.9%, 47.2% and 37.5% ($c = -1/2$) for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively. The MD method clearly obtained better results than KM.

Table 4 shows the coverage percentages and mean±sd (standard deviations below 0.00 were denoted by 0.01) of the length of the 95% symmetric confidence intervals (computed by using the naive bootstrap method) for the subdistribution function at times $t = 1/2$ and $t = 1$ using the KM and MD estimators. Observed results endorses the previous obstained ones: KM is not an estimator for the subdistribution function, especially, for larger censorship percentages. The DM estimator works adequately although it shows

Table 4. Coverage percentages and mean±sd (standard deviations below 0.00 were denoted by 0.01) of the length of the 95% symmetric confidence intervals (computed by using the naive bootstrap method with 200 iterations) for the subdistribution function at times $t = 1/2$ and $t = 1$ using the KM and MD estimators when $\mu = -1/2$.

$\rho = 0.0$		$t = 1/2$				$t = 1$			
N	c	KM		MD		KM		MD	
50	10	87.3%	0.289 ± 0.02	91.6%	0.237 ± 0.02	55.0%	0.312 ± 0.04	87.3%	0.234 ± 0.02
	-1/2	90.8%	0.348 ± 0.04	91.4%	0.265 ± 0.02	75.4%	0.452 ± 0.10	85.1%	0.266 ± 0.02
1000	10	20.0%	0.065 ± 0.01	92.9%	0.059 ± 0.01	5.1%	0.070 ± 0.01	93.6%	0.061 ± 0.01
	-1/2	37.1%	0.077 ± 0.01	96.0%	0.069 ± 0.01	3.2%	0.105 ± 0.01	94.9%	0.080 ± 0.01
$\rho = 0.75$		$t = 1/2$				$t = 1$			
50	10	92.4%	0.275 ± 0.02	92.8%	0.238 ± 0.02	77.4%	0.285 ± 0.03	91.2%	0.233 ± 0.02
	-1/2	91.9%	0.325 ± 0.03	92.3%	0.266 ± 0.02	83.3%	0.426 ± 0.08	89.5%	0.869 ± 0.03
1000	10	71.2%	0.062 ± 0.01	93.9%	0.059 ± 0.01	1.4%	0.065 ± 0.01	93.0%	0.961 ± 0.01
	-1/2	76.6%	0.073 ± 0.01	95.1%	0.069 ± 0.01	1.7%	0.095 ± 0.01	93.5%	0.083 ± 0.01

itself a little bit unconservative for the largest censorship percentage ($c = -1/2$ and $t = 1$).

Table 5 is similar to Table 3 for $\mu = 1/2$. In this case, the probability that the considered event happens is: $\mathcal{P}\{T_1 < T_2\} = 0.361$, 0.341, 0.239 and 0.387 for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively. The expected censorship percentages were 0% (for $c = 10$); for $c = -1/4$, approximately 47.2%, 49.3%, 54.3% and 43.5% for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively; and for $c = -1/2$, 54.8%, 56.7%, 61.2% and 51.5% for $\rho = 0$, 1/4, 3/4 and $-1/2$, respectively. The observed results were similar to the ones observed in the Table 3. Notice that, due to, in this case, the effect of the competing event was higher ($\mathcal{P}\{T_1 < T_2\} < 1/2$), the difference between the MD and the KM methods was bigger.

Finally, Table 6 is similar to Table 4 when $\mu = 1/2$. Although the KM estimator obtained better results, observed results are in the same way to the previous one and endorse the conclusions.

6. Real-world problem: the Myelodysplastic data

As has been claimed above, competing risk appears frequently in biomedicine researches, in fact, it is more a rule than an exception. The study of a specific cause of death and the time-free of disease are, probably, the most repeated examples. The main objectives of this section are the estimation of the time-free of leukemia and the time to death without leukemia in a cohort of patients with Myelodysplastic syndromes. The *Myelodysplastic* data was used with this goal. This dataset has been previously introduced in the Section 2 and were collected by the Spanish Group of Myelodysplastic Syndromes Registry (RESMD). Remember that a total of 968 patients (1,273.7 persons-year) were finally included in the study. There were 603 males (62.3%) and 365 females (37.7%); the median age at diagnosis was of 72.8 (ranged between 63.5 and 79.1) years. Two-hundred sixty eight patients (27.7%) developed leukemia during the follow-up and 403 died without leukemia. Figure 4 depicts a flowchart for the Myelodysplastic data. Interested readers are referred to Bernal et al. (2015) for complete information about the cohort and the problem.

By using the KM estimator and assuming as censored those events different to the studied one, the median time for developing leukemia was 3.37 years and, during the

Table 5. Mean $\pm$ standard deviation of the 1,000 Monte Carlo iterations for the quantity $100 \cdot \tau^{-1} \cdot \int_0^\tau |\hat{S}(t) - S(t)| dt$ where $S(t)$ is the real subdistribution function and $\hat{S}(t)$ its estimation based on KM and on MD estimators and τ is the maximum observed time for $\mu = 1/2$.

		$\rho = 0.0$		$\rho = 0.25$	
N	c	**KM**	**DM**	**KM**	**MD**
50	10	6.03 ± 1.96	1.16 ± 0.81	5.12 ± 1.82	0.96 ± 0.70
	-1/4	6.30 ± 3.57	2.67 ± 1.72	5.93 ± 3.18	2.21 ± 1.34
	-1/2	6.38 ± 3.68	2.78 ± 1.75	5.80 ± 3.48	2.52 ± 1.41
250	10	5.29 ± 0.99	0.34 ± 0.24	4.34 ± 0.88	0.28 ± 0.20
	-1/4	5.99 ± 1.77	0.92 ± 0.55	5.06 ± 1.67	0.83 ± 0.48
	-1/2	5.86 ± 2.08	1.06 ± 0.61	5.01 ± 2.01	1.02 ± 0.61
1000	10	4.58 ± 0.71	0.13 ± 0.09	3.60 ± 0.66	0.09 ± 0.07
	-1/4	5.92 ± 0.99	0.38 ± 0.23	5.04 ± 0.94	0.35 ± 0.21
	-1/2	5.77 ± 1.14	0.48 ± 0.27	5.07 ± 1.11	0.42 ± 0.25
		$\rho = 0.75$		$\rho = -0.50$	
50	10	3.41 ± 1.49	0.56 ± 0.43	8.38 ± 2.19	1.68 ± 1.08
	-1/4	4.05 ± 2.59	1.71 ± 1.00	7.76 ± 3.55	2.80 ± 1.59
	-1/2	4.20 ± 2.90	1.99 ± 1.23	7.78 ± 3.79	3.08 ± 1.70
250	10	2.76 ± 0.72	0.16 ± 0.12	6.43 ± 1.01	0.45 ± 0.32
	-1/4	3.35 ± 1.59	0.63 ± 0.38	7.93 ± 1.03	0.59 ± 0.40
	-1/2	3.30 ± 1.80	0.80 ± 0.47	7.82 ± 2.13	1.25 ± 0.71
1000	10	2.25 ± 0.51	0.05 ± 0.04	7.43 ± 0.73	0.23 ± 0.15
	-1/4	3.27 ± 0.94	0.26 ± 0.15	8.07 ± 0.92	0.48 ± 0.29
	-1/2	3.26 ± 1.07	0.34 ± 0.20	7.98 ± 1.04	0.54 ± 0.30

Table 6. Coverage percentages and mean±sd (standard deviations below 0.00 were denoted by 0.01) of the length of the 95% symmetric confidence intervals (computed by using the naive bootstrap method with 200 iterations) for the subdistribution function at times $t = 1/2$ and $t = 1$ using the KM and MD estimators when $\mu = 1/2$.

$\rho = 0.0$		$t = 1/2$				$t = 1$			
N	c	**KM**		**MD**		**KM**		**MD**	
50	10	92.0%	0.188 ± 0.04	92.3%	0.150 ± 0.03	76.2%	0.315 ± 0.04	91.8%	0.203 ± 0.02
	-1/2	91.2%	0.223 ± 0.07	90.6%	0.168 ± 0.05	87.3%	0.490 ± 0.14	89.4%	0.243 ± 0.04
1000	10	64.7%	0.043 ± 0.01	94.2%	0.037 ± 0.01	0.2%	0.070 ± 0.01	94.0%	0.051 ± 0.01
	-1/2	79.1%	0.051 ± 0.01	95.2%	0.044 ± 0.01	5.5%	0.107 ± 0.01	95.1%	0.071 ± 0.01
$\rho = 0.75$		$t = 1/2$				$t = 1$			
50	10	92.7%	0.137 ± 0.05	92.6%	0.110 ± 0.04	88.9%	0.240 ± 0.05	91.6%	0.159 ± 0.03
	-1/2	87.1%	0.149 ± 0.08	89.6%	0.116 ± 0.06	91.2%	0.344 ± 0.16	90.1%	0.189 ± 0.07
1000	10	81.9%	0.031 ± 0.01	92.6%	0.027 ± 0.01	11.4%	0.053 ± 0.01	95.1%	0.039 ± 0.01
	-1/2	85.5%	0.037 ± 0.01	94.9%	0.032 ± 0.01	39.6%	0.080 ± 2.26	94.0%	0.055 ± 0.01

follow-up, the estimated percentage of leukemia was 60.4%, while this estimation was only the 34.9% with the MD method (because this percentage does not lead the 50%, it is not possible to estimate the median time). In the same way, the median time to direct death (without developing leukemia) was 1.67 years when it was estimated by using the

Figure 4. Flowchart for the Myelodysplastic data.

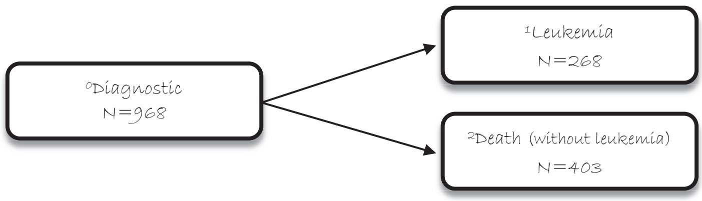
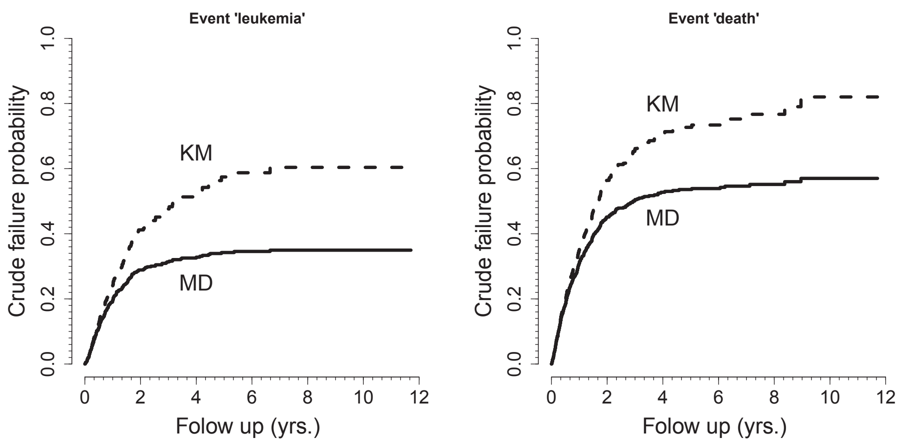


Figure 5. Crude failure probabilities computed by the KM and MD estimators for the time-free of leukemia, at left, and the time to direct death (without a previous leukemia), at right for the Myelodysplastic data.



KM estimator and 2.93 years when the MD method is employed. Figure 5 depicts the crude failure probabilities computed by the KM and MD estimators for the time-free of leukemia, at left, and the time to direct death (without a previous leukemia), at right; also called transition probabilities from the state 0 to 1 and 0 to 2, respectively.

It is worth to make note that the sum of the two KM estimations can take values larger than 1. In the considered problem, for $t \geq 3$, it does.

7. Main conclusions

Even when there exists a number of papers (see, for instance, the works of Putter et al. (2007) or Martínez-Camblor et al. (2015) among many others) trying to avoid the existing gap between theoretical and practical backgrounds, the advances in the statistical methodology are still far from the methods commonly used by practitioners. In addition physicians and basic investigators are usually reluctant to apply in their studies new statistical techniques even when they may be more appropriate to deal with the problem at hand. Multi-state and, particularity, competing risk methods are examples of this situation; in spite of these techniques are the appropriate ones in order to study complex survival schemes, direct Kaplan-Meier and Cox regression are still the used methodologies even when some of the necessary assumptions are violated.

This paper considered the Kaplan-Meier estimator behavior in the competing risk setting. Monte Carlo simulations show that the direct use of the KM estimator produces serious mistakes in those scenarios where the probability of the competing event is high. However, in this context, the MD procedure works fine. In particular, under usual and mild conditions, it is an asymptotically unbiased estimator for the subdistribution functions (see, Kalb fleisch and Prentice (2002)). In addition, and in spite of MD is not include in most popular software, this procedure is easy to implement from the KM outcomes. In addition, several specific and friendly `R packages` [18] which are freely available in the CRAN (`http://cran.r-project.org/web/package`) have been developed with this goal; for example, Meira-Machado and Roca-Pardiñas (2011) describe the `p3state.msm` package and give a complete revision about previously existing software.

Finally, it is worth to remark that friendly statistical packages make easy the data analysis process. Particularly, most of the commercial software includes routines which perform Kaplan-Meier estimations and proportional hazard Cox models. However, using these techniques without checking (and, of course, knowing) conditions required for their correct performing, can produce erroneous conclusions. Remark that, in the practical problem considered, differences between the estimations provided by the KM and the MD methods were beyond ten percent.

Acknowledgments

The author is really grateful to the Associated Editor and to the three anonymous referees for their comments and suggestions which can improve the present work. In addition, he is grateful with the Spanish Group of Myelodysplastic Syndromes Registry (RESMD) and, in special, to Teresa Bernal for the permission to use the Myelodysplastic data.

References

[1] Aalen, O. *Nonparametric estimation of partial transition probabilities in multiple decrement models*, Annals of Statistics **6** (3), 534-545, 1978.

[2] Andersen, P.K., Abildstrom, S.Z. and Rosthoj, S. *Competing risks as a multi-state model*, Statistical Methods in Medical Research **11**, 203-215, 2002.

[3] Bitouzé, D., Laurent, B. and Massart, C. *A Dvoretzky-Kiefer-Wolfowitz type inequality for the Kaplan-Meier estimator*, Ann. Inst. Henri Poincaré **35** (6), 735-763, 1999.

[4] Bernal, T., Martínez-Camblor, P., Sánchez-García, J., de Paz, R., Luño, E., Nomdedeu, B., Ardanaz, M.T., Pedro, C., Amigo, M.L., Xicoy, B., del Cañizo, C., Tormo, M., Bargay, J., Valcárcel, D., Brunet, S., Benlloch, L. and Sanz, G. *Effectiveness of azacitidine in unselected high-risk myelodysplastic syndromes: Results from the Spanish Registry*, Leukemia, `doi: 10.1038/leu`, 2015.

[5] de Boo, L., Pintilie, M., Yip, P., Baniel, J., Fleshner, N. and Margel, D. *Time from first detectable PSA following radical prostatectomy to biochemical recurrence: A competing risk analysis*, Canadian Urology Association Journal **9** (1-2), 14-21, 2015.

[6] Cox, D.R., *Regression models and life table (with discussion)*, Journal of the Royal Statitical Society, Ser. B, **34**, 187-220, 1972.

[7] Csörgő, S. *Universal Gaussian approximations under random censorship*, Annals of Statistics **24** (6), 2744-2778, 1996.

[8] Greenwood, M. *The errors of sampling of the survivorship tables*, Reports on Public Health and Statistical Subjects 33, 1-26, 1926.

[9] Kalbfleisch, R. and Prentice, R.L. *The statistical analysis of failure time data*, Wiley, New Jersey, 2002.

[10] Kaplan, E.L. and Meier, P. *Nonparametric estimation from imcomplete observations*, Journal of the American Statistical Association **53**, 457-481, 1958.

[11] Komlós, J., Major, P. and Tusnády, G. *An approximation of partial sums of independent rvs and the sample df. I*, Wahrsch verw Gebiete 32, 111-131, 1975.

[12] Martínez-Camblor, P., Larrañaga, N., Sarasqueta, C., Mitxelena, M.J. and Basterretxea, M. *That deadly and pink corporeity: Analysis of disease-free survival analysis breast cancer in Gipuzkoa (Spain) in the presence of competing risks*, Gaceta Sanitaria **23** (6), 554-557, 2009.

[13] Martínez-Camblor, P., de Uña-Álvarez, J. and Díaz Corte, C. *Expanded renal transplantation: A multi-state model approach*, Journal of Applied Statistics, `DOI:10.1080/02664763.2015.1043866`, 2015.

[14] Meira-Machado, L. and Roca-Pardiñas, J. *p3state.msm: Analyzing survival data from an illness-death mode*, Journal of Statistical Software **38** (3), 1-18, 2011.

[15] Njamen-Njomen, D. and Ngatchou-Wandji, J. *Nelson-Aalen and Kaplan-Meier estimators in competing risks*, Applied Mathematics **5**, 465-47, 2014.

[16] Peto, R., Pike, M.C., Armitage, P., Breslow, N.E., Cox, D.R., Howard, S.V., Mantel, N., McPherson, K., Peto, J. and Smith, P.G. *Design and analysis of randomised clinical trials requiring prolonged observation of each patient, II*, British Jorunal of Cancer, 35, 1-39, 1977.

[17] Putter, H., Fiocco, M. and Geskus, R. *Tutorial in biostatistics: Competing risks and multi-state models*, Statistics in Medicine 26, 2389-2430, 2007.

[18] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, 2007.

[19] Simon, R. and Lee, Y.J. *Nonparametric confidence limits for survival probabilities and median survival time*, Cancer Treatment Reports 66, 37-42, 1985.

[20] Tiatsis, AA. Competing risks, *Encyclopedia of Biostatistics*, 1, 824-834, 1988.

[21] Verduijn, M., Grootendorst, D.C., Dekker, F.W., Jager, K.J. and le Cessie, S. *The analysis of competing events like cause-specific mortalitybeware of the KaplanMeier method*, Nephrology Dialisys Transplantation 26, 5661, 2011.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1175 – 1202

The beta odd log-logistic generalized family of distributions

Gauss M. Cordeiro*, Morad Alizadeh†, M. H. Tahir‡,§, M. Mansoor¶, Marcelo Bourguignon‖ and G. G. Hamedani**

Abstract

We introduce a new family of continuous models called the *beta odd log-logistic generalized family* of distributions. We study some of its mathematical properties. Its density function can be symmetrical, left-skewed, right-skewed, reversed-J, unimodal and bimodal shaped, and has constant, increasing, decreasing, upside-down bathtub and J-shaped hazard rates. Five special models are discussed. We obtain explicit expressions for the moments, quantile function, moment generating function, mean deviations, order statistics, Rényi entropy and Shannon entropy. We discuss simulation issues, estimation by the method of maximum likelihood, and the method of minimum spacing distance estimator. We illustrate the importance of the family by means of two applications to real data sets.

Keywords: Beta-G family, characterizations, exponential distribution, generalized family, log-logistic distribution, maximum likelihood, method of minimum spacing distance.

2000 AMS Classification: 60E05; 62E10; 62N05.

Received : 13.12.2014 *Accepted* : 30.06.2015 *Doi* : 10.15672/HJMS.20157311545

*Department of Statistics, Federal University of Pernambuco, 50740-540, Recife, PE, Brazil, Email: gauss@de.ufpe.br, gausscordeiro@gmail.com

†Department of Statistics, Persian Gulf University of Bushehr, Bushehr 751691-3798, Iran, Email: moradalizadeh78@gmail.com

‡Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mtahir.stat@gmail.com, mht@iub.edu.pk

§Corresponding Author.

¶Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mansoor.abbasi143@gmail.com

‖Department of Statistics, Federal University of Pernambuco, 50740-540, Recife, PE, Brazil, Email: m.p.bourguignon@gmail.com

**Department of Mathematics, Statistics and Computer Science, Marquette University, WI 53201-1881, Milwaukee, USA, Email: g.hamedani@mu.edu

1. Introduction

There has been an increased interest in defining new generators or generalized (G) classes of univariate continuous distributions by adding shape parameter(s) to a baseline model. The extended distributions have attracted several statisticians to develop new models because the computational and analytical facilities available in programming softwares like R, Maple and Mathematica can easily tackle the problems involved in computing special functions in these extended models. Several mathematical properties of the extended distributions may be easily explored using mixture forms of the exponentiated-G ("exp-G" for short) distributions. The addition of parameter(s) has been proved useful in exploring skewness and tail properties, and also for improving the goodness-of-fit of the generated family. The well-known generators are the following: beta-G by Eugene et al. [15] and Jones [29], Kumaraswamy-G (Kw-G) by Cordeiro and de Castro [10], McDonald-G (Mc-G) by Alexander et al. [1], gamma-G type 1 by Zografos and Balakrishnan [53] and Amini et al. [6], gamma-G type 2 by Ristić and Balakrishnan [44], odd-gamma-G type 3 by Torabi and Montazari [50], logistic-G by Torabi and Montazari [51], odd exponentiated generalized (odd exp-G) by Cordeiro et al. [12], transformed-transformer (T-X) (Weibull-X and gamma-X) by Alzaatreh et al. [3], exponentiated T-X by Alzaghal et al. [5], odd Weibull-G by Bourguignon et al. [7], exponentiated half-logistic by Cordeiro et al. [13], logistic-X by Tahir et al. [47], T-X{Y}-quantile based approach by Aljarrah et al. [2] and T-R{Y} by Alzaatreh et al. [4].

This paper is organized as follows. In Section 2, we define the *beta odd log-logistic generalized* (BOLL-G) family. Some of its special cases are presented in Section 3. In Section 4, we derive some of its mathematical properties such as the asymptotics, shapes of the density and hazard rate functions, mixture representation for the density, quantile function (qf), moments, moment generating function (mgf), mean deviations, explicit expressions for the Rényi and Shannon entropies and order statistics. Section 5 deals with some characterizations of the new family. Estimation of the model parameters and simulation using maximum likelihood and the method of minimum spacing distance are discussed in Section 6. In Section 7, we illustrate the importance of the new family by means of two applications to real data. The paper is concluded in Section 8.

2. The odd log-logistic and beta odd log-logistic families

The log-logistic (LL) distribution is widely used in practice and it is an alternative to the log-normal model since it presents a hazard rate function (hrf) that increases, reaches a peak after some finite period and then declines gradually. Its properties make the distribution an attractive alternative to the log-normal and Weibull models in the analysis of survival data. If T has a logistic distribution, then $Z = \mathrm{e}^{\mathrm{T}}$ has the LL distribution. Unlike the more commonly used Weibull distribution, the LL distribution has a non-monotonic hrf which makes it suitable for modeling cancer survival data.

The *odd log-logistic (OLL) family* of distributions was originally developed by Gleaton and Lynch [18, 19]; they called this family the *generalized log-logistic (GLL) family.* They showed that:

– the set of GLL transformations form an Abelian group with the binary operation of composition;

– the transformation group partitions the set of all lifetime distributions into equivalence classes, so that any two distributions in an equivalence class are related through a GLL transformation;

– either every distribution in an equivalence class has a moment generating function, or none does;

– every distribution in an equivalence class has the same number of moments;

– each equivalence class is linearly ordered according to the transformation parameter, with larger values of this parameter corresponding to smaller dispersion of the distribution about the common class median; and

– within an equivalence class, the Kullback-Leibler information is an increasing function of the ratio of the transformation parameters.

In addition, Gleaton and Rahman obtained results about the distributions of the MLE's of the parameters of the distribution. Gleaton and Rahman [20, 21] showed that for distributions generated from either a 2-parameter Weibull distribution or a 2-parameter inverse Gaussian distribution by a GLL transformation, the joint maximum likelihood estimators of the parameters are asymptotically normal and efficient, provided the GLL transformation parameter exceeds 3.

Given a continuous baseline cumulative distribution function (cdf) $G(x;\boldsymbol{\xi})$ with a parameter vector $\boldsymbol{\xi}$, the cdf of the OLL-G family (by integrating the LL density function with an additional shape parameter $c>0$) is given by

$$F_{\text{OLL-G}}(x)=\int_0^{G(x;\xi)/\overline{G}(x;\xi)}\frac{c\,t^{c-1}}{(1+t^c)^2}\,dt=\frac{G(x;\boldsymbol{\xi})^c}{G(x;\boldsymbol{\xi})^c+\overline{G}(x;\boldsymbol{\xi})^c}. \tag{2.1}$$

If $c>1$, the hrf of the OLL-G random variable is unimodal and when $c=1$ it decreases monotonically. The fact that its cdf has closed-form is particularly important for analysis of survival data with censoring.

We can write

$$c=\frac{\log\left[F(x;\boldsymbol{\xi})/\overline{F}(x;\boldsymbol{\xi})\right]}{\log\left[G(x;\boldsymbol{\xi})/\overline{G}(x;\boldsymbol{\xi})\right]}\qquad\text{and}\qquad\overline{G}(x;\boldsymbol{\xi})=1-G(x;\boldsymbol{\xi}).$$

Here, the parameter c represents the quotient of the log-odds ratio for the generated and baseline distributions.

The probability density function (pdf) corresponding to (2.1) is

$$f_{\text{OLL-G}}(x)=\frac{c\,g(x;\boldsymbol{\xi})\,\left\{G(x;\boldsymbol{\xi})\,\overline{G}(x;\boldsymbol{\xi})\right\}^{c-1}}{\left\{G(x;\boldsymbol{\xi})^c+\overline{G}(x;\boldsymbol{\xi})^c\right\}^2}. \tag{2.2}$$

In this paper, we propose a new extension of the OLL-G family. Based on a baseline cdf $G(x;\boldsymbol{\xi})$ depending on a parameter vector $\boldsymbol{\xi}$, survival function $\overline{G}(x;\boldsymbol{\xi})=1-G(x;\boldsymbol{\xi})$ and pdf $g(x;\boldsymbol{\xi})$, we define the cdf of the BOLL-G family of distributions (for $x\in\mathbb{R}$) by

$$F(x)=F(x;a,b,c,\boldsymbol{\xi})=\frac{1}{B(a,b)}\,B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c+\overline{G}(x;\xi)^c};a,b\right), \tag{2.3}$$

where $a>0$, $b>0$ and $c>0$ are three additional shape parameters, $B(z;a,b)=\int_0^z w^{a-1}(1-w)^{b-1}\mathrm{d}w$ is the incomplete beta function, $B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function and $\Gamma(a)=\int_0^\infty t^{a-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$ is the gamma function. We also adopt the notation $I_z(a,b)=B(z;a,b)/B(a,b)$.

The pdf and hrf corresponding to (2.3) are, respectively, given by

$$f(x)\;=\;f(x;a,b,c,\boldsymbol{\xi})=\frac{c\,g(x;\boldsymbol{\xi})\,G(x;\boldsymbol{\xi})^{ac-1}\overline{G}(x;\boldsymbol{\xi})^{bc-1}}{B(a,b)\left\{G(x;\boldsymbol{\xi})^c+\overline{G}(x;\boldsymbol{\xi})^c\right\}^{a+b}} \tag{2.4}$$

and

$$h(x)\;=\;\frac{c\,g(x;\boldsymbol{\xi})\,G(x;\boldsymbol{\xi})^{ac-1}\overline{G}(x;\boldsymbol{\xi})^{bc-1}}{\left\{G(x;\boldsymbol{\xi})^c+\overline{G}(x;\boldsymbol{\xi})^c\right\}^{a+b}\left\{B(a,b)-B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c+\overline{G}(x;\xi)^c};a,b\right)\right\}}. \tag{2.5}$$

Clearly, if we take $G(x) = x/(1+x)$, equation (2.3) becomes the beta log-logistic distribution. The family (2.4) contains some sub-families listed in Table 1. The baseline G distribution is a basic exemplar of (2.4) when $a = b = c = 1$. Hereafter, $X \sim$ BOLL-G$(a, b, c, \boldsymbol{\xi})$ denotes a random variable having density function (2.4). We can omit the parameters in the pdf's and cdf's.

Table 1: Some special models of the BOLL-G family.

a	b	c	$G(x)$	Reduced distribution
-	-	1	$G(x)$	Beta-G family (Eugene et al. [15])
1	1	-	$G(x)$	Odd log-logistic family (Gleaton and Lynch[19])
1	-	1	$G(x)$	Proportional hazard rate family (Gupta et al. [26])
-	1	1	$G(x)$	Proportional reversed hazard rate family (Gupta and Gupta [25])
1	1	1	$G(x)$	$G(x)$

The BOLL-G family can easily be simulated by inverting (2.3) as follows: if V has a beta (a, b) distribution, then the random variable X can be obtained from the baseline qf, say $Q_G(u) = G^{-1}(u)$. In fact, the random variable

$$X = Q_G\Big[\frac{V^{\frac{1}{c}}}{V^{\frac{1}{c}} + (1-V)^{\frac{1}{c}}}\Big] \tag{2.6}$$

has density function (2.4).

3. Some special models

Here, we present some special models of the BOLL-G family.

3.1. The BOLL-exponential (BOLL-E) distribution. The pdf and cdf of the exponential distribution with scale parameter $\alpha > 0$ are given by $g(x;\alpha) = \alpha\,\mathrm{e}^{-\alpha\,x}$ and $G(x;\alpha) = 1 - \mathrm{e}^{-\alpha\,x}$, respectively. Inserting these expressions in (2.4) gives the BOLL-E pdf

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

3.2. The BOLL-normal (BOLL-N) distribution. The BOLL-N distribution is defined from (2.4) by taking $G(x;\boldsymbol{\xi}) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ and $g(x;\boldsymbol{\xi}) = \sigma^{-1}\,\phi\left(\frac{x-\mu}{\sigma}\right)$ for the cdf and pdf of the normal distribution with parameters μ and σ^2, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, and $\boldsymbol{\xi} = (\mu, \sigma^2)$. The BOLL-N pdf is given by

$$f(x; a, b, c, \mu, \sigma^2) = \frac{c\,\phi(\frac{x-\mu}{\sigma})\;\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{ac-1}\;\left\{1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{bc-1}}{\sigma B(a,b)\left[\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{c} + \left\{1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{c}\right]^{a+b}}, \tag{3.1}$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter and $\sigma > 0$ is a scale parameter.

We can denote by $X \sim$ BOLL-N(a, b, c, μ, σ^2) a random variable having pdf (3.1).

3.3. The BOLL-Lomax (BOLL-Lx) distribution. The pdf and cdf of the Lomax distribution with scale parameter $\beta > 0$ and shape parameter $\alpha > 0$ are given by

$g(x;\alpha,\beta) = (\alpha/\beta)\,[1+(x/\beta)]^{-(\alpha+1)}$ and $G(x;\alpha,\beta) = 1-[1+(x/\beta)]^{-\alpha}$, respectively. The BOLL-Lx pdf follows by inserting these expressions in (2.4) as

$$f(x;a,b,c,\alpha,\beta) = \frac{\frac{c\,\alpha}{\beta}\left\{1+\left(\frac{x}{\beta}\right)\right\}^{-(\alpha+1)}\left\{1+\left(\frac{x}{\beta}\right)\right\}^{-\alpha(ac-1)}}{B(a,b)\left[\left\{1-\left[1+\left(\frac{x}{\beta}\right)\right]^{-\alpha}\right\}^{c}+\left\{1+\left(\frac{x}{\beta}\right)\right\}^{-\alpha\,c}\right]^{a+b}}.$$

3.4. The BOLL-Weibull (BOLL-W) distribution. The pdf and cdf of the Weibull distribution with scale parameter $\alpha > 0$ and shape parameter $\beta > 0$ are given by $g(x;\alpha,\beta) = \alpha\beta x^{\beta-1}\,\mathrm{e}^{-\alpha x^{\beta}}$ and $G(x;\alpha,\beta) = 1-\mathrm{e}^{-\alpha x^{\beta}}$, respectively. Inserting these expressions in (2.4) yields the BOLL-W pdf

$$f(x;a,b,c,\alpha,\beta) = \frac{c\,\alpha\,\beta\,x^{\beta-1}\,\mathrm{e}^{-b\,c\,\alpha\,x^{\beta}}\left\{1-\mathrm{e}^{-\alpha\,x^{\beta}}\right\}^{ac-1}}{B(a,b)\left[\left\{1-\mathrm{e}^{-\alpha\,x^{\beta}}\right\}^{c}+\left\{\mathrm{e}^{-\alpha\,x^{\beta}}\right\}^{c}\right]^{a+b}}.$$

3.5. The BOLL-Gamma (BOLL-Ga) distribution. Consider the gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, where the pdf and cdf (for $x > 0$) are given by

$$g(x;\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\,x^{\alpha-1}\,\mathrm{e}^{-\beta x} \quad \text{and} \quad G(x;\alpha,\beta) = \frac{\gamma(\alpha,\beta\,x)}{\Gamma(\alpha)},$$

where $\gamma(\alpha,\beta\,x) = \int_0^{\beta\,x} t^{\alpha-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$ is the incomplete gamma function. Inserting these expressions in equation (2.4), the BOLL-Ga density function follows as

$$f(x;a,b,c,\alpha,\beta) = \frac{c\,\beta^{\alpha}\,x^{\alpha-1}\,\mathrm{e}^{-\beta x}\left\{\frac{\gamma(\alpha,\beta\,x)}{\Gamma(\alpha)}\right\}^{ac-1}\left\{1-\frac{\gamma(\alpha,\beta\,x)}{\Gamma(\alpha)}\right\}^{bc-1}}{\Gamma(\alpha)\,B(a,b)\left[\left\{\frac{\gamma(\alpha,\beta\,x)}{\Gamma(\alpha)}\right\}^{c}+\left\{1-\frac{\gamma(\alpha,\beta\,x)}{\Gamma(\alpha)}\right\}^{c}\right]^{a+b}}.$$

In Figures 1 and 2, we display some plots of the pdf and hrf of the BOLL-E, BOLL-N and BOLL-Lx distributions for selected parameter values. Figure 1 reveals that the BOLL-E, BOLL-N and BOLL-Lx densities generate various shapes such as symmetrical, left-skewed, right-skewed, reversed-J, unimodal and bimodal. Also, Figure 2 shows that these models can produce hazard rate shapes such as constant, increasing, decreasing, J and upside-down bathtub. This fact implies that the BOLL-G family can be very useful for fitting data sets with various shapes.

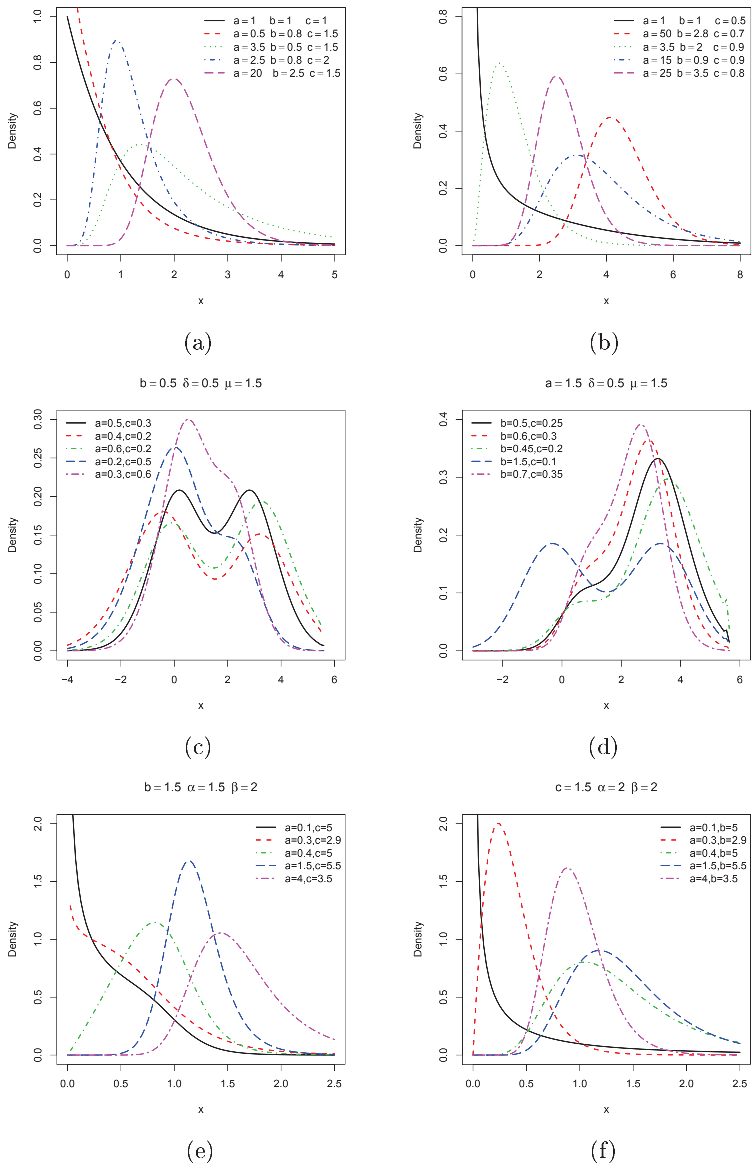


Figure 1. Density plots (a)-(b) of the BOLL-E model, (c)-(d) of the BOLL-N model and (e)-(f) of the BOLL-Lx model.

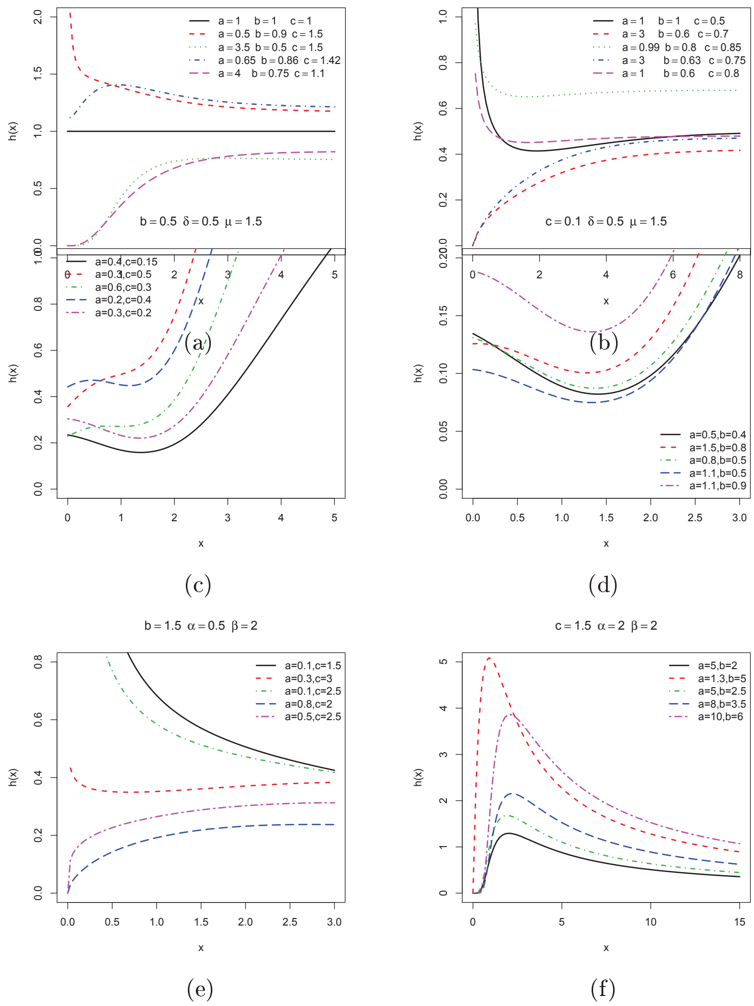


Figure 2. Hazard rate plots (a)-(b) of the BOLL-E model, (c)-(d) of the BOLL-N model and (e)-(f) of the BOLL-Lx model.

4. Mathematical properties

Here, we present some mathematical properties of the new family of distributions.

4.1. Asymptotics and shapes. The asymptotes of equations (2.3), (2.4) and (2.5) as $x \to 0$ and $x \to \infty$ are given by

$$
\begin{aligned}
&F(x) \sim I_{G(x)^c}(a,b) \quad \text{as } x \to 0,\\
&1 - F(x) \sim I_{\bar{G}(x)^c}(b,a) \quad as \;\; x \to \infty,\\
&f(x) \sim \frac{c}{B(a,b)}\, g(x) G(x)^{a\,c-1} \quad \text{as } x \to 0,\\
&f(x) \sim \frac{c}{B(a,b)}\, g(x) \bar{G}(x)^{b\,c-1} \quad \text{as } x \to \infty,\\
&h(x) \sim \frac{c\, g(x) G(x)^{a\,c-1}}{1 - I_{G(x)^c}(a,b)} \quad \text{as } x \to 0,\\
&h(x) \sim \frac{c\, g(x) \bar{G}(x)^{b\,c-1}}{I_{\bar{G}(x)^c}(b,a)} \quad \text{as } x \to \infty.
\end{aligned}
$$

The shapes of the density and hazard rate functions can be described analytically. The critical points of the BOLL-G density function are the roots of the equation:

$$
\frac{g'(x)}{g(x)} + (ac-1)\frac{g(x)}{G(x)} + (1-bc)\frac{g(x)}{\overline{G}(x)} - c(a+b)g(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c} = 0. \tag{4.1}
$$

There may be more than one root to (4.1). Let $\lambda(x) = d^2 \log[f(x)]/d\,x^2$. We have

$$
\begin{aligned}
\lambda(x) \;=\; & \frac{g''(x)g(x) - [g'(x)]^2}{g(x)^2} + (ac-1)\frac{g'(x)G(x) - g(x)^2}{G(x)^2}\\
+\; & (1-bc)\frac{g'(x)\overline{G}(x) + g(x)^2}{\overline{G}(x)^2} - c(a+b)g'(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c}\\
-\; & c(c-1)(a+b)g(x)^2\,\frac{G(x)^{c-2} + \overline{G}(x)^{c-2}}{G(x)^c + \overline{G}(x)^c}\\
-\; & (a+b)\left\{cg(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c}\right\}^2.
\end{aligned}
$$

If $x = x_0$ is a root of (4.1) then it corresponds to a local maximum if $\lambda(x) > 0$ for all $x < x_0$ and $\lambda(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\lambda(x) < 0$ for all $x < x_0$ and $\lambda(x) > 0$ for all $x > x_0$. It refers to a point of inflexion if either $\lambda(x) > 0$ for all $x \neq x_0$ or $\lambda(x) < 0$ for all $x \neq x_0$.

The critical points of the hrf $h(x)$ are obtained from the equation

$$
\begin{aligned}
& \frac{g'(x)}{g(x)} + (ac-1)\frac{g(x)}{G(x)} + (1-bc)\frac{g(x)}{\overline{G}(x)} - c(a+b)g(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c}\\
& + \frac{cg(x)G(x)^{ac-1}\overline{G}(x)^{bc-1}}{B(a,b)\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b}\left\{1 - I_{\frac{G(x)^c}{\overline{G}(x)^c + G(x)^c}}(a,b)\right\}} = 0.
\end{aligned} \tag{4.2}
$$

There may be more than one root to (4.2). Let $\tau(x) = d^2 \log[h(x)]/dx^2$. We have

$$\begin{aligned}
\tau(x) &= \frac{g''(x)g(x) - [g'(x)]^2}{g(x)^2} + (ac-1)\frac{g'(x)G(x) - g(x)^2}{G(x)^2} \\
&+ (1-bc)\frac{g'(x)\overline{G}(x) + g(x)^2}{\overline{G}(x)^2} - c(a+b)g'(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c} \\
&+ c(c-1)(a+b)g(x)^2\,\frac{G(x)^{c-2} + \overline{G}(x)^{c-2}}{G(x)^c + \overline{G}(x)^c} \\
&- (a+b)\left\{cg(x)\frac{G(x)^{c-1} - \overline{G}(x)^{c-1}}{G(x)^c + \overline{G}(x)^c}\right\}^2 \\
&+ \frac{cg'(x)G(x)^{ac-1}\overline{G}(x)^{bc-1}}{\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b}\left\{B(a,b) - B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c + \overline{G}(x;\xi)^c}; a, b\right)\right\}} \\
&+ \frac{c(ac-1)g(x)^2G(x)^{ac-2}\overline{G}(x)^{bc-1}}{\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b}\left\{B(a,b) - B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c + \overline{G}(x;\xi)^c}; a, b\right)\right\}}
\end{aligned}$$

$$\begin{aligned}
&+ \frac{c(bc-1)g(x)^2G(x)^{ac-1}\overline{G}(x)^{bc-2}}{\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b}\left\{B(a,b) - B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c + \overline{G}(x;\xi)^c}; a, b\right)\right\}} \\
&- \frac{c^2(a+b)^2g(x)G(x)^{ac-1}\overline{G}(x)^{bc-1}\left\{G(x)^{c-1} - \overline{G}(x)^{c-1}\right\}}{\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b+1}\left\{B(a,b) - B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c + \overline{G}(x;\xi)^c}; a, b\right)\right\}} \\
&+ \left\{\frac{cg(x)G(x)^{ac-1}\overline{G}(x)^{bc-1}}{\left\{G(x)^c + \overline{G}(x)^c\right\}^{a+b}\left\{B(a,b) - B\left(\frac{G(x;\xi)^c}{G(x;\xi)^c + \overline{G}(x;\xi)^c}; a, b\right)\right\}}\right\}^2.
\end{aligned}$$

If $x = x_0$ is a root of (4.2) then it refers to a local maximum if $\tau(x) > 0$ for all $x < x_0$ and $\tau(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\tau(x) < 0$ for all $x < x_0$ and $\tau(x) > 0$ for all $x > x_0$. It gives an inflexion point if either $\tau(x) > 0$ for all $x \neq x_0$ or $\tau(x) < 0$ for all $x \neq x_0$.

4.2. Useful expansions. For an arbitrary baseline cdf $G(x)$, a random variable Z has the exp-G distribution (see Section 1) with power parameter $c > 0$, say $Z \sim$exp-G(c), if its pdf and cdf are given by $h_c(x) = c\,G(x)^{c-1}\,g(x)$ and $H_c(x) = G(x)^c$, respectively. Some structural properties of the exp-G distributions are studied by Mudholkar and Srivastava [35], Mudholkar et al. [36], Mudholkar and Hutson [34], Gupta et al. [26], Gupta and Kundu [27, 28], Nadarajah and Kotz [39], Nadarajah and Gupta [40, 41] and Nadarajah [37].

We can prove that the cdf (2.3) admits the expansion

$$\begin{aligned}
F(x) &= \sum_{l=0}^{\infty} \frac{(-1)^l}{B(a,b)(a+l)} \binom{b-1}{l} \frac{G(x)^{c(a+l)}}{[G(x)^c + \bar{G}(x)^c]^{a+l}} \\
&= \sum_{l=0}^{\infty} \frac{(-1)^l}{B(a,b)(a+l)} \binom{b-1}{l} \frac{\sum_{k=0}^{\infty} \alpha_k^{(l)}\,G(x)^k}{\sum_{k=0}^{\infty} \beta_k^{(l)}\,G(x)^k}.
\end{aligned}$$

Using the power series for the ratio of two power series, we have

$$F(x) = \sum_{l=0}^{\infty} \frac{(-1)^l}{B(a,b)(a+l)} \binom{b-1}{l} \sum_{k=0}^{\infty} \gamma_k^{(l)}\,G(x)^k,$$

where (for each l) $\alpha_k^{(l)} = a_k(c(a+l))$, $\beta_k^{(l)} = h_k(c, a+l)$, $a_k(c(a+l))$ and $h_k(c, a+l)$ are defined in the Appendix A and $\gamma_k^{(l)}$ is determined recursively as

$$\gamma_k^{(l)} = \gamma_k(a, c) = \frac{1}{\beta_0^{(l)}} \left(\alpha_k^{(l)} - \frac{1}{\beta_0^{(l)}} \sum_{r=1}^{k} \beta_r^{(l)} \, \gamma_{k-r}^{(l)} \right).$$

Then, we have

$$F(x) = \sum_{k=0}^{\infty} b_k \, H_k(x),$$

where

$$b_k = \sum_{l=0}^{\infty} \frac{(-1)^l \, \gamma_k^{(l)}}{B(a,b)\,(a+l)} \binom{b-1}{l}, \tag{4.3}$$

and $H_k(x) = G(x)^k$ denotes the exp-G cdf with power parameter k. So, the density function of X can be expressed as

$$f(x) = f(x; a, b, c, \boldsymbol{\xi}) = \sum_{k=0}^{\infty} b_{k+1} \, h_{k+1}(x; \boldsymbol{\xi}), \tag{4.4}$$

where $h_{k+1}(x) = h_{k+1}(x; \boldsymbol{\xi}) = (k+1) \, g(x; \boldsymbol{\xi}) \, G(x; \boldsymbol{\xi})^k$ denotes the exp-G density function with power parameter $k+1$. Hereafter, a random variable having density function $h_{k+1}(x)$ is denoted by $Y_{k+1} \sim$ exp-G$(k+1)$. Equation (4.4) reveals that the BOLL-G density function is an infinite mixture of exp-G densities. Thus, some mathematical properties of the new model can be obtained directly from those exp-G properties. For example, the ordinary and incomplete moments, and mgf of X can be determined from those quantities of the exp-G distribution.

The formulae derived throughout the paper can be easily handled in most symbolic computation software platforms such as Maple, Mathematica and Matlab. These platforms have currently the ability to deal with analytic expressions of formidable size and complexity. Established explicit expressions to calculate statistical measures can be more efficient than computing them directly by numerical integration. The infinity limit in these sums can be substituted by a large positive integer such as 20 or 30 for most practical purposes.

4.3. Quantile function. The qf of X, say $x = Q(u) = F^{-1}(u)$, can be obtained by inverting (2.3). Let $z = Q_{a,b}(u)$ be the beta qf. Then,

$$x = Q(u) = Q_G \left\{ \frac{[Q_{a,b}(u)]^{\frac{1}{c}}}{[Q_{a,b}(u)]^{\frac{1}{c}} + [1 - Q_{a,b}(u)]^{\frac{1}{c}}} \right\}.$$

It is possible to obtain some expansions for $Q_{a,b}(u)$ from the Wolfram website `http://functions.wolfram.com/06.23.06.0004.01` such as

$$z = Q_{a,b}(u) = \sum_{i=0}^{\infty} e_i \, u^{i/a},$$

where $e_i = [a\,B(a,b)]^{1/a} \, d_i$ and $d_0 = 0$, $d_1 = 1$, $d_2 = (b-1)/(a+1)$,

$$d_3 = \frac{(b-1)\,(a^2 + 3ab - a + 5b - 4)}{2(a+1)^2(a+2)},$$

$$\begin{aligned} d_4 &= (b-1)[a^4 + (6b-1)a^3 + (b+2)(8b-5)a^2 + (33b^2 - 30b + 4)a \\ &+ b(31b-47) + 18]/[3(a+1)^3(a+2)(a+3)], \ldots \end{aligned}$$

The effects of the shape parameters a, b and c on the skewness and kurtosis of X can be based on quantile measures. The Bowley skewness (Kenney and Keeping [30]) is one of the earliest skewness measures defined by the average of the quartiles minus the median, divided by half the interquartile range, namely

$$B = \frac{Q\left(\frac{3}{4}\right) + Q\left(\frac{1}{4}\right) - 2Q\left(\frac{1}{2}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)}.$$

Since only the middle two quartiles are considered and the outer two quartiles are ignored, this adds robustness to the measure. The Moors kurtosis (Moors [33]) is based on octiles

$$M = \frac{Q\left(\frac{3}{8}\right) - Q\left(\frac{1}{8}\right) + Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)}.$$

These measures are less sensitive to outliers and they exist even for distributions without moments.

In Figure 3, we plot the measures B and M for the BOLL-N and BOLL-Lx distributions. The plots indicate the variability of these measures on the shape parameters.

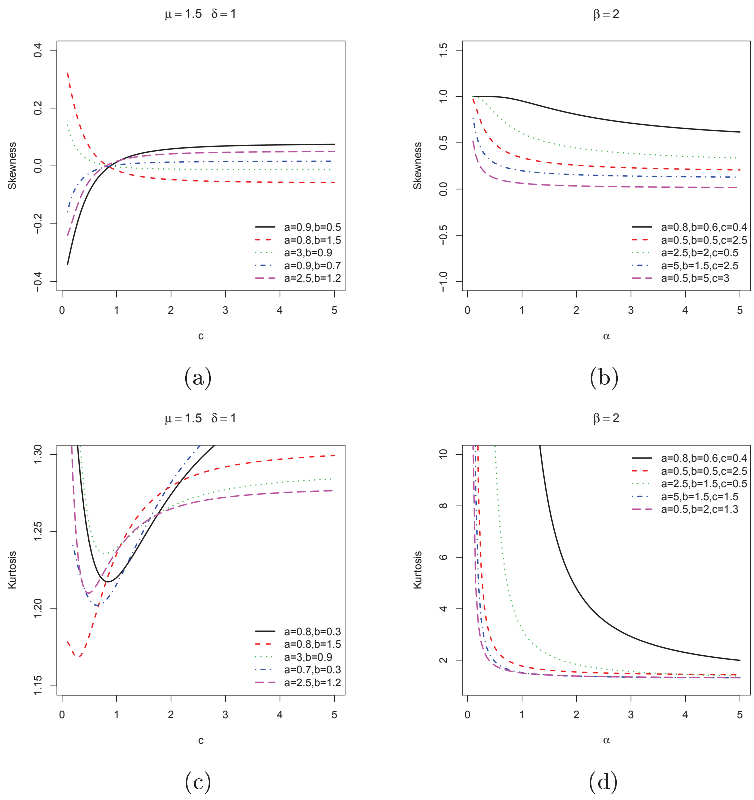


Figure 3. Skewness (a) and (b) and kurtosis (c) and (d) of X based on the quantiles for the BOLL-N and BOLL-Lx distributions, respectively.

4.4. Moments. We assume that Y is a random variable having the baseline cdf $G(x)$. The moments of X can be obtained from the (r,k)th probability weighted moment (PWM) of Y defined by Greenwood et al. [23] as

$$\tau_{r,k} = \mathrm{E}[Y^r\, G(Y)^k] = \int_{-\infty}^{\infty} x^r\, G(x)^k\, g(x) dx.$$

The PWMs are used to derive estimators of the parameters and quantiles of generalized distributions. The method of estimation is formulated by equating the population and sample PWMs. These moments have low variance and no severe biases, and they compare favorably with estimators obtained by maximum likelihood. The maximum likelihood method is adopted in Section 6.1 since it is easier to estimate the BOLL-G parameters because of several computer routines available in widely known softwares. The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used when constructing confidence intervals and regions and also in test statistics.

We can write from equation (4.4)

$$\mu'_r = \mathrm{E}(X^r) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, \tau_{r,k}, \tag{4.5}$$

where $\tau_{r,k} = \int_0^1 Q_G(u)^r\, u^k du$ can be computed at least numerically from any baseline qf.

Thus, the moments of any BOLL-G distribution can be expressed as an infinite weighted sum of the baseline PWMs. We now provide the PWMs for three distributions discussed in Section 3. For the BOLL-N and BOLL-Ga distributions discussed in subsections 3.2 and 3.5, the quantities $\tau_{r,k}$ can be expressed in terms of the Lauricella functions of type A (see Exton [16] and Trott [52]) defined by

$$F_A^{(n)}(a; b_1, \ldots, b_n; c_1, \ldots, c_n; x_1, \ldots, x_n) =$$
$$\sum_{m1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} \frac{(a)_{m_1+\ldots+m_n} (b_1)_{m_1} \ldots (b_n)_{m_n}}{(c_1)_{m_1} \ldots (c_n)_{m_n}} \frac{x_1^{m_1} \ldots x_n^{m_n}}{m_1! \ldots m_n!},$$

where $(a)_i = a(a+1)\ldots(a+i-1)$ is the ascending factorial (with the convention that $(a)_0 = 1$).

In fact, Cordeiro and Nadarajah [11] determined $\tau_{r,k}$ for the standard normal distribution as

$$\tau_{r,k} = 2^{r/2}\, \pi^{-(k+1/2)} \sum_{\substack{l=0 \\ (r+k-l)\ \text{even}}}^{k} \binom{k}{l} 2^{-l}\, \pi^l\, \Gamma\left(\frac{r+k-l+1}{2}\right) \times$$
$$F_A^{(k-l)}\left(\frac{r+k-l+1}{2}; \frac{1}{2}, \ldots, \frac{1}{2}; \frac{3}{2}, \ldots, \frac{3}{2}; -1, \ldots, -1\right).$$

This equation holds when $r+k-l$ is even and it vanishes when $r+k-l$ is odd. So, any BOLL-N moment can be expressed as an infinite weighted linear combination of Lauricella functions of type A.

For the gamma distribution, the quantities $\tau_{r,k}$ can be expressed from equation (9) of Cordeiro and Nadarajah [11] as

$$\tau_{r,k} = \frac{\Gamma(r+(k+1)\alpha)}{\alpha^k\, \beta^r\, \Gamma(\alpha)^{k+1}}\, F_A^{(k)}(r+(k+1)\alpha; \alpha, \ldots, \alpha; \alpha+1, \ldots, \alpha+1, -1, \ldots, -1).$$

Finally, for the BOLL-W distribution, the quantities $\tau_{r,k}$ are given by

$$\tau_{r,k} = \frac{\Gamma(r/\beta+1)}{\alpha^{r/\beta}} \sum_{s=0}^{k} \frac{(-1)^s}{(s+1)^{r/\beta+1}} \binom{k}{s}.$$

4.5. Generating function. Here, we provide two formulae for the mgf $M(s) = E(\mathrm{e}^{s\,X})$ of X. The first formula for $M(s)$ comes from equation (4.4) as

$$M(s) = \sum_{k=0}^{\infty} b_{k+1}\, M_{k+1}(s), \tag{4.6}$$

where $M_{k+1}(s)$ is the exp-G generating function with power parameter $k+1$.

Equation (4.6) can also be expressed as

$$M(s) = \sum_{k=0}^{\infty} (k+1)\; b_{k+1}\, \rho_k(s), \tag{4.7}$$

where the quantity $\rho_k(s) = \int_0^1 \exp\left[s\, Q_G(u)\right] u^k du$ can be computed numerically.

4.6. Mean deviations. Incomplete moments are useful for measuring inequality, for example, the Lorenz and Bonferroni curves and Pietra and Gini measures of inequality all depend upon the incomplete moments of the distribution. The nth incomplete moment of X is defined by $m_n(y) = \int_{-\infty}^{y} x^n\, f(x)dx$. Here, we propose two methods to determine the incomplete moments of the new family. First, the nth incomplete moment of X can be expressed as

$$m_n(y) = \sum_{k=0}^{\infty} b_{k+1} \int_0^{G(y;\,\xi)} Q_G(u)^n\, u^k\, du. \tag{4.8}$$

The integral in (4.8) can be computed at least numerically for most baseline distributions.

The mean deviations about the mean ($\delta_1 = E(|X - \mu_1'|)$) and about the median ($\delta_2 = E(|X - M|)$) of X are given by

$$\delta_1 = 2\mu_1'\, F\left(\mu_1'\right) - 2m_1\left(\mu_1'\right) \qquad \text{and} \qquad \delta_2 = \mu_1' - 2m_1(M), \tag{4.9}$$

respectively, where $M = Q(0.5)$ is the median of X, $\mu_1' = \mathrm{E}(X)$ comes from equation (4.5), $F(\mu_1')$ can easily be calculated from (2.3) and $m_1(z) = \int_{-\infty}^{z} x\, f(x)dx$ is the first incomplete moment.

Next, we provide two alternative ways to compute δ_1 and δ_2. A general equation for $m_1(z)$ can be derived from equation (4.4) as

$$m_1(z) = \sum_{k=0}^{\infty} b_{k+1}\, J_{k+1}(z), \tag{4.10}$$

where

$$J_{k+1}(z) = \int_{-\infty}^{z} x\, h_{k+1}(x)dx.$$

Equation (4.10) is the basic quantity to compute the mean deviations in (4.9). A simple application of (4.10) refers to the BOLL-W model. The exponentiated Weibull density function (for $x > 0$) with power parameter $k+1$, shape parameter α and scale parameter β, is given by

$$h_{k+1}(x) = (k+1)\,\alpha\,\beta^{\alpha}\, x^{\alpha-1}\, \exp\left\{-(\beta x)^{\alpha}\right\}\, \left[1 - \exp\left\{-(\beta x)^{\alpha}\right\}\right]^k,$$

and then

$$J_{k+1}(z) = c\,(k+1)\,\beta^{\alpha}\, \sum_{r=0}^{\infty} (-1)^r \binom{k}{r} \int_0^z x^{\alpha}\, \exp\left\{-(r+1)(\beta x)^{\alpha}\right\} dx.$$

The last integral reduces to the incomplete gamma function and then

$$J_{k+1}(z) = \beta^{-1} \sum_{r=0}^{\infty} \frac{(-1)^r\,(k+1)\binom{k}{r}}{(r+1)^{1+\alpha^{-1}}}\, \gamma\left(1+\alpha^{-1}, (r+1)(\beta z)^{\alpha}\right).$$

A second general formula for $m_1(z)$ can be derived by setting $u = G(x)$ in (4.4)

$$m_1(z) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\; T_k(z),$$

where $T_k(z) = \int_0^{G(z)} Q_G(u)\, u^k du$.

The main application of the first incomplete moment refers to the Bonferroni and Lorenz curves which are very useful in economics, reliability, demography, insurance and medicine. For a given probability π, applications of these equations can be addressed to obtain these curves defined by $B(\pi) = m_1(q)/(\pi\, \mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $q = Q(\pi)$ is calculated from the parent qf.

4.7. Entropies. An entropy is a measure of variation or uncertainty of a random variable X. Two popular entropy measures are the Rényi [43] and Shannon [45]. The Rényi entropy of a random variable with pdf $f(x)$ is defined by

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left(\int_0^{\infty} f^{\gamma}(x) dx\right),$$

for $\gamma > 0$ and $\gamma \neq 1$. The Shannon entropy of a random variable X is given by $I_S = E\left\{-\log\left[f(X)\right]\right\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$. Direct calculation yields

$$\begin{aligned} I_S \;&=\; -\log\left[\frac{c}{B(a,b)}\right] - \mathrm{E}\left\{\log\left[g(X;\boldsymbol{\xi})\right]\right\} + (1-ac)\,\mathrm{E}\left\{\log\left[G(x;\boldsymbol{\xi})\right]\right\} \\ &+\; (1-bc)\,\mathrm{E}\left\{\log\left[\bar{G}(x;\boldsymbol{\xi})\right]\right\} + (a+b)\,\mathrm{E}\left\{\log\left[G(x;\boldsymbol{\xi})^c + \bar{G}(x;\boldsymbol{\xi})^c\right]\right\}. \end{aligned}$$

First, we define and compute

$$\begin{aligned} A(a_1,a_2,a_3;a) &= \int_0^1 \frac{u^{a_1}(1-u)^{a_2}}{\left[u^a + (1-u)^a\right]^{a_3}} du \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{a_2}{i} \int_0^1 \frac{u^{a_1+i}}{\left[u^a + (1-u)^a\right]^{a_3}} du \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{a_2}{i} \int_0^1 \frac{\sum_{k=0}^{\infty} \delta_{1,k} u^k}{\sum_{k=0}^{\infty} \delta_{2,k}\, u^k} du \\ &= \sum_{i=0}^{\infty} (-1)^i \binom{a_2}{i} \int_0^1 \sum_{k=0}^{\infty} \delta_{3,k} u^k \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i\, \delta_{3,k}}{(k+1)} \binom{a_2}{i}, \end{aligned}$$

where $\delta_{1,k} = a_k(a_1+i)$, $\delta_{2,k} = h_k(a, a_3)$ and

$$\delta_{3,k} = \frac{1}{\delta_{2,0}}\left(\delta_{1,k} - \frac{1}{\delta_{2,0}} \sum_{r=1}^{k} \delta_{2,r}\, \delta_{3,k-r}\right).$$

After some algebraic manipulations, we obtain:

4.1. Theorem. *Let X be a random variable with pdf* (2.4)*. Then,*

$$\mathrm{E}\left\{\log\left[G(X)\right]\right\} = \frac{c}{B(a,b)}\, \frac{\partial}{\partial t} A(a\,c+t-1, b\,c-1, a+b; c)\Big|_{t=0},$$

$$\mathrm{E}\left\{\log\left[\bar{G}(X)\right]\right\} = \frac{c}{B(a,b)}\, \frac{\partial}{\partial t} A(a\,c-1, b\,c+t-1, a+b; c)\Big|_{t=0},$$

$$\mathrm{E}\left\{G(x;\boldsymbol{\xi})^a+\bar{G}(X;\boldsymbol{\xi})^a\right\}=\frac{c}{B(a,b)}\,\frac{\partial}{\partial t}A(a\,c-1,b\,c-1,a+b-t;c)\Big|_{t=0}.$$

The simplest formula for the entropy of X is given by

$$\begin{aligned}I_S&=-\log\left[\frac{c}{B(a,b)}\right]-\mathrm{E}\left\{\log\left[g(X;\boldsymbol{\xi})\right]\right\}\\&+\frac{c\,(1-a\,c)}{B(a,b)}\,\frac{\partial}{\partial t}A(a\,c+t-1,b\,c-1,a+b;c)\Big|_{t=0}\\&+\frac{c\,(1-b\,c)}{B(a,b)}\,\frac{\partial}{\partial t}A(a\,c-1,b\,c+t-1,a+b;c)\Big|_{t=0}\\&+\frac{c\,(a+b)}{B(a,b)}\,\frac{\partial}{\partial t}A(a\,c-1,b\,c-1,a+b-t;c)\Big|_{t=0}.\end{aligned}$$

After some algebraic developments, we have an alternative expression for $I_R(\gamma)$:

$$I_R(\gamma)=\frac{\gamma}{1-\gamma}\log\left[\frac{c}{B(a,b)}\right]+\frac{1}{1-\gamma}\log\left[\sum_{i,k=0}^{\infty}t_{i,k}\;\mathrm{E}_{V_k}(g^{\gamma-1}[G^{-1}(Y)])\right].$$

Here, V_k has a beta distribution with parameters $k+1$ and one,

$$t_{i,k}=\frac{(-1)^i\,\gamma_{3,k}(a,b,c,i)}{(k+1)}\binom{c(a-1)}{i},$$

$$\gamma_{1,k}=a_k\Big((a\,c-1)\gamma+i\Big),\ \gamma_{2,k}=h_k\Big(c,(a+b)\gamma\Big)$$

and

$$\gamma_{3,k}=\frac{1}{\gamma_{2,0}}\left(\gamma_{1,k}-\frac{1}{\gamma_{2,0}}\sum_{r=1}^{k}\gamma_{2,r}\gamma_{3,k-r}\right),$$

where $a_k((a\,c-1)\gamma+i)$ and $h_k\Big(c,(a+b)\gamma\Big)$ are defined in equation (8.6) given in Appendix A.

4.8. Order statistics. Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1,\ldots,X_n$ is a random sample from the BOLL-G family of distributions. We can write the density of the ith order statistic, say $X_{i:n}$, as

$$f_{i:n}(x)=\;K\,f(x)\,F^{i-1}(x)\;\{1-F(x)\}^{n-i}=K\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}f(x)\,F(x)^{j+i-1},$$

where $K=n!/[(i-1)!\,(n-i)!]$.

Following similar algebraic developments of Nadarajah et al. [38], we can write the density function of $X_{i:n}$ as

$$f_{i:n}(x)=\sum_{r,k=0}^{\infty}m_{r,k}\,h_{r+k+1}(x),\tag{4.11}$$

where $h_{r+k}(x)$ denotes the exp-G density function with power parameter $r+k+1$ (for $r,k\geq 0$)

$$m_{r,k}=\frac{n!\,(r+1)\,(i-1)!\,b_{r+1}}{(r+k+1)}\sum_{j=0}^{n-i}\frac{(-1)^j\,f_{j+i-1,k}}{(n-i-j)!\,j!},$$

and b_k is defined in equation (4.3). The quantities $f_{j+i-1,k}$ can be obtained recursively by $f_{j+i-1,0} = b_0^{j+i-1}$ and

$$f_{j+i-1,k} = (k\, b_0)^{-1} \sum_{m=1}^{k} [m(j+i) - k]\, b_m\, f_{j+i-1,k-m}, \quad k \geq 1.$$

Equation (4.11) is the main result of this section. It reveals that the pdf of the BOLL-G order statistics is a linear combination of exp-G density functions. So, several mathematical quantities of the BOLL-G order statistics such as ordinary, incomplete and factorial moments, mgf, mean deviations and several others can be determined from those quantities of the exp-G distribution.

5. Characterizations of the new family based on two truncated moments

The problem of characterizing distributions is an important problem which has attracted the attention of many researchers recently. An investigator will, generally, be interested to know if their chosen model fits the requirements of a particular distribution. Hence, one will depend on the characterizations of this distribution which provide conditions under which one can check to see if the underlying distribution is indeed that particular distribution. Various characterizations of distributions have been established in many different directions. In this section, we present characterizations of the BOLL-G distribution based on a simple relationship between two truncated moments. Our characterization results will employ a theorem due to Glänzel [24] (Theorem 5.1, below). The advantage of the characterizations given here is that the cdf F is not required to have a closed-form and is given in terms of an integral whose integrand depends on the solution of a first order differential equation, which can serve as a bridge between probability and differential equation. We believe that other characterizations of the BOLL-G family may not be possible.

5.1. Theorem. *Let $(\Omega, \Sigma, \mathbf{P})$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ ($a = -\infty, b = \infty$ might as well be allowed). Let $X : \Omega \to H$ be a continuous random variable with distribution function $F(x)$ and let q_1 and q_2 be two real functions defined on H such that*

$$\mathbb{E}\left[q_1(X) \,|\, X \geq x\right] = \mathbb{E}\left[q_2(X) \,|\, X \geq x\right]\, \eta(x), \quad x \in H,$$

is defined with some real function η. Consider that q_1, $q_2 \in C^1(H)$, $\eta \in C^2(H)$ and $F(x)$ is twice continuously differentiable and strictly monotone function on the set H. Further, we assume that the equation $q_2 \eta = q_1$ has no real solution in the interior of H. Then, F is uniquely determined by the functions q_1, q_2 and η, particularly

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u)\; q_2(u) - q_1(u)} \right| \; \mathrm{e}^{-s(u)}\, du\,,$$

where the function s is a solution of the differential equation $s' = \frac{\eta'\, q_2}{\eta\, q_2 - q_1}$ and C is a constant chosen to make $\int_H dF = 1$.

We have to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence. In particular, let us assume that there is a sequence $\{X_n\}$ of random variables with distribution functions $\{F_n\}$ such that the functions $q_{1,n}$, $q_{2,n}$ and η_n ($n \in \mathbb{N}$) satisfy the conditions of Theorem 5.1 and let $q_{1,n} \to q_1$, $q_{2,n} \to q_2$ for some continuously differentiable real functions q_1 and q_2. Finally, let X be a random variable with distribution F. Under the condition that $q_{1,n}(X)$ and $q_{2,n}(X)$ are uniformly integrable and the family $\{F_n\}$ is relatively compact, the sequence X_n converges to X in distribution if and only if η_n converges to η, where

$$\eta(x) = \frac{E[q_1(X) | X \geq x]}{E[q_2(X) | X \geq x]}.$$

5.2. Remark. (a) In Theorem 5.1, the interval H need not be closed since the condition is only on the interior of H.

(b) Clearly, Theorem 5.1 can be stated in terms of two functions q_1 and η by taking $q_2(x) = 1$, which will reduce the condition in Theorem 5.1 to $E[q_1(X) | X \geq x] = \eta(x)$. However, adding an extra function will give a lot more flexibility, as far as its application is concerned.

5.3. Proposition. *Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let $q_1(x) = q_2(x)\, G(x;\xi)^{ac}$ and $q_2(x) = \left\{G(x;\xi)^c + \overline{G}(x;\xi)^c\right\}^{-(a+b)} \overline{G}(x;\xi)^{1-bc}$ for $x \in \mathbb{R}$. The pdf of X is* (2.4) *if and only if the function η defined in Theorem 5.1 has the form*

$$\eta(x) = \frac{1}{2}\left[1 + G(x;\xi)^{ac}\right], \quad x \in \mathbb{R}.$$

Proof. If X has pdf (2.4), then

$$[1 - F(x)]\, \mathbf{E}[q_2(X) | X \geq x] = \frac{1}{aB(a,b)}\left[1 - G(x;\xi)^{ac}\right], \quad x \in \mathbb{R}$$

and

$$[1 - F(x)]\, \mathbf{E}[q_1(X) | X \geq x] = \frac{1}{2aB(a,b)}\left[1 - G(x;\xi)^{2ac}\right], \quad x \in \mathbb{R}.$$

Finally,

$$\eta(x)\, q_2(x) - q_1(x) = \frac{1}{2} q_2(x)\left[1 - G(x;\xi)^{ac}\right] > 0, \quad \text{for} \quad x \in \mathbb{R}.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)\, q_2(x)}{\eta(x)\, q_2(x) - q_1(x)} = \frac{a\,c\,g(x)\, G(x;\xi)^{ac-1}}{\left[1 - G(x;\xi)^{ac}\right]}, \quad x \in \mathbb{R},$$

and hence

$$s(x) = -\log\left[1 - G(x;\xi)^{ac}\right], \quad x \in \mathbb{R}.$$

Now, in view of Theorem 5.1, X has pdf (2.4). □

5.4. Corollary. *Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let $q_2(x)$ be as in Proposition 5.3. The pdf of X is (2.4) if and only if there exist functions q_1 and η defined in Theorem 5.1 satisfying the differential equation*

$$\frac{\eta'(x)\, q_2(x)}{\eta(x)\, q_2(x) - q_1(x)} = \frac{a\,c\,g(x)\, G(x;\xi)^{ac-1}}{[1 - G(x;\xi)^{ac}]}, \quad x \in \mathbb{R}.$$

5.5. Remark. (a) The general solution of the differential equation in Corollary 5.4 is

$$\eta(x) = [1 - G(x;\xi)^{ac}]^{-1}\left[-\int a\,c\,g(x)\, G(x;\xi)^{ac-)}\, q_1(x)\, q_2(x)^{-1}\, dx + D\right],$$

for $x \in \mathbb{R}$, where D is a constant. One set of appropriate functions is given in Proposition 5.3 with $D = 1/2$.

(b) Clearly there are other triplets of functions (q_1, q_2, η) satisfying the conditions of Theorem 5.1, e.g.,

$$q_1(x) = q_2(x)\, \overline{G}(x;\xi)^{bc}$$

and

$$q_2(x) = \left[G(x;\xi)^c + \overline{G}(x;\xi)^c\right]^{-(a+b)} G(x;\xi)^{1-ac}, \quad x \in \mathbb{R}.$$

Then, $\eta(x) = \frac{1}{2}\,\overline{G}(x;\xi)^{bc}$ and $s'(x) = \frac{\eta'(x)\,q_2(x)}{\eta(x)\,q_2(x) - q_1(x)} = b\,c\,g(x)\,\overline{G}(x)^{-1}, \quad x \in \mathbb{R}$.

6. Different methods of estimation

Here, we discuss parameter estimation using the methods of maximum likelihood and of minimum spacing distance estimator proposed by Torabi [48].

6.1. Maximum likelihood estimation. We consider the estimation of the unknown parameters of this family from complete samples only by the method of maximum likelihood. Let $x_1, \ldots, x_n$ be observed values from the BOLL-G distribution with parameters a, b, c and $\boldsymbol{\xi}$. Let $\Theta = (a, b, c, \boldsymbol{\xi})^\top$ be the $r \times 1$ parameter vector. The total log-likelihood function for Θ is given by

$$\begin{aligned}\ell_n \quad = \quad & n\,\log(c) - n\log[B(a,b)] + \sum_{i=1}^n \log[g(x_i;\boldsymbol{\xi})] + (ac-1)\sum_{i=1}^n \log[G(x_i;\boldsymbol{\xi})] \\ & + (bc-1)\sum_{i=1}^n \log[\bar{G}(x_i;\boldsymbol{\xi})] - (a+b)\sum_{i=1}^n \log\left\{G(x_i;\boldsymbol{\xi})^c + \bar{G}(x_i;\boldsymbol{\xi})^c\right\}. \end{aligned} \tag{6.1}$$

The log-likelihood function can be maximized either directly by using the R (AdequacyModel or Maxlik) (see R Development Core Team [42]), SAS (PROC NLMIXED), Ox program (sub-routine MaxBFGS) (see Doornik [14]), Limited-Memory quasi-Newton code for bound-constrained optimization (L-BFGS-B) or by solving the nonlinear likelihood equations obtained by differentiating (6.1).

Let $U_n(\Theta) = (\partial\ell_n/\partial a, \partial\ell_n/\partial b, \partial\ell_n/\partial c, \partial\ell_n/\partial\boldsymbol{\xi})^\top$ be the score function. Its components are given by

$$\begin{aligned}
\frac{\partial\ell_n}{\partial a} &= -n\psi(a) + n\psi(a+b) + c\sum_{i=1}^n \log[G(x_i;\boldsymbol{\xi})] - \sum_{i=1}^n \log\left\{G(x_i;\boldsymbol{\xi})^c + \bar{G}(x_i;\boldsymbol{\xi})^c\right\}, \\
\frac{\partial\ell_n}{\partial b} &= -n\psi(b) + n\psi(a+b) + c\sum_{i=1}^n \log[\bar{G}(x_i;\boldsymbol{\xi})] - \sum_{i=1}^n \log\left\{G(x_i;\boldsymbol{\xi})^c + \bar{G}(x_i;\boldsymbol{\xi})^c\right\}, \\
\frac{\partial\ell_n}{\partial c} &= \frac{n}{c} + a\sum_{i=1}^n \log[G(x_i;\boldsymbol{\xi})] + b\sum_{i=1}^n \log[\bar{G}(x_i;\boldsymbol{\xi})] \\
&\quad -(a+b)\sum_{i=1}^n \frac{G(x_i;\boldsymbol{\xi})^c \log[G(x_i;\boldsymbol{\xi})] + \bar{G}(x_i;\boldsymbol{\xi})^c \log[\bar{G}(x_i;\boldsymbol{\xi})]}{G(x_i;\boldsymbol{\xi})^c + \bar{G}(x_i;\boldsymbol{\xi})^c}, \\
\frac{\partial\ell_n}{\partial\boldsymbol{\xi}} &= \sum_{i=1}^n \frac{g(x_i;\boldsymbol{\xi})^{(\xi)}}{g(x_i;\boldsymbol{\xi})} + (ac-1)\sum_{i=1}^n \frac{G(x_i;\boldsymbol{\xi})^{(\xi)}}{G(x_i;\boldsymbol{\xi})} + (1-bc)\sum_{i=1}^n \frac{G(x_i;\boldsymbol{\xi})^{(\xi)}}{\bar{G}(x_i;\boldsymbol{\xi})} \\
&\quad -c(a+b)\sum_{i=1}^n G(x_i;\boldsymbol{\xi})^{(\xi)}\,\frac{G(x_i;\boldsymbol{\xi})^{c-1} - \bar{G}(x_i;\boldsymbol{\xi})^{c-1}}{G(x_i;\boldsymbol{\xi})^c + \bar{G}(x_i;\boldsymbol{\xi})^c},
\end{aligned}$$

where $h^{(\xi)}(\cdot)$ means the derivative of the function h with respect to ξ.

For interval estimation and hypothesis tests, we can use standard likelihood techniques based on the observed information matrix, which can be obtained from the authors upon request.

6.2. Minimum spacing distance estimator (MSDE). Torabi [48] introduced a general method for estimating parameters through spacing called maximum spacing distance estimator (MSDE). Torabi and Bagheri [49] and Torabi and Montazeri [51] used different MSDEs to compare with the MLEs. Here, we used two MSDEs, "minimum spacing absolute distance estimator" (MSADE) and "minimum spacing absolute-log distance estimator" (MSALDE) and compared them with the MLEs of the BOLL-E distribution. For mathematical details, the reader is referred to Torabi and Bagheri [49] and Torabi and Montazeri [51].

Table 2: The AEs, biases and MSEs of the MLEs, MSADEs and MSALDEs of the parameters based on 1,000 simulations of the BOLL-E(2, 1.5, 0.5, 1) distribution for n = 100, 200, 300 and 400.

		MLE			MSADE			MSALDE		
n		AE	Bias	MSE	AE	Bias	MSE	AE	Bias	MSE
100	a	3.158	1.158	5.743	2.271	0.271	5.404	2.361	0.361	14.717
	b	2.826	1.326	5.933	1.870	0.370	5.206	2.053	0.553	14.854
	c	0.587	2.658	0.301	0.509	1.771	0.027	0.582	1.861	0.133
	α	1.203	0.203	0.817	1.074	0.074	0.303	1.145	0.145	0.485
200	a	2.862	0.862	3.915	2.179	0.179	2.771	2.072	0.072	2.715
	b	2.461	0.961	3.758	1.750	0.250	2.855	1.651	0.151	2.837
	c	0.539	2.362	0.126	0.535	1.679	0.048	0.582	1.572	0.081
	α	1.114	0.114	0.440	1.078	0.078	0.245	1.141	0.141	0.334
300	a	2.112	0.112	2.492	2.666	0.666	2.609	2.133	0.133	3.709
	b	1.695	0.195	2.331	2.217	0.717	2.475	1.695	0.195	3.368
	c	0.554	1.612	0.072	0.519	2.166	0.080	0.583	1.633	0.080
	α	1.051	0.051	0.176	1.097	0.097	0.310	1.130	0.130	0.248
400	a	2.587	0.587	1.956	2.048	0.048	0.956	2.143	0.143	3.588
	b	2.109	0.609	1.869	1.602	0.102	0.970	1.669	0.169	3.383
	c	0.498	2.087	0.049	0.534	1.548	0.026	0.558	1.643	0.039
	α	1.080	0.080	0.232	1.062	0.062	0.161	1.135	0.135	0.220

We simulate the BOLL-E distribution for n=100, 200, 300 and 400 with $a = 2$, $b = 1.5$, $c = 0.5$ and $\alpha = 1$. For each sample size, we compute the MLEs, MSADEs and MSALDEs of the parameters. We repeat this process 1,000 times and obtain the average estimates (AEs), biases and mean square error (MSEs). The results are reported in Table 2. From the figures in this table, we note that the performances of the MLEs and MSADEs are better than MSALDEs.

7. Applications

In this section, we provide two applications to real data to illustrate the importance of the BOLL-G family through the special models: BOLL-E, BOLL-N and BOLL-Lx. The MLEs of the parameters are computed and the goodness-of-fit statistics for these models are compared with other competing models.

7.1. Data set 1: Strength of glass fibres. The first data set represents the strength of 1.5 cm glass fibres, measured at National physical laboratory, England (see, Smith and Naylor [46]). The data are: 0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2.00, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.50, 1.54, 1.60, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.50, 1.55, 1.61, 1.62, 1.66, 1.70, 1.77, 1.84, 0.84, 1.24, 1.30, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.70, 1.78, 1.89.

We fit the BOLL-E, BOLL-N, McDonald-Normal (McN) (Cordeiro et al. [9]), beta-normal (BN) (Famoye et al. [17]) and beta-exponential (BE) (Nadarajah and Kotz [39]) models to data set 1 and also compare them through seven goodness-of-fit statistics. The densities of the McN, BN and BE models are, respectively, given by:

$$\text{McN}: \; f_{\text{McN}}(x;a,b,c,\mu,\sigma) = \tfrac{c}{\sigma\, B(a,b)}\, \phi\left(\tfrac{x-\mu}{\sigma}\right) \Phi\left(\tfrac{x-\mu}{\sigma}\right)^{ac-1} \left[1-\Phi\left(\tfrac{x-\mu}{\sigma}\right)^{c}\right]^{b-1}, \quad \mu \in \Re, \quad a,b,c,\sigma > 0,$$

$$\text{BN}: \; f_{\text{BN}}(x;a,b,\mu,\sigma) = \tfrac{1}{\sigma\, B(a,b)} \phi\left(\tfrac{x-\mu}{\sigma}\right) \Phi\left(\tfrac{x-\mu}{\sigma}\right)^{a-1} \left[1-\Phi\left(\tfrac{x-\mu}{\sigma}\right)\right]^{b-1}, \quad \mu \in \Re, \quad a,b,\sigma > 0,$$

$$\text{BE}: \; f_{\text{BE}}(x;a,b,\alpha) = \tfrac{\alpha}{B(a,b)}\, \mathrm{e}^{-\alpha\, b\, x} \; \left(1-\mathrm{e}^{-\alpha\, x}\right)^{a-1}, \quad a,b,\alpha > 0.$$

7.2. Data set 2: Bladder cancer patients. The second data set represents the uncensored remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang [31]. The data are: 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33,5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

We fit the BOLL-E, BOLL-Lx, McDonald-Lomax (McLx) and beta-Lomax (BLx) (Lemonte and Cordeiro [32]) and BE models to these data and also compare their goodness-of-fit statistics. The densities of the McLx and BLx models are, respectively, given by

$$\begin{aligned} \text{McLx}: \; f_{\text{McLx}}(x;a,b,c,\alpha,\beta) &= \tfrac{c\,\alpha}{\beta\, B(a,b)} \left[1+\left(\tfrac{x}{\beta}\right)\right]^{-(\alpha+1)} \\ &\times \; \left\{1-\left[1+\left(\tfrac{x}{\beta}\right)\right]^{-\alpha}\right\}^{ac-1} \left[1-\left\{1-\left[1+\left(\tfrac{x}{\beta}\right)\right]^{-\alpha}\right\}^{c}\right]^{b-1}, \\ & \qquad a,b,c,\alpha,\beta > 0, \end{aligned}$$

$$\begin{aligned} \text{BLx}: \; f_{\text{BLx}}(x;a,b,\alpha,\beta) &= \tfrac{\alpha}{\beta\, B(a,b)} \left[1+\left(\tfrac{x}{\beta}\right)\right]^{-(\alpha b+1)} \left\{1-\left[1+\left(\tfrac{x}{\beta}\right)\right]^{-\alpha}\right\}^{a-1} \\ &\times \; \left[1-\left\{1-\left[1+\left(\tfrac{x}{\beta}\right)\right]^{-\alpha}\right\}^{a}\right]^{b-1}, \quad a,b,\alpha,\beta > 0. \end{aligned}$$

For all models, the MLEs are computed using the Limited-Memory Quasi-Newton Code for Bound-Constrained Optimization (L-BFGS-B). Further, the log-likelihood function evaluated at the MLEs ($\hat{\ell}$), Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), Anderson-Darling (A^*), Cramér–von Mises (W^*) and Kolmogorov-Smirnov (K-S) statistics are calculated to compare the fitted models. The statistics A^* and W^* are defined by Chen and Balakrishnan [8]. In general, the smaller the values of these statistics, the better the fit to the data. The required computations are carried out in R-language.

Table 3: MLEs and their standard errors (in parentheses) for the first data set.

Distribution	a	b	c	μ	σ	α
BOLL-E	0.0698	0.1834	50.4548	-	-	0.4118
	(0.0931)	(0.2712)	(66.9766	-	-	(0.0125)
BOLL-N	0.0358	0.0764	34.7642	1.6597	0.6056	-
	(0.0660)	(0.1384)	(65.6410)	(0.0381)	(0.5323)	-
McN	0.5298	17.2226	1.2924	2.3850	0.4773	-
	(0.5249)	(48.8078)	(6.2595)	(1.8112)	(0.9820)	-
BN	0.5836	21.9402	-	2.5679	0.4658	-
	(0.6444)	(79.8234)	-	(1.3451)	(0.4546)	-
BE	17.4548	38.3856	-	-	-	0.2514
	(3.1323)	(65.8297)	-	-	-	(0.3684)

Table 4: The statistics $\hat{\ell}$, AIC, CAIC , BIC , HQIC, A^* and W^* for the first data set.

Distribution	$\hat{\ell}$	AIC	CAIC	BIC	HQIC	A^*	W^*
BOLL-E	**−10.4852**	**28.9703**	**29.6599**	**37.5429**	**32.3419**	**0.3923**	**0.0681**
BOLL-N	**−9.9976**	**29.9953**	**31.0479**	**40.7110**	**34.2098**	2.0245	**0.2858**
McN	−14.0577	38.1154	39.1680	48.8311	42.3299	0.9289	0.1659
BN	−14.0560	36.1119	36.8016	44.6845	39.4836	0.9179	0.1637
BE	−24.0256	54.0511	54.4579	60.4805	56.5798	3.1307	0.5708

Table 5: The K-S statistics and p-values for the first data set.

Distribution	K-S	p-value (K-S)
BOLL-E	**0.1126**	**0.4013**
BOLL-N	**0.0928**	**0.6496**
McN	0.1369	0.1886
BN	0.1356	0.1973
BE	0.2168	0.0053

Table 6: MLEs and their standard errors (in parentheses) for the second data set.

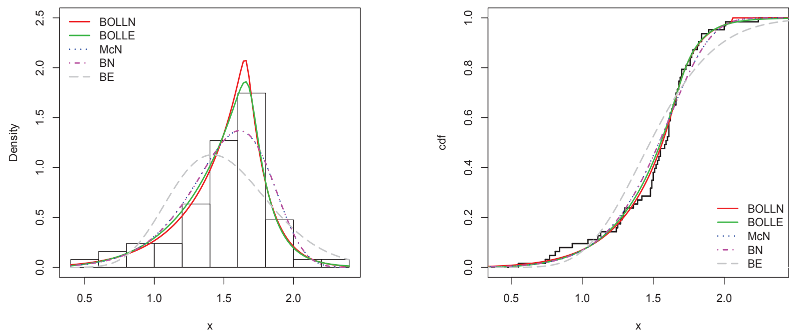
Distribution	a	b	c	α	β
BOLL-E	0.2772	0.1548	3.7895	0.1563	-
	(0.2529)	(0.1441)	(3.1996)	(0.0413)	-
BOLL-Lx	0.4507	0.3046	2.5267	8.5700	57.6246
	(0.4279)	(0.3573)	(2.0183)	(14.4135)	(88.4252)
McLx	1.5052	5.9638	2.0608	0.7177	10.9267
	(0.2831)	(30.1616)	(2.9944)	(3.0698)	(16.6896)
BLx	1.5882	12.0014	-	0.3859	20.4693
	(0.2830)	(319.2372)	-	(10.0697)	(14.0657)
BE	1.3781	0.2543	-	0.4595	-
	(0.2162)	(0.0251)	-	(0.0028)	-

Table 7: The statistics $\hat{\ell}$, AIC, CAIC , BIC , HQIC, A^* and W^* for the second data set.

Distribution	$\hat{\ell}$	AIC	CAIC	BIC	HQIC	A^*	W^*
BOLL-E	**−409.8323**	**827.6646**	**827.9898**	**839.0727**	**832.2998**	**1.5745**	**0.2022**
BOLL-Lx	**−409.2256**	**828.4513**	**828.9431**	**842.7115**	**834.2453**	**0.0800**	**0.0126**
McLx	−409.9128	829.8256	830.3174	844.0858	835.6196	0.1688	0.0254
BLx	−410.0813	828.1626	828.4878	839.5708	832.7978	0.1917	0.0285
BE	−412.1016	830.2033	830.3968	838.7594	833.6797	0.5475	0.0896

Table 8: The K-S statistics and p-values for the second data set.

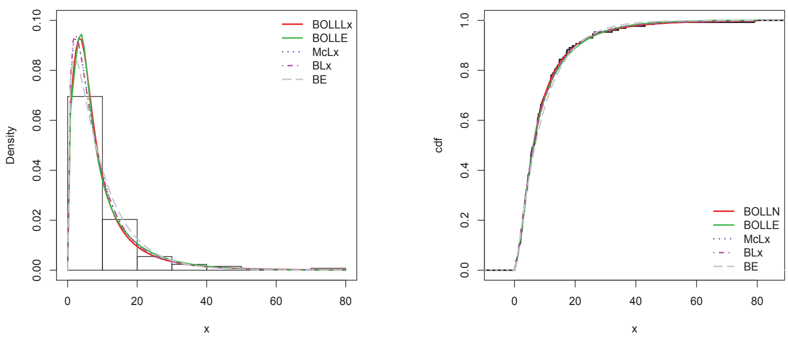
Distribution	K-S	p-value (K-S)
BOLL-E	**0.0295**	**0.9999**
BOLL-Lx	**0.0341**	**0.9984**
McLx	0.0391	0.9896
BLx	0.0407	0.9840
BE	0.0688	0.5793



(a) Estimated pdfs (b) Estimated cdfs

Figure 4. Plots (a) and (b) of the estimated pdfs and cdfs of the BOLL-E and BOLL-N and other competitive models.

Tables 3 and 6 list the MLEs and their corresponding standard errors (in parentheses) of the parameters. The values of the model selection statistics AIC, CAIC, BIC, HQIC, A^*, W^* and K-S are listed in Tables 4-5 and 7-8. We note from Tables 4 and 5 that the BOLL-E and BOLL-N models have the lowest values of the AIC, CAIC, BIC, HQIC, W^* and K-S statistics (for the first data set) among the fitted McN, BN and BE models, thus suggesting that the BOLL-E and BOLL-N models provide the best fits, and therefore could be chosen as the most adequate models for the first data set. The histogram of these data and the estimated pdfs and cdfs of the BOLL-E and BOLL-N models and their competitive models are displayed in Figure 4. Similarly, it is also evident from the results in Tables 7 and 8 that the BOLL-E and BOLL-Lx models give the lowest values for the $\hat{\ell}$, AIC, CAIC, BIC, HQIC, A^*, W^* and K-S statistics (for the second data set) among the fitted McLx, BLx, KwLx and Lx distributions. Thus, the BOLL-E and BOLL-Lx models can be chosen as the best models. The histogram of the second data set and the estimated pdfs and cdfs of the BOLL-E and BOLL-Lx models and other competitive models are displayed in Figure 5.



(a) Estimated pdfs (b) Estimated cdfs

Figure 5. Plots (a) and (b) of the estimated pdfs and cdfs of the BOLL-E and BOLL-Lx models and other competitive models.

It is clear from the figures in Tables 4-5 and 7-8, and Figures 4 and 5 that the BOLL-E, BOLL-N and BOLL-Lx models provide the best fits to these two data sets as compared to other models.

8. Concluding remarks

The generalized continuous univariate distributions have been widely studied in the literature. We propose a new class of distributions called the *beta odd log-logistic-G* family. We study some of its structural properties including an expansion for its density function and explicit expressions for the moments, generating function, mean deviations, quantile function and order statistics. The maximum likelihood method and the method of minimum spacing distance are employed to estimate the model parameters. We fit three special models of the proposed family to two real data sets to demonstrate its usefulness. We use some goodness-of-fit statistics in order to determine which distribution fits better to these data. We conclude that these special models provide consistently better fits than other competing models. We hope that the new family and its generated models will attract wider applications in several areas such as reliability engineering, insurance, hydrology, economics and survival analysis.

Appendix A

We present four power series expansions required for the proof of the general result in Section 4. First, for $a > 0$ real non-integer, we have the binomial expansion

$$(1-u)^a = \sum_{j=0}^{\infty} (-1)^j \binom{a}{j} u^j, \tag{8.1}$$

where the binomial coefficient is defined for any real a as $a(a-1)(a-2), \ldots, (a-j+1)/j!$.

Second, the following expansion holds for any $\alpha > 0$ real non-integer

$$G(x)^\alpha = \sum_{r=0}^{\infty} a_r(\alpha)\, G(x)^r, \tag{8.2}$$

where $a_r(\alpha) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{\alpha}{j}\binom{j}{r}$. The proof of (8.2) follows from $G(x)^\alpha = \{1 - [1 - G(x)]\}^\alpha$ by applying (8.1) twice.

Third, by expanding z^λ in Taylor series (when k is a positive integer), we have

$$z^\lambda = \sum_{k=0}^{\infty} (\lambda)_k \, (z-1)^k / k! = \sum_{i=0}^{\infty} f_i \, z^i, \tag{8.3}$$

where

$$f_i = f_i(\lambda) = \sum_{k=0}^{\infty} \frac{(-1)^{k-i}}{k!} \binom{k}{i} (\lambda)_k$$

and $(\lambda)_k = \lambda(\lambda-1)\ldots(\lambda-k+1)$ is the descending factorial.

Fourth, we use throughout an equation of Gradshteyn and Ryzhik [22] for a power series raised to a positive integer i given by

$$\left(\sum_{j=0}^{\infty} a_j \, v^j \right)^i = \sum_{j=0}^{\infty} c_{i,j} \, v^j, \tag{8.4}$$

where the coefficients $c_{i,j}$ (for $j = 1, 2, \ldots$) are obtained from the recurrence equation (for $j \geq 1$)

$$c_{i,j} = (j a_0)^{-1} \sum_{m=1}^{j} [m(j+1) - j] \, a_m \, c_{i,j-m} \tag{8.5}$$

and $c_{i,0} = a_0^i$. Hence, $c_{i,j}$ can be calculated directly from $c_{i,0}, \ldots, c_{i,j-1}$ and, therefore, from $a_0, \ldots, a_j$.

We now obtain an expansion for $[G(x)^c + \bar{G}(x)^c]^a$. We can write from equations (8.1) and (8.2)

$$[G(x)^c + \bar{G}(x)^c] = \sum_{j=0}^{\infty} t_j \, G(x)^j,$$

where

$$t_j = (-1)^j \left[\binom{c}{j} + \sum_{i=j}^{\infty} (-1)^i \binom{c}{i} \binom{c}{j} \right].$$

Then, using (8.3), we have

$$[G(x)^c + \bar{G}(x)^c]^a = \sum_{i=0}^{\infty} f_i \, \left(\sum_{j=0}^{\infty} t_j \, G(x)^j \right)^i,$$

where $f_i = f_i(a)$ is defined before.

Finally, using equations (8.4) and (8.5), we obtain

$$[G(x)^c + \bar{G}(x)^c]^a = \sum_{j=0}^{\infty} h_j \, G(x)^j, \tag{8.6}$$

where

$$h_j = h_j(c, a) = \sum_{i=0}^{\infty} \, f_i \, m_{i,j},$$

$$m_{i,j} = (j \, t_0)^{-1} \sum_{m=1}^{j} [m(j+1) - j] \, t_m \, m_{i,j-m} \quad \text{(for} \quad j \geq 1)$$

and $m_{i,0} = t_0^i$.

Acknowledgments

The authors gratefully acknowledge the help of Professor Hamzeh Torabi in developing R-code for the MSDEs. The authors are also grateful to the Editor-in-Chief (Professor Cem Kadilar) and two referees for their helpful comments and suggestions.

References

[1] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M. *Generalized beta-generated distributions*, Computational Statistics and Data Analysis **56**, 1880–1897, 2012.

[2] Aljarrah, M.A., Lee, C. and Famoye, F. *On generating T-X family of distributions using quantile functions*, Journal of Statistical Distributions and Applications **1**, Article 2, 2014.

[3] Alzaatreh, A., Lee, C. and Famoye, F. *A new method for generating families of continuous distributions*, Metron **71**, 63–79, 2013.

[4] Alzaatreh, A., Lee, C. and Famoye, F. *T-normal family of distributions: A new approach to generalize the normal distribution*, Journal of Statistical Distributions and Applications **1**, Article 16, 2014.

[5] Alzaghal, A., Lee, C. and Famoye, F. *Exponentiated T-X family of distributions with some applications*, International Journal of Probability and Statistics **2**, 31–49, 2013.

[6] Amini, M., MirMostafaee, S.M.T.K. and Ahmadi, J. *Log-gamma-generated families of distributions*, Statistics **48**, 913–932, 2014.

[7] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. *The Weibull–G family of probability distributions*, Journal of Data Science **12**, 53–68, 2014.

[8] Chen, G. and Balakrishnan, N. *A general purpose approximate goodness-of-fit test*, Journal of Quality Technology **27**, 154–161, 1995.

[9] Cordeiro, G.M., Cintra, R.J., Rego, L.C. and Ortega, E.M.M. *The McDonald normal distribution*, Pakistan Journal of Statistics and Operations Research **8**, 301–329, 2012.

[10] Cordeiro, G.M. and de Castro, M. *A new family of generalized distributions*, Journal of Statistical Computation and Simulation **81**, 883–893, 2011.

[11] Cordeiro, G.M. and Nadarajah, S. *Closed-form expressions for moments of a class of beta generalized distributions*, Brazilian Journal of Probability and Statistics **25**, 14–33, 2011.

[12] Cordeiro, G.M., Ortega, E.M.M. and da Cunha, D.C.C. *The exponentiated generalized class of distributions*, Journal of Data Science **11**, 1–27, 2013.

[13] Cordeiro, G.M., Alizadeh, M. and Ortega, E.M.M. *The exponentiated half-logistic family of distributions: Properties and applications*, Journal of Probability and Statistics Article ID 864396, 21 pages, 2014.

[14] Doornik, J.A. *Ox 5: An Object-Oriented Matrix Programming Language*, Fifth edition (Timberlake Consultants, London, 2007)

[15] Eugene, N., Lee, C. and Famoye, F. *Beta-normal distribution and its applications*, Communications in Statistics–Theory and Methods **31**, 497–512, 2002.

[16] Exton, H. *Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs* (Ellis Horwood, New York, 1978)

[17] Famoye, F., Lee, C. and Eugene, N. *Beta-normal distribution: Biomodality properties and application*, Journal of Modern Applied Statistical Methods **3**, 85–103, 2004.

[18] Gleaton, J.U. and Lynch, J.D. *On the distribution of the breaking strain of a bundle of brittle elastic fibers*, Advances in Applied Probability **36**, 98–115, 2004.

[19] Gleaton, J.U. and Lynch, J.D. *Properties of generalized log-logistic families of lifetime distributions*, Journal of Probability and Statistical Science **4**, 51–64, 2006.

[20] Gleaton, J.U. and Rahman, M. *Asymptotic properties of MLE's for distributions generated from a 2-parameter Weibull distribution by a generalized log-logistic transformation*, Journal of Probability and Statistical Science **8**, 199–214, 2010.

[21] Gleaton, J.U. and Rahman, M. *Asymptotic properties of MLE's for distributions generated from a 2-parameter inverse Gaussian distribution by a generalized log-logistic transformation*, Journal of Probability and Statistical Science **12**, 85–99, 2014.

[22] Gradshteyn, I.S. and Ryzhik, I.M. *Table of Integrals, Series, and Products*, Sixth edition (Academic Press, San Diego, 2000)

[23] Greenwood, J.A., Landwehr, J.M. and Matalas, N. C. *Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form*, Water Resources Research **15**, 1049–1054, 1979.

[24] Glänzel, W. *A characterization theorem based on truncated moments and its application to some distribution families*, In: Mathematical Statistics and Probability Theory, Volume B, pp. 75–84 (Reidel, Dordrecht, 1987)

[25] Gupta, R.C. and Gupta, R.D. *Proportional revered hazard rate model and its applications*, Journal of Statistical Planning and Inference **137**, 3525–3536, 2007.

[26] Gupta, R. C., Gupta, P. I. and Gupta, R. D. *Modeling failure time data by Lehmann alternatives*, Communications in statistics–Theory and Methods **27**, 887–904, 1998.

[27] Gupta, R.D. and Kundu, D. *Generalized exponential distribution*, Australian and New Zealand Journal of Statistics **41**, 173–188, 1999.

[28] Gupta, R.D. and Kundu, D. *Generalized exponential distribution: An alternative to Gamma and Weibull distributions*, Biometrical Journal **43**, 117–130, 2001.

[29] Jones, M.C. *Families of distributions arising from the distributions of order statistics*, Test **13**, 1–43, 2004.

[30] Kenney, J. and Keeping, E. *Mathematics of Statistics*, Volume 1, Third edition (Van Nostrand, Princeton, 1962)

[31] Lee, E.T. and Wang, J.W. *Statistical Methods for Survival Data Analysis*, Third edition (Wiley, New York, 2003)

[32] Lemonte, A.J. and Cordeiro, G.M. *An extended Lomax distribution*, Statistics **47**, 800–816, 2013.

[33] Moors, J.J.A. *A quantile alternative for kurtosis*, The Statistician **37**, 25–32, 1998.

[34] Mudholkar, G.S. and Hutson, A.D. *The exponentiated Weibull family: Some properties and a flood data application*, Communications in Statistics–Theory and Methods **25**, 3059–3083, 1996.

[35] Mudholkar, G. S. and Srivastava, D.K. *Exponentiated Weibull family for analyzing bathtub failure data*, IEEE Transactions on Reliability **42**, 299–302, 1993.

[36] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. *The exponentiated Weibull family: A reanalysis of the bus-motor failure data*, Technometrics **37**, 436–445, 1995.

[37] Nadarajah, S. *The exponentiated exponential distribution: a survey*, AStA Advances in Statistical Analysis **95**, 219–251, 2011.

[38] Nadarajah, S., Cordeiro, G.M., Ortega and E.M.M. *The Zografos-Balakrishnan–G family of distributions: Mathematical properties and applications*, Communications in Statistics–Theory and Methods **44**, 186–215, 2015.

[39] Nadarajah, S. and Kotz, S. *The exponentiated-type distributions*, Acta Applicandae Mathematicae **92**, 97–111, 2006.

[40] Nadarajah, S. and Gupta, A.K. *The exponentiated gamma distribution with application to drought data*, Calcutta Statistical Association Bulletin **59**, 29–54, 2007.

[41] Nadarajah, S. and Gupta, A.K. *A generalized gamma distribution with application to drought data*, Mathematics in Computer and Simulation **74**, 1–7, 2007.

[42] R Development Core Team *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (Vienna, Austria, 2009)

[43] Rényi, A. *On measures of entropy and information*, In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability–I, University of California Press, Berkeley, pp. 547–561, 1961.

[44] Ristić, M.M. and Balakrishnan, N. *The gamma-exponentiated exponential distribution*, Journal of Statistical Computation and Simulation **82**, 1191–1206, 2012.

[45] Shannon, C.E. *A mathematical theory of communication*, Bell System Technical Journal **27**, 379–432, 1948.

[46] Smith, R.L. and Naylor, J.C. *A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution*, Applied Statistics **36**, 358–369, 1987.

[47] Tahir, M.H., Cordeiro, G.M., Alzaatreh, A., Zubair, M. and Mansoor, M. *The Logistic-X family of distributions and its applications*, Communications in Statistics–Theory and Methods **45**, 7326–7349, 2016.

[48] Torabi, H. *A general method for estimating and hypotheses testing using spacings*, Journal of Statistical Theory and Applications **8**, 163–168, 2008.

[49] Torabi, H. and Bagheri, F.L. *Estimation of parameters for an extended generalized half-logistic distribution based on complete and densored data*, Journal of the Iranian Statistical Society **9**, 171–195, 2010.

[50] Torabi, H. and Montazari, N.H. *The gamma-uniform distribution and its application*, Kybernetika **48**, 16–30, 2012.

[51] Torabi, H. and Montazari, N.H. *The logistic-uniform distribution and its application*, Communications in Statistics–Simulation and Computation **43**, 2551–2569, 2014.

[52] Trott, M. *The Mathematica Guidebook for Symbolics* (Springer, New York, 2006)

[53] Zografos K. and Balakrishnan, N. *On families of beta- and generalized gamma-generated distributions and associated inference*, Statistical Methodology **6**, 344–362, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1203 – 1229

The Kumaraswamy exponential-Weibull distribution: theory and applications

Gauss M. Cordeiro*, Abdus Saboor†, Muhammad Nauman Khan‡, Gamze Ozel§ and Marcelino A.R. Pascoa¶

Abstract

Significant progress has been made towards the generalization of some well–known lifetime models, which have been successfully applied to problems arising in several areas of research. In this paper, some properties of the new Kumaraswamy exponential-Weibull (KwEW) distribution are provided. This distribution generalizes a number of well-known special lifetime models such as the Weibull, exponential, Rayleigh, modified Rayleigh, modified exponential and exponentiated Weibull distributions, among others. The beauty and importance of the new distribution lies in its ability to model monotone and non-monotone failure rate functions, which are quite common in environmental studies. We derive some basic properties of the KwEW distribution including ordinary and incomplete moments, skewness, kurtosis, quantile and generating functions, mean deviations and Shannon entropy. The method of maximum likelihood and a Bayesian procedure are used for estimating the model parameters. By means of a real lifetime data set, we prove that the new distribution provides a better fit than the Kumaraswamy Weibull, Marshall-Olkin exponential-Weibull, extended Weibull, exponential-Weibull and Weibull models. The application indicates that the proposed model can give better fits than other well-known lifetime distributions.

Keywords: Exponential–Weibull distribution, Fox–Wright generalized ${}_p\Psi_q$ function, generalized distribution, lifetime data, maximum likelihood, moment.

2000 AMS Classification: 62E15, 60E05

*Departamento de Estatística, Universidade Federal de Pernambuco, 50740-540, Brazil, Email: gausscordeiro@gmail.com

†Corresponding Author
Department of Mathematics, Kohat University of Science & Technology, Kohat, 26000, Pakistan, Email: saboorhangu@gmail.com; dr.abdussaboor@kust.edu.pk

‡Department of Mathematics, Kohat University of Science & Technology, Kohat, 26000, Pakistan, Email: Zaybasdf@gmail.com

§Department of Statistics, Hacettepe University, 06800, Turkey, Email: gamzeozl@gmail.com

¶Departamento de Estatística, Universidade Federal de Mato Grosso, 78075-850, Brazil, Email: marcelino.pascoa@gmail.com

Received : 12.12.2014 *Accepted* : 13.07.2015 *Doi* : 10.15672/HJMS.20157612083

1. Introduction

In many applied areas like lifetime analysis, finance, insurance and biology, there is a clear need for extended forms of the classical distributions, i.e., new distributions more flexible to model real data that present a high degree of skewness and kurtosis in these areas. Recent developments focus on new techniques by adding parameters to existing distributions for building classes of more flexible distributions. Following this idea, Cordeiro et al. [6] introduced an interesting method by adding two new parameters to a parent distribution to model data with a high degree of skewness and kurtosis. The generated family can provide more flexibility to model various types of data. If $G(x)$ is the cumulative distribution function (cdf) of a baseline model, then the *Kumaraswamy generalized* (Kw-G) family has cdf given by

$$F(x) = 1 - \{1 - G^{\alpha}(x)\}^{\gamma}. \tag{1.1}$$

The probability density function (pdf) corresponding to (1.1) is given by

$$f(x) = \alpha\,\gamma\,g(x)\,G^{\alpha-1}(x)\,\{1 - G^{\alpha}(x)\}^{\gamma-1}\,. \tag{1.2}$$

Each new Kw-G distribution can be obtained from a specified G distribution. For $\alpha = \gamma = 1$, the G distribution is a basic exemplar of the Kw-G family with a continuous crossover towards cases with different shapes (e.g., a particular combination of skewness and kurtosis). One major benefit of equation (1.2) is its ability of fitting skewed data that can not be properly fitted by existing distributions. Further, it allows for greater flexibility of its tails and can be widely applied in many areas of reliability and biology.

The Weibull distribution is a very popular distribution for modeling lifetime data. When modeling monotone hazard rates, it may be an initial choice because of its skewed density shapes. However, it does not have a bathtub or upside–down bathtub shaped hazard rate function (hrf) and can not be used to model the lifetime of certain systems. Such bathtub hazard curves have nearly flat middle portions and the corresponding densities have a positive anti-mode. An example of the bathtub-shaped failure rate is the human mortality experience with a high infant mortality rate which reduces rapidly to reach a low level. Unimodal failure rates can be observed in course of a disease whose mortality reaches a peak after some finite period and then declines gradually. Thus, it cannot be used to model lifetime data with a bathtub shaped hazard function, such as human mortality and machine life cycles. Therefore, several researchers have developed various extensions and modified forms of the Weibull distribution having a number of parameters ranging from two to five parameters.

In the last few years, new classes of distributions aim to define generalized Weibull distributions to cope with bathtub shaped failure rates. Mudholkar and Srivastava [17] and Mudholkar et al. [18] pioneered and studied the exponentiated Weibull (ExpW) distribution to analyze bathtub failure data. A good review of some of these extended models is presented in Pham and Lai [25]. Also, the additive Weibull distribution was proposed by Xie and Lai [27], the modified Weibull distribution by Lai et al. [12] and the generalized modified Weibull distribution by Carrasco et al. [2]. Further, Lee et al. [13] and Silva et al. [23] defined two extensions of the Weibull model called the beta Weibull (BW) and beta modified Weibull (BMW) distributions, respectively.

The *exponential–Weibull* (EW) distribution proposed by Cordeiro *et al.* [5] has cdf and pdf given by

$$G(x) = 1 - \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\, 1_{\mathbb{R}_+}(\mathrm{x}),\ \ \lambda > 0,\ \beta > 0,\ \mathrm{k} > 0 \tag{1.3}$$

and

$$g(x) = (\lambda + \beta\, k\, x^{k-1})\, \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\, 1_{\mathbb{R}_+}(\mathrm{x}), \tag{1.4}$$

respectively, where $\lambda > 0$ and $k > 0$ are shape parameters, $\beta > 0$ is a scale parameter and $1_A(x)$ denotes the characteristic function of the set A, i.e. $1_A(x) = 1$ when $x \in A$ and equals 0 elsewhere.

We generalize the EW model by defining the *Kumaraswamy exponential–Weibull* (KwEW) distribution. The cdf and pdf of the KwEW distribution, for which the EW is the baseline model, are given by

$$F(x) = 1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma} 1_{\mathbb{R}_+}(x) \tag{1.5}$$

and

$$\begin{aligned} f(x) = {} & \alpha\gamma\left(\lambda + k\beta x^{k-1}\right)\mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\left(1 - \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\right)^{-1+\alpha} \\ & \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\right)^{\alpha}\right\}^{-1+\gamma} 1_{\mathbb{R}_+}(x), \end{aligned} \tag{1.6}$$

respectively, where $\lambda > 0$, $\beta > 0$, $k > 0$, $\alpha > 0$ and $\gamma > 0$. Hereafter, we denote by $X \sim \mathrm{KwEW}_{\alpha,\gamma}(\lambda, \beta, k)$ a random variable having the pdf (1.6).

The density (1.6) is much more flexible than the EW density and can allow for greater flexibility of the tails. It can exhibit different behavior depending on the parameter values. In fact, Figure 1 (a,c) and Figure 2 (d) reveal that the mode of the pdf increases as α and λ increases, respectively. Figure 2 (e) also shows that the mode of the pdf increases as k increases. The new parameter γ behaves somewhat as a scale parameter as shown in Figure 1(b). The structure of the density function (1.6) can be motivated as it provides more flexible distribution than the two-parameter Weibull and many other extended Weibull distributions (see Table 1).

The rest of the paper is organized as follows. In Section 2, twelve widely-known special models of the proposed distribution are presented. A useful expansion for the KwEW density and explicit expressions for certain mathematical quantities of X are obtained in Section 3. We demonstrate in Section 4 that the KwEW density is an infinite mixture of EW densities. Further, we obtain alternative expressions for the moments and generating function. The estimation of the model parameters by maximum likelihood and a Bayesian procedure are addressed in Section 5. We prove in Section 6 the flexibility of the new distribution for modeling lifetime data by means of a real data set. A bivariate extension is given in Section 7. The paper is concluded in Section 8.

2. Special Distributions

We point out some special cases of the $\mathrm{KwEW}_{\alpha,\gamma}(\lambda, \beta, k)$ distribution by specifying its parameters values. Table 1 lists twelve important special models of the new distribution. For example, the $\mathrm{KwEW}_{\alpha,\gamma}(0, \beta, k)$ model reduces to the *Kw-modified Weibull* [12], the $\mathrm{KwEW}_{1,1}(\lambda, \beta, k)$ refers to the *exponential–Weibull* [5], the $\mathrm{KwEW}_{1,1}(\lambda, \beta, 2)$ is the *modified Rayleigh*, the $\mathrm{KwEW}_{1,1}(\lambda, \beta, 1)$ is the *modified exponential* and the $\mathrm{KwEW}_{1,1}(0, \beta, k)$ becomes the classical two-parameter Weibull. If $k = 1$ and $k = 2$ in addition to $\alpha = 1, \gamma = 1$ and $\lambda = 0$, it coincides with the exponential and Rayleigh distributions, respectively. Finally, the $\mathrm{KwEW}_{1,\gamma}(0, \beta, k)$ model becomes the ExpW distribution pioneered by [17, 18].

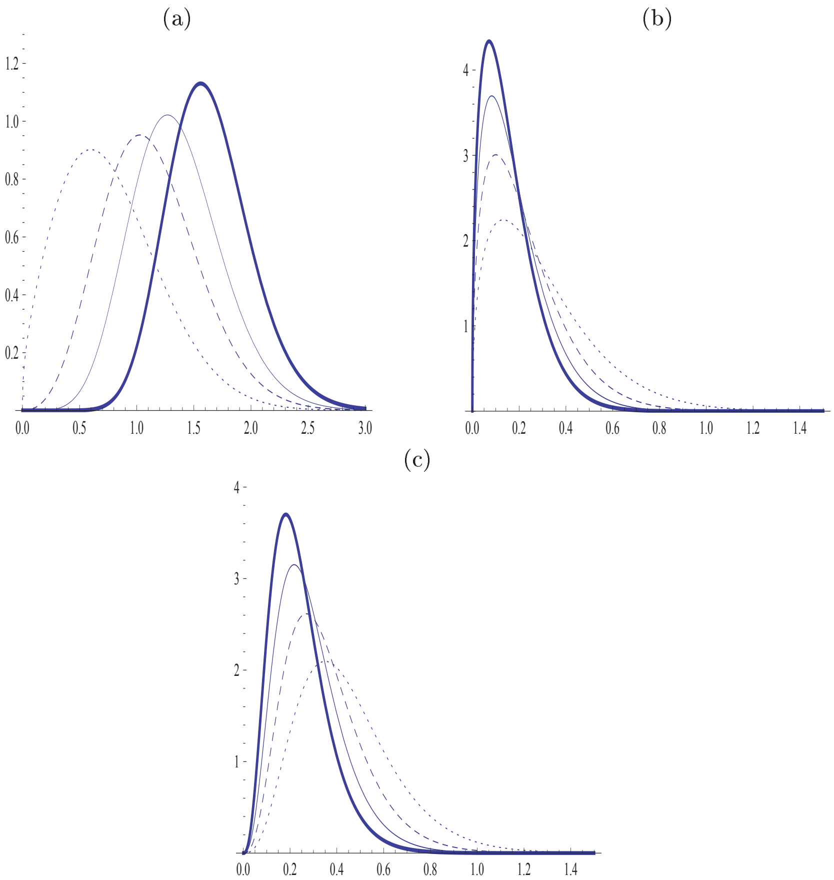
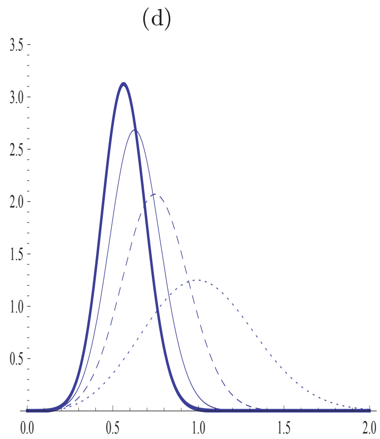


Figure 1. Plots of the KwEW density function. (a) $\lambda = 0.5$, $\beta = 0.6$, $k = 2$, $\gamma = 1.5$ and $\alpha = 1.4$ (dotted line), $\alpha = 3$ (dashed line), $\alpha = 5$ (solid line), $\alpha = 10$ (thick line). (b) $\lambda = 3.5$, $\beta = 1.6$, $k = 2$, $\alpha = 1.5$ and $\gamma = 1$ (dotted line), $\gamma = 1.5$ (dashed line), $\gamma = 2$ (solid line), $\gamma = 2.5$ (thick line). (c) $\beta = 2.6$, $k = 1.2$, $\alpha = 3.5$, $\gamma = 1.7$ and $\lambda = 1$ (dotted line) $\lambda = 2$, (dashed line), $\lambda = 3$ (solid line), $\lambda = 4$ (thick line).

3. Main Properties

We derive computational sum–representations and explicit expressions for the ordinary and central moments, skewness, kurtosis, generating and quantile functions, Shannon entropy and mean deviations of X. These expressions can be evaluated analytically or numerically using packages such as *Mathematica*, *Matlab* and *Maple*. In numerical applications, the infinite sums can be truncated whenever convergence is observed.



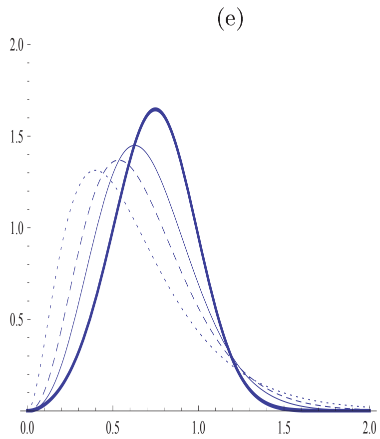


Figure 2. Plots of the KwEW density function. (d) $\lambda = 1.3$, $k = 3$, $\alpha = 5$, $\gamma = 1.3$ and $\beta = 0.5$ (dotted line), $\beta = 2$ (dashed line), $\beta = 4$ (solid line), $\beta = 6$ (thick line). (e) $\lambda = 1$, $\beta = 1.5$, $\alpha = 3$, $\gamma = 1.3$ and $k = 1$ (dotted line), $k = 1.5$ (dashed line), $k = 2$ (solid line), $k = 3$ (thick line).

Table 1. Some special distributions

Model	λ	β	k	α	γ
Kw-Modified Weibull	0	–	–	–	–
Kw-Exponential	–	0	–	–	–
Kw-Rayleigh	0	–	2	–	–
Exponentiated Weibull	0	–	–	1	–
Kw-Linear Failure Rate	–	–	2	–	–
Exponential Weibull	–	–	–	1	1
Two Parameter Weibull	0	–	–	1	1
Exponential	0	–	1	1	1
Rayleigh	0	–	2	1	1
Modified Rayleigh	–	–	2	1	1
Modified Exponential	–	–	1	1	1
Linear Failure Rate	–	–	2	1	1

3.1. A Useful Expansion. Here, we provide a useful expansion for the KwEW pdf (1.6). By using the power series

$$(1-z)^{\beta-1} = \sum_{n=0}^{\infty} a_n\, z^n,\ |z|<1,\ \beta>0, \tag{3.1}$$

we obtain

$$f(x) = \alpha\gamma\left(\lambda + k\beta x^{k-1}\right)\sum_{m=0}^{\infty} W_m\left(\mathrm{e}^{-\lambda\,\mathrm{x}-\beta\,\mathrm{x}^{\mathrm{k}}}\right)^{m+1}, \tag{3.2}$$

where

$$a_n = \frac{(-1)^n \Gamma(\beta)}{\Gamma(\beta - n) n!}, \quad W_m = \sum_{n=0}^{\infty} \frac{(-1)^{n+m}\, \Gamma(\gamma)\, \Gamma\{(n+1)\gamma\}}{\Gamma(\gamma - n)\Gamma\{(n+1)\gamma - m\} m!\, n!}.$$

3.2. Moments. Some key features of a distribution such as skewness and kurtosis can be studied through its moments. We derive closed-form expressions for the ordinary and central moments, generating function, skewness and kurtosis of X.

First, we introduce the Fox-Wright function ${}_p\Psi_q$, which is an extension of the usual generalized hypergeometric function ${}_pF_q$, with $p \in \mathbb{N}_0$ numerator parameters $a_1, \cdots, a_p \in \mathbb{C}$ and $q \in \mathbb{N}_0$ denominator parameters $b_1, \cdots, b_q \in \mathbb{C} \setminus \mathbb{Z}_0^-$, defined by

$${}_p\Psi_q \left[\begin{array}{c} (a_1, A_1), \cdots, (a_p, A_p) \\ \\ (b_1, B_1), \cdots, (b_q, B_q) \end{array} \middle| z \right] = \sum_{n \ge 0} \frac{\Gamma(a_1 + A_1 n) \cdots \Gamma(a_p + A_p n)}{\Gamma(b_1 + B_1 n) \cdots \Gamma(b_q + B_q n)} \frac{z^n}{n!},$$

where the empty products are conventionally taken to be equal to one, and

$$A_j > 0,\ j = \overline{1,p},\ B_k > 0,\ k = \overline{1,q}, \quad \Delta = 1 + \sum_{j=1}^{q} B_j - \sum_{j=1}^{p} A_j \ge 0,$$

(see, for instance [11, p. 56]). The convergence will occur for suitably bounded values of $|z|$ such that

$$|z| < \nabla = \left(\prod_{j=1}^{p} A_j^{-A_j} \right) \left(\prod_{j=1}^{q} B_j^{B_j} \right).$$

We derive closed-form expressions for the real order moments of X. We have

$$\begin{aligned} \mathsf{E}(X^r) &= \alpha\, \gamma \sum_{m=0}^{\infty} W_m \int_0^{\infty} x^r \left(\lambda + \beta\, k\, x^{k-1} \right) \mathrm{e}^{-\lambda\, (\mathrm{m}+1)\, \mathrm{x}}\, \mathrm{e}^{-\beta\, (\mathrm{m}+1)\, \mathrm{x}^{\mathrm{k}}}\, \mathrm{dx} \\ &= \alpha\, \gamma\, \lambda \sum_{m=0}^{\infty} W_m \int_0^{\infty} x^r\, \mathrm{e}^{-\lambda\, (\mathrm{m}+1)\, \mathrm{x}}\, \mathrm{e}^{-\beta\, (\mathrm{m}+1)\, \mathrm{x}^{\mathrm{k}}}\, \mathrm{dx} \\ &\quad + \alpha\gamma\beta k \sum_{m=0}^{\infty} W_m \int_0^{\infty} x^{r+k-1} \mathrm{e}^{-\lambda(\mathrm{m}+1)\mathrm{x}} \mathrm{e}^{-\beta(\mathrm{m}+1)\mathrm{x}^{\mathrm{k}}} \mathrm{dx}. \end{aligned}$$

The rth moment is a linear combination of integrals of the type $\mathcal{I}(\omega)$ based on a similar approach by [19, Eq. (2.1)], where $\omega = (\kappa, \mu, a, \eta)$ and all components are positive parameters,

$$\mathcal{I}(\omega) = \int_0^{\infty} x^{\kappa-1}\, \mathrm{e}^{-(\mu \mathrm{x} + \mathrm{a}\mathrm{x}^{\eta})}.$$

A representation for this integral is given by [21, p. 515, Corollary 1.1]:

$$\mathcal{I}(\omega) = \begin{cases} \mu^{-\kappa}\, {}_1\Psi_0 \left[\begin{array}{c} (\kappa, \eta) \\ \text{---} \end{array} \middle| -\frac{a}{\mu^{\eta}} \right], & 0 < \eta < 1, \\ \dfrac{\Gamma(\kappa)}{(\mu + a)^{\kappa}}, & \eta = 1, \\ \dfrac{1}{\eta a^{\kappa/\eta}}\, {}_1\Psi_0 \left[\begin{array}{c} \left(\dfrac{\kappa}{\eta}, \dfrac{1}{\eta} \right) \\ \text{---} \end{array} \middle| -\frac{\mu}{a^{1/\eta}} \right], & \eta > 1. \end{cases} \tag{3.3}$$

Thus, for all $k \in (0,1)$, we can write

$$
\begin{aligned}
\mathsf{E}(X^r) &= \alpha\,\gamma\,\lambda \sum_{m=0}^{\infty} W_m\, \mathfrak{I}(r+1, \lambda\,(m+1), \beta\,(m+1), k) \\
&\quad + \alpha\gamma\beta k \sum_{m=0}^{\infty} W_m\, \mathfrak{I}(r+k-1, \lambda(m+1), \beta(m+1), k) \\
&= \sum_{m=0}^{\infty} W_m \frac{\gamma\,\alpha}{\lambda^r\,(m+1)^{r+1}}\, {}_1\Psi_0\left[\begin{array}{c} (r+1,k) \\ \overline{} \end{array} \middle| -\frac{\beta}{\lambda^k\,(m+1)^{k-1}} \right] \\
&\quad + \sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\beta\,k}{(\lambda\,(m+1))^{r+k}}\, {}_1\Psi_0\left[\begin{array}{c} (r+k,k) \\ \overline{} \end{array} \middle| -\frac{\beta}{\lambda^k (m+1)^{k-1}} \right]. \qquad (3.4)
\end{aligned}
$$

For $k=1$, we have

$$
\mathsf{E}(X^r) = \frac{\lambda\,\alpha\,\gamma\,\Gamma(r+1)}{(\lambda+\beta)^{r+1}} \sum_{m=0}^{\infty} \frac{W_m}{(m+1)^{r+1}}. \qquad (3.5)
$$

The remaining values of the parameter $k>1$ lead to

$$
\begin{aligned}
\mathsf{E}(X^r) &= \sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\lambda}{k\{\beta(m+1)\}^{\frac{r+1}{k}}}\, {}_1\Psi_0\left[\begin{array}{c} \left(\frac{r+1}{k}, \frac{1}{k}\right) \\ \overline{} \end{array} \middle| -\frac{(m+1)^{1-\frac{1}{k}}\,\lambda}{\beta^{\frac{1}{k}}} \right] \\
&\quad + \sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\beta\,k}{\{\beta(m+1)\}^{\frac{r+k}{k}}}\, {}_1\Psi_0\left[\begin{array}{c} \left(\frac{r}{k}+1, \frac{1}{k}\right) \\ \overline{} \end{array} \middle| -\frac{(m+1)^{1-\frac{1}{k}}\,\lambda}{\beta^{\frac{1}{k}}} \right]. \qquad (3.6)
\end{aligned}
$$

Hence, we have the following result:

3.1. Theorem. *If $X \sim \mathrm{KwEW}_{\alpha,\gamma}(\lambda, \beta, k)$, then (for all $r > -1$) we have*

$$
(3.7) \qquad \mathsf{E}(X^r) = \begin{cases}
\sum_{m=0}^{\infty} W_m \frac{\gamma\,\alpha}{\lambda^r\,(m+1)^{r+1}} \\
\quad \times {}_1\Psi_0\left[\begin{array}{c} (r+1,k) \\ \hline \end{array} \middle| -\frac{\beta}{\lambda^k\,(m+1)^{k-1}} \right] \\
\quad + \sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\beta\,k}{\{\lambda\,(m+1)\}^{r+k}} \\
\quad \times\ {}_1\Psi_0\left[\begin{array}{c} (r+k,k) \\ \hline \end{array} \middle| -\frac{\beta}{\lambda^k(m+1)^{k-1}} \right], & 0 < k < 1, \\
\sum_{m=0}^{\infty} W_m \frac{\lambda\,\alpha\,\gamma\,\Gamma(r+1)}{(\lambda+\beta)^{r+1}\,(m+1)^{r+1}}, & k = 1, \\
\sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\lambda}{k\{\beta(m+1)\}^{\frac{r+1}{k}}} \\
\quad \times {}_1\Psi_0\left[\begin{array}{c} \left(\frac{r+1}{k}, \frac{1}{k}\right) \\ \hline \end{array} \middle| -\frac{(m+1)^{1-\frac{1}{k}}\,\lambda}{\beta^{\frac{1}{k}}} \right] \\
\quad + \sum_{m=0}^{\infty} W_m \frac{\alpha\,\gamma\,\beta\,k}{\{\beta(m+1)\}^{\frac{r+k}{k}}} \\
\quad \times {}_1\Psi_0\left[\begin{array}{c} \left(\frac{r}{k}+1, \frac{1}{k}\right) \\ \hline \end{array} \middle| -\frac{(m+1)^{1-\frac{1}{k}}\,\lambda}{\beta^{\frac{1}{k}}} \right], & k > 1.
\end{cases}
$$

Proof. It only remains to verify the convergence conditions of the Fox–Wright series, which depends only on the parameter k. Note that, when $k \in (0,1)$, $\Delta = 1 - k > 0$, so that both series in (3.4) converge. So, it does when $k = 1$. Finally, for $k > 1$, the value $\Delta = 1 - \frac{1}{k} > 0$ ensures that the moment $\mathsf{E}(X^r)$ is finite for any $r > -1$. □

3.2. Remark. For certain integer and rational values of the parameter k, we adopt a representation of the Fox–Wright ${}_1\Psi_0$ function in terms of the generalized hypergeometric ${}_pF_q$ functions, which is discussed in detail in [16]. By their Eq. (3.3), for all positive rational $A = \frac{m}{M}$, one has

$$
{}_1\Psi_0\left[\begin{array}{c} (a, \frac{m}{M}) \\ \hline \end{array} \middle| z \right] = \Gamma(a) + \sum_{j=1}^{M} \frac{\Gamma(a + \frac{m}{M}j)\,z^j}{j!} \times\ {}_{m+1}F_M\left[\begin{array}{c} 1, \frac{j}{M} + \frac{a}{m}, \cdots, \frac{j}{M} + \frac{a+m-1}{m} \\ \hline \end{array} \middle| \frac{m^m\,z^M}{M^M} \right],
$$

where ${}_pF_q$ stands for the generalized hypergeometric function which is a built–in *Mathematica* function specified by

```
HypergeometricPFQ[{a_1,\ldots, a_p},{b_1,\ldots, b_q},z].
```

On the other hand, the same authors also give an insight into transforming Fox–Wright Ψ functions into Meijer G–functions for rational arguments. Referring to [16, Eq. (5.1)], one has

$$
{}_1\Psi_0\left[\begin{array}{c}(a,\,\frac{m}{M})\\ \hline\end{array}\middle|\, z\right] = \frac{2\sqrt{M}\,m^a}{\Gamma(a)\,\sqrt{m}\,\pi^{\frac{M+m-1}{2}}} \times G^{M,m}_{m,M}\left(\frac{m^m\,(-z)^M}{M^M}\,\middle|\,\begin{array}{c}1-\frac{a}{m},\cdots,1-\frac{a+m-1}{m}\\ 0,\frac{1}{M},\cdots,\frac{M-1}{M}\end{array}\right).
$$

See, for example, the monographs [14, Ch. V] and [11] for an introduction to the G–function. ■

3.3. Remark. The nth factorial moment of order of X is given by

$$
\Phi_n = \mathsf{E}[X(X-1)(X-2)\cdots(X-n+1)] = \left.\frac{\mathrm{d}^n\left[\mathsf{E}(t^X)\right]}{\mathrm{d}t^n}\right|_{t=1}.
$$

Based on the Viète–Girard formula for expanding the polynomial $X(X-1)(X-2)\cdots(X-n+1)$, we obtain

$$
\Phi_n = \sum_{r=1}^{n}(-1)^{n-r}\left\{\sum_{1\le\ell_1<\cdots<\ell_r\le n-1}\ell_1\cdots\ell_r\right\}\mathsf{E}(X^r),
$$

where the second sum represents elementary symmetric polynomials:

$$
e_r = e_r(\ell_1,\cdots,\ell_r) = \sum_{1\le\ell_1<\cdots<\ell_r\le n-1}\ell_1\cdots\ell_r, \qquad r=\overline{0,n-1}.
$$

This in conjunction with positive integer rth order moment expression given in equation (3.7) provides an exact power series for the fractional order moments. ■

3.4. Remark. The moment generating function (mgf) $M(t) = E(\mathrm{e}^{\mathrm{t\,X}})$ of X can be obtained by setting $r=0$ and replacing $[\lambda\,(m+1)]$ by $[\lambda\,(m+1)-t]$ in equation (3.7).■

3.5. Remark. The central moments (μ_n) and cumulants (κ_n) of X are easily obtained from (3.7) as

$$
\mu_n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}\mu_1'^k\,\mu_{n-k}' \qquad\text{and}\qquad \kappa_n = \mu_n' - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\kappa_k\,\mu_{n-k}',
$$

respectively, where $\kappa_1 = \mu_1'$. Thus, $\kappa_2 = \mu_2' - \mu_1'^2$, $\kappa_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$, etc. Clearly, the skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships.■

Some numerical values for the skewness and kurtosis of X are listed in Table 2. The figures in this table indicate a large range for the skewness of X, although the kurtosis does not vary much.

Next, we discuss some other structural properties of X, i.e., survival, hazard rate, mean residual life, entropy, mean deviations and quantile function (qf).

3.3. Survival, Hazard rate, Quantile function, Skewness and Kurtosis. Central role is playing in the reliability theory by the ratio of the pdf and survival function. The survival function of X is given by

$$
S(x) = \left\{1-\left(1-\mathrm{e}^{-\lambda\,\mathrm{x}-\beta\,\mathrm{x}^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma}\,1_{\mathbb{R}_+}(x)\,. \tag{3.8}
$$

Then, the hrf of X reduces to

$$
h(x) = \frac{\alpha\gamma\left(kx^{-1+k}\beta+\lambda\right)\mathrm{e}^{-\lambda\,\mathrm{x}-\beta\,\mathrm{x}^{\mathrm{k}}}\left(1-\mathrm{e}^{-\lambda\,\mathrm{x}-\beta\,\mathrm{x}^{\mathrm{k}}}\right)^{-1+\alpha}}{\left\{1-\left(1-\mathrm{e}^{-\lambda\,\mathrm{x}-\beta\,\mathrm{x}^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma}}\,1_{\mathbb{R}_+}(x)\,. \tag{3.9}
$$

Table 2. Skewness and kurtosis of the KwEW distribution for selected parameter values.

λ	β	k	α	γ	Skewness	Kurtosis
1.0	2.6	1.5	10	30	-0.001	1.229
2.0	2.6	1.5	10	30	-0.001	1.229
3.0	2.6	1.5	10	30	-0.001	1.229
4.0	2.6	1.5	10	30	-0.001	1.229
–	–	–	–	–	–	–
1.3	0.5	2.5	25	18	-0.002	1.234
1.3	2.0	2.5	25	18	-0.002	1.234
1.3	4.0	2.5	25	18	-0.002	1.234
1.3	6.0	2.5	25	18	-0.002	1.234
–	–	–	–	–	–	–
0.2	3.4	1.0	2.0	3.0	0.150	1.251
0.2	3.4	1.5	2.0	3.0	0.150	1.251
0.2	3.4	2.0	2.0	3.0	0.150	1.251
0.2	3.4	3.0	2.0	3.0	0.150	1.251
–	–	–	–	–	–	–
0.7	0.7	2.0	0.2	5.0	0.914	5.283
0.7	0.7	2.0	1.2	5.0	0.218	1.275
0.7	0.7	2.0	1.8	5.0	0.149	1.245
0.7	0.7	2.0	10	5.0	0.049	1.238
–	–	–	–	–	–	–
3.5	1.6	3.0	5.0	0.5	0.190	1.306
3.5	1.6	3.0	5.0	1.0	0.146	1.277
3.5	1.6	3.0	5.0	1.5	0.123	1.263
3.5	1.6	3.0	5.0	2.0	0.108	1.254

Figures 3 (a), (b) and (c) display some plots of $h(x)$.

The qf of X is determined by inverting (1.5) as

$$Q(u) = F^{-1}(u) = -\frac{\log[1-\{1-(1-u)^{1/\gamma}\}^{1/\alpha}]}{\lambda+\beta}. \tag{3.10}$$

Simulating KwEW random variable is straightforward. Let U be a uniform variable on the unit interval $(0,1)$. Thus, by means of the inverse transformation method, the random variable X given by

$$X = -\frac{\log[1-\{1-(1-U)^{1/\gamma}\}^{1/\alpha}]}{(\lambda+\beta)} \tag{3.11}$$

follows the density (1.6). In particular, the median of X is

$$M = -\frac{\log[1-\{1-0.5^{1/\gamma}\}^{1/\alpha}]}{(\lambda+\beta)}.$$

Further, the mode of $f(x)$ is obtained as

$$MO = -\frac{\log\left\{1-\left(\frac{2-\alpha}{1-\alpha\,\gamma}\right)^{1/\alpha}\right\}}{(\lambda+\beta)}$$

The shortcomings of the classical kurtosis measure are well-known. There are many heavy tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. Indeed, our motivation to use quantile-based measures stemmed from the non-existence of the classical kurtosis for many of the Kw-G distributions. The Bowley's skewness is based on quartiles

$$S = \frac{Q(3/4)-2Q(1/2)+Q(1/4)}{Q(3/4)-Q(1/4)}$$

and the Moors' kurtosis is based on octiles

$$K = \frac{\{Q(7/8) - Q(5/8)\} + \{Q(3/8) - Q(1/8)\}}{Q(6/8) - Q(2/8)},$$

where $Q(\cdot)$ is given by (3.10).

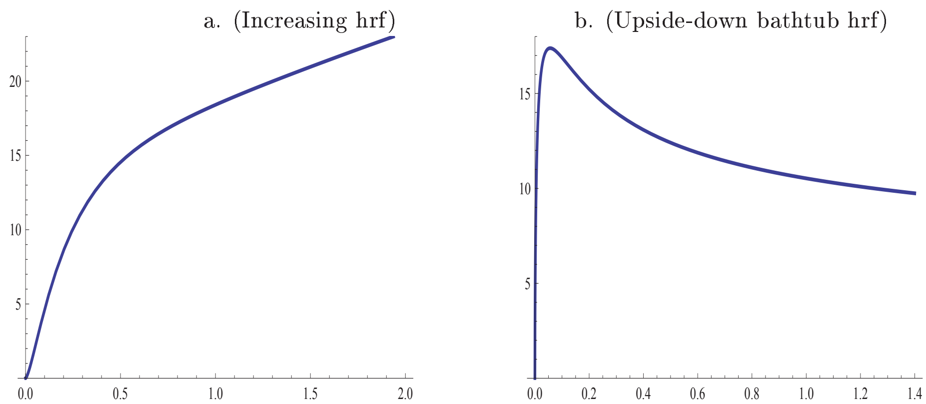


Figure 3. (a) $\alpha = 2.5$, $\gamma = 3$, $\lambda = 4$, $\beta = 1.2$, $k = 1.8$ (b) $\alpha = 2.3$, $\gamma = 1.3$, $\lambda = 1.8$, $\beta = 9$, $k = 0.7$.

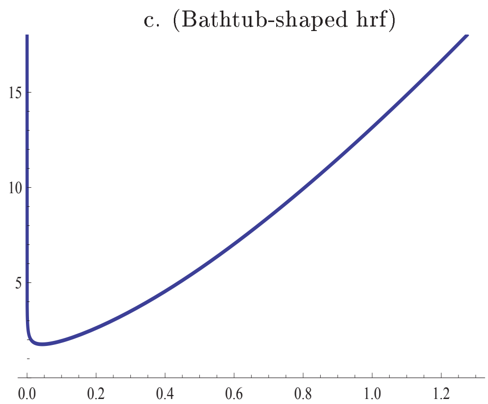


Figure 4. (c) $\alpha = 0.8$, $\gamma = 0.5$, $\lambda = 2.3$, $\beta = 10$, $k = 2.4$.

3.4. Mean residual life function. The mean residual life function (mrlf) is defined by

$$K(x) = \frac{1}{S(x)}\,[E(X) - m_1(x)] - x,$$

where $f(x)$, $E(X)$ and $S(x)$ are given in (1.6), (3.7) and (3.8), respectively, and

$$m_1(x) = \int_0^x y\, f(y)\,\mathrm{d}y = \alpha\,\gamma \sum_{m=0}^{\infty} W_m \times \int_0^x y\, \left(\lambda + \beta\, k\, y^{k-1}\right)\, \mathrm{e}^{-\lambda\,(\mathrm{m}+1)\,\mathrm{y}}\, \mathrm{e}^{-\beta\,(\mathrm{m}+1)\,\mathrm{y}^{\mathrm{k}}}\,\mathrm{dy}$$

is the first incomplete moment of X. By expanding the exponential in the last expression, we obtain

$$\begin{aligned} m_1(x) &= \alpha\,\gamma \sum_{m=0}^{\infty} W_m \sum_{j=0}^{\infty} \frac{(-1)^j\,[\lambda\,(m+1)]^j}{j!} \times \int_0^x y^{j+1}\, \left(\lambda + \beta\, k\, y^{k-1}\right)\, e^{-\beta\,(m+1)\,y^k}\,\mathrm{d}y \\ &= \alpha\,\gamma \sum_{m=0}^{\infty} W_m \sum_{j=0}^{\infty} \frac{(-1)^j\,[\lambda\,(m+1)]^j}{j!} \\ &\quad \times \Bigg(\lambda \int_0^x y^{j+1}\, G_{0,1}^{1,0}\left(\beta\,(m+1) y^{p/q} \,\middle|\, \begin{matrix} - \\ 0 \end{matrix}\right)\mathrm{d}y \\ &\quad + \beta\,\frac{p}{q} \int_0^x y^{j+p/q}\, G_{0,1}^{1,0}\left(\beta\,(m+1)\, y^{p/q} \,\middle|\, \begin{matrix} - \\ 0 \end{matrix}\right)\mathrm{d}y\Bigg), \end{aligned} \tag{3.12}$$

where $\mathrm{e}^{-\mathrm{g}(\mathrm{x})} = \mathrm{G}_{0,1}^{1,0}\left(\mathrm{g}(\mathrm{x}) \,\middle|\, \begin{matrix} - \\ 0 \end{matrix}\right)$, $k = p/q$ and $p \geq 1$ and $q \geq 1$ are natural co-prime numbers and

$$\begin{aligned} &\int_0^x y^t\, G_{0,1}^{1,0}\left(\beta\,(m+1)\, y^{p/q} \,\middle|\, \begin{matrix} - \\ 0 \end{matrix}\right)\mathrm{d}y \\ &= \frac{q\, x^{p\,(t+1)}}{p(2\,\pi)^{(q-1)/2}}\, G_{p,p+q}^{q,p}\left(\frac{\{\beta\,(m+1)\}^q\, x^p}{q^q} \,\middle|\, \begin{matrix} \frac{-t}{p}, \frac{1-t}{p}, \ldots, \frac{p-t-1}{p}, - \\ 0, \frac{-t-1}{p}, \frac{t}{p}, \ldots, \frac{p-t-2}{p} \end{matrix}\right). \end{aligned} \tag{3.13}$$

Equation (3.13) is obtained by using (13) of [5]. So, the first incomplete moment of X is easily obtained from (3.12) and (3.13).

Some applications of $m_1(x)$ refer to the Bonferroni and Lorenz curves of X defined, for a given probability π, by $B(\pi) = m_1(q)/(\pi\mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $\mu_1' = E(X)$ and $q = Q(\pi)$ is the value of (3.10) at $u = \pi$.

3.5. Entropy. An entropy is a concept encountered in physics and engineering. It is a measure of variation or uncertainty of a random variable X. An extension of Shannon's entropy for the continuous case can be defined as follows:

$$H(f) = -\int_0^{\infty} \log[f(x)]\, f(x)\,\mathrm{d}x\,. \tag{3.14}$$

Combining (1.6) and (3.14), we can write

$$
\begin{aligned}
H(f) = & -\alpha\,\gamma \sum_{m=0}^{\infty} W_m \,\log\left(\alpha\,\gamma \sum_{m=0}^{\infty} W_m\right) \\
& \times \int_0^{\infty} \left(\lambda + \beta\,k\,x^{k-1}\right)\, \mathrm{e}^{-\lambda\,(\mathrm{m}+1)\,\mathrm{x}}\ \mathrm{e}^{-\beta\,(\mathrm{m}+1)\ \mathrm{x}^{\mathrm{k}}}\, \mathrm{dx} \\
& - \alpha\,\gamma \sum_{m=0}^{\infty} W_m \\
& \times \int_0^{\infty} \left(\lambda + \beta\,k\,x^{k-1}\right) \log\left(\lambda + \beta\,k\,x^{k-1}\right) \mathrm{e}^{-\lambda\,(\mathrm{m}+1)\,\mathrm{x}}\,\mathrm{e}^{-\beta\,(\mathrm{m}+1)\ \mathrm{x}^{\mathrm{k}}}\mathrm{dx} \\
& + \lambda\,\alpha\,\gamma \sum_{m=0}^{\infty} (m+1)\,W_m \\
& \times \int_0^{\infty} x\,\left(\lambda + \beta\,k\,x^{k-1}\right)\, \mathrm{e}^{-\lambda\,(\mathrm{m}+1)\,\mathrm{x}}\ \mathrm{e}^{-\beta\,(\mathrm{m}+1)\ \mathrm{x}^{\mathrm{k}}}\, \mathrm{dx} \\
& + \beta\,\alpha\,\gamma \sum_{m=0}^{\infty} (m+1)\,W_m \\
& \times \int_0^{\infty} x^k\,\left(\lambda + \beta\,k\,x^{k-1}\right)\, \mathrm{e}^{-\lambda\,(\mathrm{m}+1)\,\mathrm{x}}\ \mathrm{e}^{-\beta\,(\mathrm{m}+1)\ \mathrm{x}^{\mathrm{k}}}\, \mathrm{dx}.
\end{aligned}
\tag{3.15}
$$

Note that the first, third and fourth integrals on the right–hand side of (3.15) can be determined by using (3.7) for $r = 0, 1$ and k, respectively. The second one can be evaluated by numerical integration.

3.6. Order statistics. Let $X_1, X_2, \ldots, X_n$ be a random sample from the KwEW distribution and $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ denote the corresponding order statistics. Let $f_{i:n}\,(x)$ and $F_{i:n}\,(x)$ denote, respectively, the pdf and the cdf of the ith order statistic $X_{i:n}$. We can write

$$
f_{i:n}\,(x) = \frac{n!f\,(x)}{(i-1)!\,(n-i)!} \sum_{l=0}^{n-i} \binom{n-i}{l} (-1)^l\ F\,(x)^{i-1+l}\,,
$$

and

$$
F_{i:n}\,(x) = \frac{n!}{(i-1)!\,(n-i)!} \sum_{l=0}^{n-i} \frac{(-1)^l}{i+l} \binom{n-i}{l} F\,(x)^{i+l}\,,
$$

where $F\,(x)$ and $f\,(x)$ are given by equations (1.5) and (1.6), respectively. Using (3.1) and after some algebra, we obtain

$$
f_{i:n}\,(x) = \frac{n!\,\alpha\,\gamma\,\left(\lambda + \beta\,k\,x^{k-1}\right)}{(i-1)!\,(n-i)!} \sum_{l=0}^{n-i}\sum_{u=0}^{\infty} \binom{n-i}{l} W_u\,\mathrm{e}^{-\lambda(\mathrm{u}+1)\,\mathrm{x}}\ \mathrm{e}^{-\beta(\mathrm{u}+1)\,\mathrm{x}^{\mathrm{k}}}
$$

and

$$
\begin{aligned}
F_{i:n}\,(x) = & \frac{n!}{(i-1)!\,(n-i)!} \sum_{l=0}^{n-i}\sum_{s=0}^{\infty} \binom{n-i}{l} \frac{\Gamma(i+l)\,(-1)^{l+s}}{\Gamma(i+l-s)s!\,(i+l)} \\
& \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma\,s},
\end{aligned}
$$

where

$$
W_u = \sum_{t=0}^{\infty}\sum_{s=0}^{\infty} \frac{(-1)^{l+s+t+u}\,\Gamma(i+l)\,\Gamma\{(s+1)\gamma\}\Gamma\{(t+1)\alpha\}}{\Gamma(i+l-s)\,\Gamma\{(s+1)\gamma - t\}\Gamma\{(t+1)\alpha - u\}s!\,t!\,u!}.
$$

The sth moment of $X_{i:n}$ is given by

$$E\left(X_{i:n}^{s}\right)=\int_{0}^{\infty} x^{s}\, f_{i:n}\left(x\right)\mathrm{d}x.$$

By using $f_{i:n}\left(x\right)$ and equation (3.3), the moments of $X_{i:n}$ can be easily obtained.

3.7. Mean deviations. The mean deviations provide important information about characteristics of a population and they can be calculated from the first incomplete moment. Indeed, the amount of dispersion in a population may be measured to some extent by the deviations from the mean and median. The mean deviations of X about the mean $\mu_1'=E(X)$ and about the median M can be expressed as $\delta_1=2\mu F(\mu_1')-2m_1(\mu_1')$ and $\delta_2=\mu_1'-2m_1(M)$, where $F(\mu_1')$ is calculated from (1.5) and $m_1(z)=\int_0^z x\, f(x)dx$ can be determined from (3.12) and (3.13).

4. Alternative Properties

In this section, we provide an alternative mixture representation for the pdf of X. By combining (1.4) and (3.2), we can write

$$f(x)=\sum_{m=0}^{\infty} V_m\, g_{m+1}(x), \tag{4.1}$$

where (for $m\geq 0$) $V_m=\alpha\gamma\, W_m/(m+1)$ and $g_{m+1}(x)$ is the pdf of the EW model with parameters $\lambda^\star=(m+1)\lambda$, $\beta^\star=(m+1)\beta$ and k. So, the KwEW density function is a mixture of EW densities.

Based on equation (4.1) and the results by Cordeiro *et al.* [5], we can obtain the following properties of X.

4.1. Moments. The calculations in this section involve some special functions. In particular, the gamma function $\Gamma(r)=\int_0^\infty w^{r-1}\mathrm{e}^{-\mathrm{w}}\mathrm{dw}$ $(r>0)$, and other functions given in Appendices A and B. In order to obtain μ_s', we require an integral of the type

$$I(s;\lambda^\star,\beta^\star,k)=\int_{0}^{\infty} x^{s}\,\mathrm{e}^{-(\lambda^\star \mathrm{x}+\beta^\star \mathrm{x}^{\mathrm{k}})}\,\mathrm{dx}. \tag{4.2}$$

We provide four representations for (4.2). First, by expanding $\mathrm{e}^{-\lambda^\star \mathrm{x}}$ in Taylor series, we obtain

$$\begin{aligned}I(s;\lambda^\star,\beta^\star,k)&=\sum_{j=0}^{\infty}\frac{(-\lambda^\star)^j}{j!}\int_0^\infty x^{s+j}\,\mathrm{e}^{-\beta^\star \mathrm{x}^{\mathrm{k}}}\mathrm{dx}\\&=\frac{1}{k\beta^{\star(s+1)/k}}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\left(\frac{\lambda^\star}{\beta^{\star 1/k}}\right)^j\Gamma\left(\frac{s+1+j}{k}\right).\end{aligned}$$

The above sum can be expressed in a simple form for $k>1$ using the Fox–Wright generalized hypergeometric function defined in Appendix A. We have

$$I(s;\lambda^\star,\beta^\star,k)=\frac{1}{k\beta^{\star(s+1)/k}}\,{}_1\Psi_0\left[\begin{array}{c}\left(\frac{s+1}{k},\frac{1}{k}\right)\\ -\end{array};-\frac{\lambda^\star}{\beta^{\star 1/k}}\right]. \tag{4.3}$$

Applying (4.3) to (4.1), we can write

$$\mu_s'=E(X^s)=\sum_{m=0}^{\infty}V_m\left[\lambda^\star\, I(s;\lambda^\star,\beta^\star,k)+\beta^\star\, k\, I(s+\lambda^\star-1;\lambda^\star,\beta^\star,k)\right]. \tag{4.4}$$

Secondly, we offer two formulae for the integral (4.2) provided that $k=p/q$, where $p\geq 1$ and $q\geq 1$ are relatively natural co-prime numbers. We use equation (2.3.2.13) in [26, p. 321] to obtain formulae for $I(s;\lambda^\star,\beta^\star,k)$ when $0<k<1$ and $k>1$. We exclude the case $k = 1$ since the model is non-identifiable. For irrational k, an approximation of vanishingly small error can be made using increasingly accurate rational approximations for k. Let $z = (p^p \beta^{\star q})/(q^q \lambda^{\star p})$, ${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; x)$ be the well-known generalized hypergeometric function and $\Delta(\tau, a) = (a/\tau, (a+1)/\tau, \ldots, (a+\tau-1)/\tau)$. The generalized hypergeometric functions are available in *Mathematica*. For $0 < k < 1$, we obtain

$$\begin{aligned} I(s; \lambda^\star, \beta^\star, k) = &\sum_{j=0}^{q-1} \frac{(-\beta^\star)^j\, \Gamma(s+1+jp/q)}{\lambda^{\star(s+1+jp/q)}\, j!} \\ &\times {}_{p+1}F_q\left(1, \Delta(p, s+1+jp/q); \Delta(q, 1+j); (-1)^q z\right). \end{aligned} \tag{4.5}$$

For $\gamma > 1$, we have

$$\begin{aligned} I(s; \lambda^\star, \beta^\star, k) = &\sum_{j=0}^{p-1} \frac{(-1)^j\, q\, \Gamma\left([s+1+j]q/p\right)}{p\, \beta^{\star(s+1+j)q/p}\, j!} \\ &\times {}_{q+1}F_p\left(1, \Delta(q, [s+1+j]q/p); \Delta(p, 1+j); \frac{(-1)^p}{z}\right). \end{aligned} \tag{4.6}$$

A fourth representation for the integral (4.2) also holds when $k = p/q$, where $p \geq 1$ and $q \geq 1$ are natural co-prime numbers. It follows in terms of the Meijer $G_{p,q}^{m,n}$ function defined in Appendix B and also available in *Mathematica*. For an arbitrary function $g(\cdot)$, we use the result

$$\exp\{-g(x)\} = G_{0,1}^{1,0}\left(g(x) \Big| \begin{matrix} - \\ 0 \end{matrix}\right), \tag{4.7}$$

and then equation (4.2) can be expressed in the same form of equation (2.24.3.1) given by [26, p. 350]. Hence, we obtain

$$I(s; \lambda^\star, \beta^\star, k) = \frac{p^{s+1/2}}{(2\pi)^{(p+q)/2-1}\, \lambda^{\star s+1}}\, G_{p,q}^{q,p}\left(\frac{\beta^{\star q}\, p^p}{\lambda^{\star p}\, q^q} \Bigg| \begin{matrix} \frac{-s}{p}, \frac{1-s}{p}, \ldots, \frac{p-s-1}{p} \\ 0 \end{matrix}\right). \tag{4.8}$$

Further, if $q = 1$, using equation (9.31.2) in [10]

$$G_{p,q}^{m,n}\left(z^{-1} \Big| \begin{matrix} a_r \\ b_s \end{matrix}\right) = G_{q,p}^{n,m}\left(z \Big| \begin{matrix} 1-b_s \\ 1-a_r \end{matrix}\right),$$

we have, as a special case of (4.8), the following result [3]

$$I(s; \lambda^\star, \beta^\star, k) = \frac{p^{s+1/2}}{(2\pi)^{(p-1)/2}\, \lambda^{\star s+1}}\, G_{1,p}^{p,1}\left(\frac{\lambda^{\star p}}{\beta^\star\, p^p} \Bigg| \begin{matrix} 1 \\ \frac{(s+1)}{p}, \frac{(s+2)}{p}, \ldots, \frac{(s+p)}{p} \end{matrix}\right).$$

Equations (4.3), (4.4), (4.5), (4.6) and (4.8) are the main results of this section.

4.2. Incomplete Moments. For lifetime models, it is useful to obtain the sth incomplete moment of X given by $T_s(y) = \int_0^y x^s\, f(x)dx$. We define $J(s,a) = J(s,a;\beta,\gamma) = \int_0^a x^s\, \mathrm{e}^{-\beta x^\gamma}\mathrm{dx}$. Moreover, it is simple to verify from (1.6) that $T_s(y)$ can be expressed as

$$T_s(y) = \int_0^y x^s\, (\lambda^\star + \beta^\star\, k\, x^{k-1})\, \mathrm{e}^{-(\lambda^\star x + \beta^\star x^k)}\mathrm{dx}.$$

By expanding the exponential in the last expression, we have

$$T_s(y) = \sum_{j=0}^{\infty} \frac{(-1)^j \lambda^{\star j}}{j!} \left[\lambda^\star\, J(s+j, y) + \beta^\star k\, J(s+k-1, y)\right]. \tag{4.9}$$

We now provide a formula for $T_s(y)$ in terms of the Meijer $G_{p,q}^{m,n}$ function (see Appendix B) which holds only when $k = p/q$, where $p \geq 1$ and $q \geq 1$ are natural co-prime numbers. By using (4.7), we can write

$$J(s,a) = \int_0^a x^s\, G_{0,1}^{1,0}\left(\beta^\star x^{p/q} \mid \begin{matrix} - \\ 0 \end{matrix}\right) \mathrm{d}x.$$

By using equation (2.24.2.2) in [26, p. 348], we can express $J(s,a)$ as

$$J(s,a) = \frac{q\, a^{p(s+1)}}{p\,(2\pi)^{(q-1)/2}}\, G_{p,p+q}^{q,p}\left(\frac{\beta^{\star q}\, a^p}{q^q} \middle| \begin{matrix} \frac{-s}{p}, \frac{1-s}{p}, \ldots, \frac{p-s-1}{p}, - \\ 0, \frac{-s-1}{p}, \frac{s}{p}, \ldots, \frac{p-s-2}{p} \end{matrix}\right). \tag{4.10}$$

Combining equations (4.9) and (4.10), we obtain the incomplete moments of X.

4.3. Generating Function. For $t < \lambda^\star$, the mgf of X follows from (4.1) as

$$M(t) = \sum_{m=0}^{\infty} V_m\, I(s; \lambda^\star - t, \beta^\star, k).$$

Thus, we can use the results in Section 4.1 to obtain an explicit expression for $M(t)$

$$M(t) = \sum_{m=0}^{\infty} V_m \left[\frac{1}{k\beta^{\star (s+1)/k}} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \left(\frac{\lambda^\star - t}{\beta^{\star 1/k}}\right)^j \Gamma\left(\frac{s+1+j}{k}\right)\right].$$

5. Parameter Estimation

5.1. Maximum likelihood estimation. Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimates (MLEs) enjoy desirable properties and can be used when constructing confidence intervals and test statistics. Large sample theory for these estimates delivers simple approximations that work well in finite samples. However, we can approximate quantities such as the density of test statistics that depend on the sample size in order to obtain better approximation for the MLEs, which can be easily handled either analytically or numerically.

Let $\theta = (\lambda, \beta, k, \alpha, \gamma)$ be the parameter vector of the KwEW distribution. The log-likelihood for θ given the data set $x_1, \ldots, x_n$ obtained from (1.6) is given by

$$\begin{aligned} \ell(\theta) = n\,[\log(\alpha) + \log(\gamma)] &+ \sum_{i=1}^{n} \log\left(\mathrm{e}^{-\,\mathrm{x_i^k}\beta - \mathrm{x_i}\lambda}\right) \\ &- (1-\alpha)\sum_{i=1}^{n} \log\left(1 - \mathrm{e}^{-\mathrm{x_i^k}\beta - \mathrm{x_i}\lambda}\right) + \sum_{i=1}^{n} \log\left(k x_i^{k-1}\beta + \lambda\right) \\ &- (1-\gamma)\sum_{i=1}^{n} \log\left\{1 - \left(1 - \mathrm{e}^{-\mathrm{x_i^k}\beta - \mathrm{x_i}\lambda}\right)^{\alpha}\right\}. \end{aligned} \tag{5.1}$$

The associated nonlinear log-likelihood equations $\frac{\partial \ell(\theta)}{\partial \theta}=0$ are given by

$$\frac{\partial \ell(\theta)}{\partial \lambda}=\sum_{i=1}^{n}-x_i+(\alpha-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\mathrm{x_i}}{1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}}+\sum_{i=1}^{n}\frac{1}{\lambda+k\beta x_i^{-1+k}}$$
$$-(\gamma-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{-1+\alpha}\alpha \mathrm{x_i}}{1-\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}}=0,$$

$$\frac{\partial \ell(\theta)}{\partial \beta}=\sum_{i=1}^{n}-x_i^k+(\alpha-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\mathrm{x_i^k}}{1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}}+\sum_{i=1}^{n}\frac{kx_i^{-1+k}}{\lambda+k\beta x_i^{-1+k}}$$
$$-(\gamma-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{-1+\alpha}\alpha \mathrm{x_i^k}}{1-\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}}=0,$$

$$\frac{\partial \ell(\theta)}{\partial k}=\sum_{i=1}^{n}-\beta\log(x_i)\,x_i^k+(\alpha-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\beta\log(\mathrm{x_i})\,\mathrm{x_i^k}}{1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}}$$
$$-(\gamma-1)\sum_{i=1}^{n}\frac{\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{-1+\alpha}\alpha\beta\log(\mathrm{x_i})\,\mathrm{x_i^k}}{1-\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}}$$
$$+\sum_{i=1}^{n}\frac{\beta\,x_i^{-1+k}+k\,\beta\,\log(x_i)\,x_i^{-1+k}}{\lambda+k\,\beta\,x_i^{-1+k}}=0,$$

$$\frac{\partial \ell(\theta)}{\partial \alpha}=\frac{n}{\alpha}+\sum_{i=1}^{n}\log\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)-(\gamma-1)$$
$$\times\sum_{i=1}^{n}\frac{\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}\log\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)}{1-\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}}=0\,,$$

$$\frac{\partial \ell(\theta)}{\partial \gamma}=\frac{n}{\gamma}+\sum_{i=1}^{n}\mathrm{Log}\left\{1-\left(1-\mathrm{e}^{-\lambda \mathrm{x_i}-\beta \mathrm{x_i^k}}\right)^{\alpha}\right\}=0. \tag{5.2}$$

For estimating the model parameters, numerical iterative techniques should be employed to solve these equations. We can investigate the global maximum of the log-likelihood by setting different starting values for the parameters. The information matrix will be required for interval estimation. The elements of the 5×5 total observed information matrix $J(\theta)=\{J_{rs}(\theta)\}$ (for $r,s=\lambda,\beta,k,\alpha,\gamma$) can be obtained from the authors upon request. The asymptotic distribution of $(\hat{\theta}-\theta)$ is $N_5(O,K(\theta)^{-1})$, under standard regularity conditions, where $K(\theta)=E\{J(\theta)\}$ is the expected information matrix and $J(\hat{\theta})$ is the observed information matrix evaluated at $\hat{\theta}$. The multivariate normal $N_5(O,J(\hat{\theta})^{-1})$ distribution can be used to construct approximate confidence intervals for the individual parameters.

5.2. Bayesian analysis. In the Bayesian approach, the information referring to the model parameters is obtained through a posterior marginal distribution. Here, we use the simulation method of Markov Chain Monte Carlo (MCMC) by the Metropolis-Hastings algorithm. Since we have no prior information from historical data or from previous experiment, we assign conjugate but weakly informative prior distributions to the parameters. We assume informative (but weakly) prior distribution and then the posterior distribution is a well-defined proper distribution. We also assume that the elements of the parameter vector are independent and that the joint prior distribution for all unknown parameters has a pdf given by

$$\pi(\lambda, \beta, k, \alpha, \gamma) \propto \pi(\lambda) \times \pi(\beta) \times \pi(k) \times \pi(\alpha) \times \pi(\gamma). \tag{5.3}$$

Here, $\lambda \sim \Gamma(a_1, b_1)$, $\beta \sim \Gamma(a_2, b_2)$, $k \sim \Gamma(a_3, b_3)$, $\alpha \sim \Gamma(a_4, b_4)$ and $\gamma \sim \Gamma(a_5, b_5)$, where $\Gamma(a_i, b_i)$ denotes a gamma distribution with mean a_i/b_i, variance a_i/b_i^2 and density function given by

$$f(\upsilon; a_i, b_i) = \frac{b_i^{a_i} \upsilon^{a_i - 1} \exp(-\upsilon b_i)}{\Gamma(a_i)},$$

where $\upsilon > 0$, $a_i > 0$ and $b_i > 0$. All hyper-parameters are specified. Combining the likelihood function (5.1) and the prior distribution (5.3), the joint posterior distribution for λ, β, k, α and γ reduces to

$$\begin{aligned}\pi(\lambda, \beta, k, \alpha, \gamma | x) \propto{}& (\alpha\gamma)^n \, \mathrm{e}^{-\lambda \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i} - \beta \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i^k}} \prod_{\mathrm{i}=1}^{\mathrm{n}} \Bigg\{ \left(\lambda + \mathrm{k}\beta \mathrm{x_i^{k-1}}\right) \\ & \times \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{-1+\alpha} \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{-1+\gamma} \Bigg\} \\ & \times \pi(\lambda, \beta, k, \alpha, \gamma). \end{aligned} \tag{5.4}$$

The joint posterior density above is analytically intractable because the integration of the joint posterior density is not easy to perform. In this direction, we first obtain the full conditional distributions of the unknown parameters given by

$$\begin{aligned}\pi(\lambda | x, \beta, k, \alpha, \gamma) \propto{}& \mathrm{e}^{-\lambda \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i}} \prod_{\mathrm{i}=1}^{\mathrm{n}} \Bigg\{ \left(\lambda + \mathrm{k}\beta \mathrm{x^{k-1}}\right) \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{-1+\alpha} \\ & \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{-1+\gamma} \Bigg\} \times \pi(\lambda),\end{aligned}$$

$$\begin{aligned}\pi(\beta | x, \lambda, k, \alpha, \gamma) \propto{}& \mathrm{e}^{-\beta \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i^k}} \prod_{\mathrm{i}=1}^{\mathrm{n}} \Bigg\{ \left(\lambda + \mathrm{k}\beta \mathrm{x_i^{k-1}}\right) \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{-1+\alpha} \\ & \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{-1+\gamma} \Bigg\} \times \pi(\beta), \\ \pi(k | x, \lambda, \beta, \alpha, \gamma) \propto{}& \mathrm{e}^{-\beta \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{x_i^k}} \prod_{\mathrm{i}=1}^{\mathrm{n}} \Bigg\{ \left(\lambda + \mathrm{k}\beta \mathrm{x_i^{k-1}}\right) \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{-1+\alpha} \\ & \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{-1+\gamma} \Bigg\} \times \pi(k), \\ \pi(\alpha | x, \lambda, \beta, k, \gamma) \propto{}& \alpha^n \prod_{i=1}^{n} \Bigg\{ \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha} \\ & \times \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{-1+\gamma} \Bigg\} \times \pi(\alpha)\end{aligned}$$

and

$$\pi(\gamma | x, \lambda, \beta, k, \alpha) \propto \gamma^n \prod_{i=1}^{n} \left\{1 - \left(1 - \mathrm{e}^{-\lambda \mathrm{x_i} - \beta \mathrm{x_i^k}}\right)^{\alpha}\right\}^{\gamma} \times \pi(\gamma).$$

Since the full conditional distributions for λ, β, k, α and γ do not have explicit expressions, we require the use of the Metropolis-Hastings algorithm.

Table 3. Empirical means and the RMSEs in parentheses for $\beta = 1$

n	$\hat{\lambda}$	$\hat{\beta}$	$\hat{k}$	$\hat{\alpha}$	$\hat{\gamma}$
		$\lambda = 2.3$	$k = 1.6$	$\alpha = 1.5$	$\gamma = 1$
250	1.218	1.689	1.467	1.702	1.346
	(0.954)	(0.517)	(0.978)	(1.597)	(1.510)
350	1.214	1.574	1.428	1.503	1.335
	(0.897)	(0.501)	(0.834)	(1.453)	(1.478)
450	1.213	1.572	1.346	1.548	1.217
	(0.895)	(0.498)	(0.740)	(1.404)	(1.156)
		$\lambda = 3.4$	$k = 1.8$	$\alpha = 2$	$\gamma = 2.3$
250	1.414	1.023	1.101	2.471	2.601
	(1.221)	(0.742)	(0.456)	(2.102)	(2.102)
350	1.367	1.367	1.084	2.495	2.495
	(0.918)	(0.904)	(0.285)	(2.104)	(2.104)
450	1.278	1.278	1.053	2.348	2.348
	(1.012)	(0.843)	(0.324)	(1.945)	(1.945)
		$\lambda = 0.4$	$k = 2$	$\alpha = 2.5$	$\gamma = 1.4$
250	2.203	1.146	2.142	2.104	1.925
	(0.962)	(0.765)	(0.978)	(1.231)	(1.024)
350	2.458	1.107	2.154	2.116	1.823
	(0.784)	(0.452)	(0.450)	(1.114)	(0.978)
450	1.067	1.047	2.045	2.123	1.450
	(0.452)	(0.596)	(0.258)	(1.080)	(0.856)
		$\lambda = 3.2$	$k = 2.5$	$\alpha = 1.5$	$\gamma = 3$
250	1.854	1.256	1.478	1.149	1.853
	(0.927)	(0.451)	(0.301)	(0.856)	(1.420)
350	1.745	1.024	1.201	1.131	1.741
	(0.847)	(0.237)	(0.214)	(0.723)	(1.204)
450	1.680	1.345	1.635	1.085	1.658
	(0.784)	(0.478)	(0.481)	(0.456)	(1.004)

5.3. Simulation study. We also assess the performance of the MLEs in terms of the sample size n. The simulation is performed using the Ox matrix programming language. The number of Monte Carlo replications is $10,000$. For maximizing the log-likelihood function, we use the MaxBFGS subroutine with analytical derivatives. The evaluation of the estimates is performed based on the following quantities for each sample size: the empirical mean squared errors (MSEs) and the root MSEs (RMSEs) using the R package from the Monte Carlo replications. The inversion method is used to generate samples, i.e., the variates having the KwEW distribution are generated using (3.10). The MLEs are evaluated for each simulated data, say $(\hat{\lambda}_i, \hat{\beta}_i, \hat{k}_i, \hat{\alpha}_i, \hat{\gamma}_i)$ (for $i = 1, \ldots, 10,000$) and the biases and MSEs are computed by

$$bias_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h_i} - h) \quad \text{and} \quad MSE_h(n) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{h_i} - h)^2$$

for $h = \lambda,\ \beta,\ k,\ \alpha,\ \gamma$.
Let the sample size be $n = 250$, 350 and 450 and consider different values for the shape parameters λ, k, α and γ, whereas the scale parameter β is fixed at one. The empirical results are given in Table 3.

Table 4. MLEs of the parameters (standard errors in parentheses) for the Aarset data

Distributions	Estimates				
Weibull(k , λ)	3.441197	47.05054			
	(0.000248)	(0.036047)			
E-W(λ, β, k)	0.018620	0.040483	0.373635		
	(0.003771)	(0.031143)	(0.188693)		
ExtW(a , b , c)	0.027836	0.942137	0.020278		
	(0.033196)	(0.285026)	(0.319463)		
MO-EW(a , b , c , α)	0.027083	0.161829	0.328829	3.599999	
	(0.006184)	(0.124196)	(0.143844)	(1.87102)	
Kw-W(a , b , c , λ)	0.340211	0.145696	1.209999	0.089756	
	(0.201699)	(0.106772)	(0.294355)	(0.079873)	
KwEW(λ , β , k , α , γ)	0.004366	0.209999	0.116764	3.516432	18.99999
	(0.001879)	(0.175644)	(0.057365)	(1.61287)	(15.3596)

The figures in this table indicate that the estimates are quite stable and, more importantly, are close to the true values for these sample sizes. Additionally, as the sample size increases, the RMSEs decrease as expected. We can conclude that the MLEs are robust.

6. Application

Here, we prove the potentiality of the KwEW distribution by means of a real data set using both MLEs and Bayesian approaches.

6.1. The MLEs approach. By using MLEs method, we fit the two-parameter Weibull (Weibull), exponential-Weibull (EW) [5], extended Weibull (ExtW) [20], Marshall-Olkin exponential-Weibull (MO-EW) [22], Kumaraswamy Weibull (Kw-W) [4] and KwEW distributions to the Aarset data [1] on lifetimes of 50 components, which possess a bathtub-shaped failure rate property. The density functions of these models are given below (for $x > 0$):

- The Weibull density function
$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \mathrm{e}^{-(\mathrm{x}/\lambda)^{\mathrm{k}}},\ \mathrm{k} > 0,\ \lambda > 0;$$
- The EW density function
$$f(x) = \left(\lambda + \beta\, k\, x^{k-1}\right) \mathrm{e}^{-\lambda\,\mathrm{x} - \beta\,\mathrm{x}^{\mathrm{k}}},\ \ \lambda, \beta, \mathrm{k} > 0;$$
- The ExtW density function
$$f(x) = a\,(c + b\,x)\,x^{-2+b}\,\mathrm{e}^{-\mathrm{c}/\mathrm{x} - \mathrm{a}\mathrm{x}^{\mathrm{b}}\mathrm{e}^{-\mathrm{c}/\mathrm{x}}},\ \ \mathrm{a}, \mathrm{b} > 0,\ \mathrm{c} \geq 0;$$
- The MO-EW density function
$$f(x) = \frac{\alpha\left(a + b\,c\,x^{-1+c}\right)\,\mathrm{e}^{-(\mathrm{ax}+\mathrm{bx}^{\mathrm{c}})}}{\left[1 - (1-\alpha)\,\mathrm{e}^{-(\mathrm{ax}+\mathrm{bx}^{\mathrm{c}})}\right]^2},\ \ \lambda, \beta, k, \alpha > 0;$$
- The Kw-W density function
$$f(x) = a\,b\,c\,\lambda^c\,x^{c-1}\,\mathrm{e}^{-(\mathrm{x}\lambda)^{\mathrm{c}}}\,\left\{1 - \mathrm{e}^{-(\mathrm{x}\lambda)^{\mathrm{c}}}\right\}^{\mathrm{a}-1}\left[1 - \left\{1 - \mathrm{e}^{-(\mathrm{x}\lambda)^{\mathrm{c}}}\right\}^{\mathrm{a}}\right]^{\mathrm{b}-1},$$
$$a, b, c, \lambda > 0.$$

The parameters of the above distributions are estimated by maximizing the log-likelihoods using the *NMaximize* command in the symbolic computational package *Mathematica.* Table 4 lists the MLEs (and the corresponding standard errors in parentheses) of the parameters. Table 5 gives the values of minus the maximized log-likelihood $(-\hat{\ell})$,

Table 5. Goodness-of-fit statistics for the Aarset data

Distributions	$-\hat{\ell}$	AIC	BIC	A^*	W^*
Weibull(k, λ)	240.98	485.959	489.783	3.53566	0.532984
E-W(λ, β, k)	239.463	484.927	490.663	2.92873	0.513036
ExtW(a, b, c)	240.957	487.914	493.65	3.5425	0.53549
MO-EW(a, b, c, α)	235.515	479.03	486.678	2.21706	0.34524
Kw-W(a, b, c, λ)	235.925	479.851	487.499	2.48043	0.424629
KwEW(λ, β, k, α, γ)	**233.087**	**476.175**	**485.735**	**2.11894**	**0.32768**

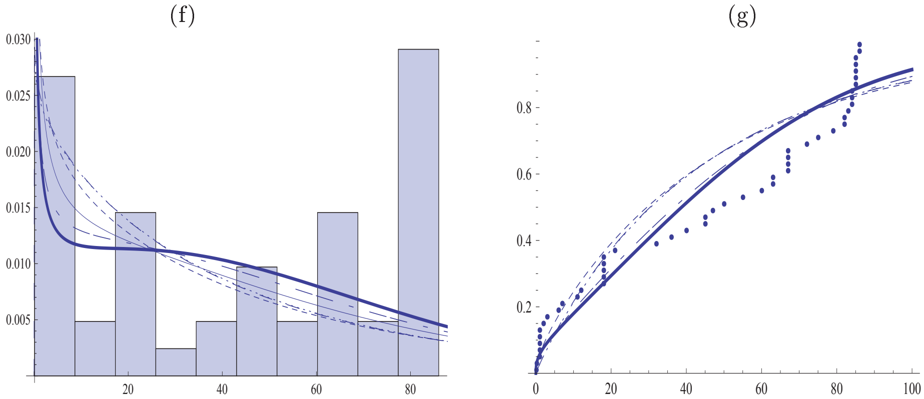


Figure 5. (f) The estimated KwEW density superimposed on the histogram for the Aarset data with other models. (g) The empirical cdf and the estimated cdf's of other models, where Kw-Ew is represented by (Thick line), Kw-W by (Thin line), MO-EW by (Long and short dashed line), ExtW by (Long dashed line), E-W by (dashed line) and Weibull by (Dotted line)

Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Anderson-Darling (A^*) and Cramér-von Mises (W^*) goodness-of-fit statistics for some fitted models. Since the values of these statistics are smaller for the KwEW distribution compared to those values of the Weibull, EW, ExtW, MO-EW and Kw-W distributions, the proposed distribution is a very competitive model for lifetime data analysis. Plots of the fitted KwEW, Weibull, E-W, ExtW, Mo-EW and Kw-W densities and the histogram of the data are displayed in Figure 5(f). In Figure 5(g), we plot the empirical cumulative function and the estimated cdf's for the KwEW and other distributions, which shows a satisfactory fit of the new model.

6.2. Bayesian approach. The following independent priors are considered to perform the Metropolis-Hastings algorithm: $\lambda \sim \Gamma(0.01, 0.01)$, $\beta \sim \Gamma(0.01, 0.01)$, $k \sim \Gamma(0.01, 0.01)$, $\alpha \sim \Gamma(0.01, 0.01)$ and $\gamma \sim \Gamma(0.01, 0.01)$, so that we have vague prior distributions. Considering these prior density functions, we generate two parallel independent runs of the Metropolis-Hastings with size 150,000 for each parameter, disregarding the first 15.000 iterations to eliminate the effect of the initial values and, to avoid correlation problems, we consider a spacing of size 10, obtaining a sample of size 13,500 from each chain. To monitor the convergence of the Metropolis-Hastings, we perform the methods suggested by Cowles and Carlin [7]. To monitor the convergence of the Metropolis-Hastings, we use

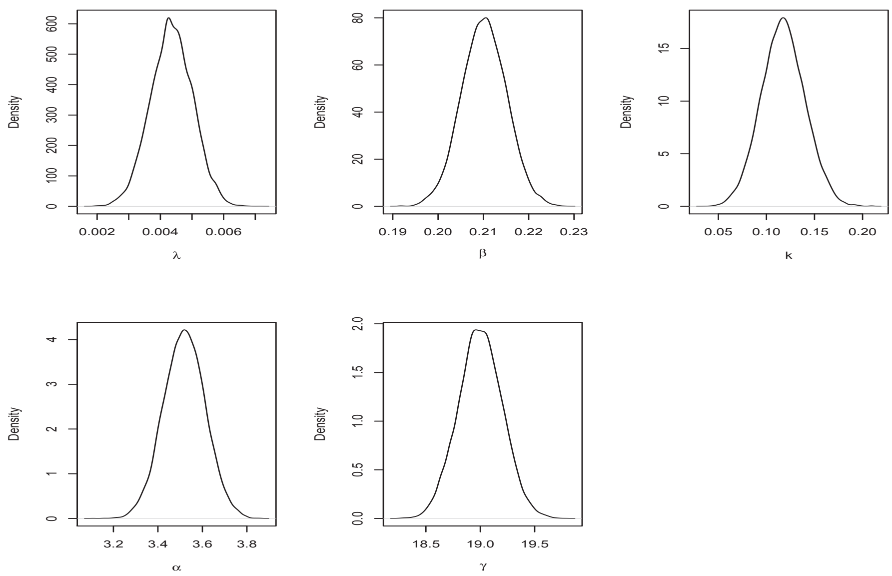


Figure 6. Approximate posterior marginal densities for the parameters from the KwEW model for the Aarset data.

the between and within sequence information, following the approach developed in Gelman and Rubin [9], to obtain the potential scale reduction, $\widehat{R}$. In all cases, these values were close to one, indicating the convergence of the chain. The approximate posterior marginal density functions for the parameters are presented in Figure 6. In Table 6, we report posterior summaries for the parameters of the new model. We note that the values for the means a posteriori (Table 6) are quite close (as expected) to the MLEs given in Table 5. Here, SD represents the standard deviation from the posterior distributions of the parameters and HPD represents the 95% highest posterior density (HPD) intervals.

Table 6. Posterior summaries for the parameters from the KwEW model for the Aarset data.

Parameter	Mean	SD	HPD (95%)	$\hat{R}$
λ	0.0044	0.0007	(0.0031; 0.0057)	1.0052
β	0.2102	0.0050	(0.2005; 0.2200)	1.0002
k	0.1175	0.0227	(0.0740; 0.1630)	1.0018
α	3.5188	0.0934	(3.3338; 3.7012)	0.9999
γ	19.0003	0.2027	(18.6049; 19.3974)	1.0008

7. Bivariate KwEW Distribution

Suppose $U_1 \sim \text{KwEW}(\gamma_1, \alpha, \lambda, \beta, k)$, $U_2 \sim \text{KwEW}(\gamma_2, \alpha, \lambda, \beta, k)$ and $U_3 \sim \text{KwEW}(\gamma_3, \alpha, \lambda, \beta, k)$ are independently distributed. Define $X_1 = max(U_1, U_3)$ and

$X_2 = max(U_2, U_3)$. Then the bivariate vector $(X_1, X_2) \sim$ KwEW $(\gamma_1, \gamma_2, \gamma_3, \alpha, \lambda, \beta, k)$.

Now, we construct the joint CDF of X_1 and X_2. Since

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)\,,$$

we have

$$\begin{aligned} F(x_1, x_2) &= P(max(U_1, U_3) \leq x_1, (max(U_2, U_3) \leq x_2) \\ &= P(U_1 \leq x_1, U_3 \leq x_1, U_2 \leq x_2, U_3 \leq x_2) \\ &= P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq min(x_1, x_2)\,. \end{aligned}$$

Since $U_i, i = 1, 2, 3$ are independent, one gets

$$\begin{aligned} F(x_1, x_2) &= P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq min(x_1, x_2) \\ &= F(x_1, \gamma_1, \alpha, \lambda, \beta, k)\, F(x_2, \gamma_2, \alpha, \lambda, \beta, k)\, F(z, \gamma_3, \alpha, \lambda, \beta, k) \\ &= \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_1}\right] \\ &\quad \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_2 - \beta\, \mathrm{x}_2^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_2}\right] \\ &\quad \times 1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{z} - \beta\, \mathrm{z}^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_3}, \end{aligned} \tag{7.1}$$

where $z = min(x_1, x_2)\,$.

Combining (1.5) and (7.1), we obtain the joint cdf of the bivariate KwEW distribution as:

$$F(x_1, x_2) = \begin{cases} \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_1 + \gamma_3}\right] & \\ \times \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_2 - \beta\, \mathrm{x}_2^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_2}\right], & x_1 \leq x_2 \\ \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_1}\right] & \\ \times \left[1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_2 - \beta\, \mathrm{x}_2^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_2 + \gamma_3}\right], & x_2 \leq x_1 \\ 1 - \left\{1 - \left(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\right)^{\alpha}\right\}^{\gamma_1 + \gamma_2 + \gamma_3}, & x_1 = x_2 = x \end{cases}. \tag{7.2}$$

The joint pdf of (X_1, X_2) is given by

$$f(x_1, x_2) = \begin{cases} f_1(x_1, x_2)\,, & x_1 \leq x_2 \\ f_2(x_1, x_2)\,, & x_2 \leq x_1 \\ f_3(x_1, x_2)\,, & x_1 = x_2 = x \end{cases}.$$

Now, $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ can easily be obtained by taking second order partial differentiation (i.e $f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}$) of the bivariate KwEW cdf given in (7.2) and obtain the following forms:

$$f_1(x_1,x_2)=\alpha^2\gamma_2(\gamma_1+\gamma_3)\left(\beta(-k)\,x_1^{k-1}-\lambda\right)\left(\beta(-k)\,x_2^{k-1}-\lambda\right)$$
$$\times\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha-1}\mathrm{e}^{-\lambda(x_1+x_2)-\beta(x_1^k+x_2^k)}\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha}\right)^{\gamma_2-1}\left(1-\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha}\right)^{\gamma_1+\gamma_3-1} \tag{7.3}$$

and

$$f_2(x_1,x_2)=\alpha^2\gamma_1(\gamma_2+\gamma_3)\left(\beta(-k)\,x_1^{k-1}-\lambda\right)\left(\beta(-k)\,x_2^{k-1}-\lambda\right)$$
$$\times\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha-1}\mathrm{e}^{-\lambda(x_1+x_2)-\beta(x_1^k+x_2^k)}\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha}\right)^{\gamma_1-1}\left(1-\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha}\right)^{\gamma_2+\gamma_3-1}. \tag{7.4}$$

But $f_3(x_1,x_2)$ can not be derived in the similar way. For this, we use the following identity

$$\int_0^\infty\int_0^{x_2} f_1(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2+\int_0^\infty\int_0^{x_1} f_2(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2$$
$$+\int_0^\infty f_3(x,x)\,\mathrm{d}x=1$$
$$=I_1+I_2+\int_0^\infty f_3(x,x)\,\mathrm{d}x=1\,.$$

Let

$$I_1=\alpha\gamma_2\int_0^\infty\left(\beta(-k)\,x_2^{k-1}-\lambda\right)\mathrm{e}^{-\lambda x_2-\beta x_2^k}\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha}\right)^{\gamma_2-1}\mathrm{d}x_2$$
$$\times\,\alpha\,(\gamma_1+\gamma_3)\int_0^{x_2}\left(\beta(-k)\,x_1^{k-1}-\lambda\right)\mathrm{e}^{-\lambda x_1-\beta x_1^k}\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha}\right)^{\gamma_1+\gamma_3-1}\mathrm{d}x_1\,,$$

then

$$I_1=\alpha\gamma_2\int_0^\infty\left(\beta(-k)\,x_2^{k-1}-\lambda\right)\mathrm{e}^{-\lambda x_2-\beta x_2^k}\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha}\right)^{\gamma_1+\gamma_2+\gamma_3-1}\mathrm{d}x_2\,. \tag{7.5}$$

Similarly,

$$I_2=\alpha\gamma_1\int_0^\infty\left(\beta(-k)\,x_1^{k-1}-\lambda\right)\mathrm{e}^{-\lambda x_1-\beta x_1^k}\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_1-\beta x_1^k}\right)^{\alpha}\right)^{\gamma_1-1}\mathrm{d}x_2$$
$$\times\,\alpha\,(\gamma_2+\gamma_3)\int_0^{x_1}\left(\beta(-k)\,x_2^{k-1}-\lambda\right)\mathrm{e}^{-\lambda x_2-\beta x_2^k}\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha-1}$$
$$\times\left(1-\left(1-\mathrm{e}^{-\lambda x_2-\beta x_2^k}\right)^{\alpha}\right)^{\gamma_2+\gamma_3-1}\mathrm{d}x_2\,,$$

then

$$I_2 = \alpha\gamma_1 \int_0^\infty \Big(\beta(-k)\, x^{k-1} - \lambda\Big)\; \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}} \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\Big)^{\alpha-1}$$

$$\times \Big(1 - \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x}_1 - \beta\, \mathrm{x}_1^{\mathrm{k}}}\Big)^{\alpha}\Big)^{\gamma_1+\gamma_2+\gamma_3-1}\; \mathrm{dx}_2\,. \tag{7.6}$$

From (7.5) and (7.6), one obtains

$$\int_0^\infty f_3(x,x)\,\mathrm{dx} = \alpha\gamma_3 \int_0^\infty \Big(\beta(-\mathrm{k})\, \mathrm{x}^{\mathrm{k}-1} - \lambda\Big)\; \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}}$$

$$\times \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}}\Big)^{\alpha-1} \Big(1 - \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}}\Big)^{\alpha}\Big)^{\gamma_1+\gamma_2+\gamma_3-1}\; \mathrm{dx}\,.$$

Thus,

$$f_3(x,x) = \alpha\gamma_3 \int_0^\infty \Big(\beta(-k)\, x^{k-1} - \lambda\Big)\; \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}} \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}}\Big)^{\alpha-1}$$

$$\times \Big(1 - \Big(1 - \mathrm{e}^{-\lambda\, \mathrm{x} - \beta\, \mathrm{x}^{\mathrm{k}}}\Big)^{\alpha}\Big)^{\gamma_1+\gamma_2+\gamma_3-1}\,. \tag{7.7}$$

8. Conclusions

In the last two decades, several authors have been interested in developing methods for generating distributions with more flexibility in applications and data modeling. There has been a growing interest among statisticians and applied researchers in constructing flexible lifetime models in order to improve the modeling of survival data. In particular, some authors proposed new extensions of the classical Weibull model. In this paper, we introduce a five–parameter distribution obtained by applying the Kumaraswamy generator defined by Cordeiro *et al.* [6] to the exponential-Weibull model given by Cordeiro *et al.* [5]. Interestingly, the proposed model has increasing, upside-down bathtub and bathtub shaped hazard rate functions. We study some of its mathematical properties. We discuss the maximum likelihood method and a Bayesian approach to make inference on the model parameters. In the Bayesian approach, the selection of proper priors is difficult to examine and it is left to the interested readers for further study. Also, the monitoring the rate of convergence of the associated MCMC method will be an important issue to look after. An application proves its flexibility to analysis of real data. We also discuss a bivariate extension of the KwEW distribution. The distributional results developed in this paper can have numerous applications in the physical and biological sciences, reliability theory, hydrology, medicine, meteorology, engineering and survival analysis.

Appendix A. The unified Fox–Wright generalized hypergeometric function

Here,

$${}_p\Psi_q^*\Big[\begin{array}{c} (a,A)_p \\ (b,B)_q \end{array}\Big|\, z\Big] = \sum_{n=0}^\infty \frac{\prod_{j=1}^p (a_j)_{A_j n}}{\prod_{j=1}^q (b_j)_{B_j n}}\, \frac{z^n}{n!} \tag{8.1}$$

stands for the *unified variant of the Fox–Wright generalized hypergeometric function* with p upper and q lower parameters; $(a,A)_p$ denotes the parameter p–tuple $(a_1,A_1),\cdots,(a_p,A_p)$ and $a_j \in \mathbb{C}$, $b_i \in \mathbb{C}\setminus\mathbb{Z}_0^-$, $A_i, B_j > 0$ for all $j = \overline{1,p}, i = \overline{1,q}$. The power series converges for suitably bounded values of $|z|$ when

$$\Delta_{p,q} = 1 - \sum_{j=1}^p A_j + \sum_{j=1}^q B_j > 0\,.$$

In the case $\Delta = 0$, the convergence holds in the open disc $|z| < \beta = \prod_{j=1}^q B_j^{B_j} \cdot \prod_{j=1}^p A_j^{-A_j}$.

The function ${}_1\Psi_0^*$ is called *confluent*. The convergence condition $\Delta_{1,0} = 1 - A_1 > 0$ is of special interest for us.

We point out that the original definition of the Fox–Wright function ${}_p\Psi_q[z]$ (consult formula collection [8] and the monographs [11], [15]) contains gamma functions instead of the generalized Pochhammer symbols used here. However, these two functions differ only up to constant multiplying factor, that is

$$ {}_p\Psi_q\Big[\begin{array}{c} (a,A)_p \\ (b,B)_q \end{array} \Big| z \Big] = \frac{\prod_{j=1}^p \Gamma(a_j)}{\prod_{j=1}^q \Gamma(b_j)} \, {}_p\Psi_q^*\Big[\begin{array}{c} (a,A)_p \\ (b,B)_q \end{array} \Big| z \Big]. $$

The unification's motivation is clear - for $A_1 = \cdots = A_p = B_1 = \cdots = B_q = 1$, the fucntion ${}_p\Psi_q^*[z]$ reduces exactly to the well-known generalized hypergeometric function ${}_pF_q[z]$.

Appendix B. Meijer G–function

The symbol $G_{p,q}^{m,n}(\cdot|\cdot)$ denotes Meijer's $G-$function [24] defined in terms of the Mellin–Barnes integral as

$$ G_{p,q}^{m,n}\Big(z\Big| \begin{array}{c} a_1,\cdots,a_p \\ b_1,\cdots,b_q \end{array} \Big) = \frac{1}{2\pi \mathrm{i}} \oint_{\mathfrak{C}} \frac{\prod_{j=1}^m \Gamma(b_j - s) \prod_{j=1}^n \Gamma(1-a_j+s)\, z^s}{\prod_{j=m+1}^q \Gamma(1-b_j+s) \prod_{j=n+1}^p \Gamma(a_j - s)} \mathrm{ds}, \tag{8.2} $$

where $0 \le m \le q$, $0 \le n \le p$ and the poles a_j, b_j are such that no pole of $\Gamma(b_j - s), j = \overline{1,m}$ coincides with any pole of $\Gamma(1-a_j+s), j = \overline{1,n}$; i.e. $a_k - b_j \notin \mathbb{N}$, while $z \neq 0$. $\mathfrak{C}$ is a suitable integration contour which startes at $-\mathrm{i}\infty$ and goes to $\mathrm{i}\infty$ separating the poles of $\Gamma(b_j - s), j = \overline{1,m}$ which lie to the right of the contour, from all poles of $\Gamma(1-a_j+s), j = \overline{1,n}$, which lie to the left of $\mathfrak{C}$. The integral converges if $\delta = m + n - \frac{1}{2}(p+q) > 0$ and $|\arg(z)| < \delta\pi$, see [14, p. 143] and [24].

The G function's Mathematica code reads

$\texttt{MeijerG[}\{\{a_1,...,a_n\},\{a_{n+1},...,a_p\}\},\{\{b_1,...,b_m\},\{b_{m+1},...,b_q\}\},z\texttt{]}.$

Acknowledgement

The research of Gauss M. Cordeiro and Abdus Saboor have been supported by CNPq agency, Brazil and in part by the Higher Education Commission of Pakistan under NRPU project No. 3104, respectively. All the authors are grateful to both referees for their valuable comments, which improve the quality of the paper.

References

[1] Aarset, M. V. *How to identify a bathtub hazard rate*, IEEE Transactions on Reliability **R-36**, 106–108, (1987).

[2] Carrasco, J.M.F., Edwin, M.M. and Cordeiro, G.M. *A generalized modified Weibull distribution for lifetime modeling*, Comput. Statist. Data Anal. **53**, 450–462, (2008).

[3] Cheng, J., Tellambura, C. and Beaulieu, N.C. *Performance analysis of digital modulations on Weibull fading channels*,IEEE: Vehicular Technology Conference, **1**, 236–240, (2003).

[4] Cordeiro, G.M., Edwin, M.M. and Nadarajah, S. *The Kumaraswamy Weibull distribution with application to failure data*, Journal of the Franklin Institute **347**, 1399-1429, (2010).

[5] Cordeiro, G.M., Edwin, M.M. Ortega and Lemonte, A.J. *The exponential-Weibull lifetime distribution*, J Statist Comput Simulation, **84**, 2592–2606, (2014).

[6] Cordeiro, G.M. and Castro, M.D. *A new family of generalized distributions*, J Statist Comput Simulation, **81**, 883–898, (2011).

[7] Cowles, M.K. and Carlin, B.P. *Markov chain Monte Carlo convergence diagnostics: a comparative review*, Journal of the American Statistical Association, **91**, 133–169, (1996).

[8] Erdélyi, A, Magnus, W, Oberhettinger, F, Tricomi, FG. *Higher Transcendental Functions*, Vol. 1, McGraw–Hill, New York, Toronto & London, (1955).

[9] Gelman, A. and Rubin, D.B. *Inference from iterative simulation using multiple sequences (with discussion)*, Statistical Science, **7**, 457–472, (1992).

[10] Gradshteyn, I.S., Ryzhik, I.M. *Table of Integrals, Series, and Products*. Academic Press, New York, (2007).

[11] Kilbas, A. A., Srivastava, H. M. and Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*, North-Holland Mathematical Studies Vol. 204, Amsterdam: Elsevier (North-Holland) Science Publishers, (2006).

[12] Lai, C. D., Xie, M. and Murthy, DNP. *A modified Weibull distribution*, IEEE Transactions on Reliability, **52**, 3–7, (2003).

[13] Lee, C., Famoye, F. and Olumolade, O. *Beta Weibull distribution: Some properties and applications to censored data*, Journal of Modern Applied Statistical Methods, **6**, 173–186, (2007).

[14] Luke, Y.L. *The Special Functions and Their Approximations*. San Diego: Academic Press, (1969).

[15] Mathai, A, Saxena, R. *The H–function with Applications in Statistics and Other Disciplines*. Wiley Halsted, New York, (1978).

[16] Miller, A. R. and Moskowitz, I.S. *Reduction of a class of Fox-Wright Psi functions for certain rational parameters*, Comput. Math. Appl., **30**, 73–82, (1996).

[17] Mudholkar, G.S., Srivastava, D.K., and Friemer, M. *The exponentiated Weibull family: A reanalysis of the bus-motor-failure data*, Technometrics, **37**, 436–445, (1995).

[18] Mudholkar, G.S., Srivastava, D.K., and Kollia, G.D. *A generalization of the Weibull distribution with application to the analysis of survival data*, Journal of the American Statistical Association, **91**, 1575–1583, (1996).

[19] Nadarajah, S. and Kotz, S. *On a distribution of Leipnik and Pearce*, ANZIAM. J., **48**, 405–407, (2007).

[20] Peng, X. and Yan, Z. *Estimation and application for a new extended Weibull distribution*, Reliab Eng Syst Safe, **121**, 34–42, (2014).

[21] Pogány, T. K. and Saxena, R. K. *The gamma–Weibull distribution revisited*, Anais. Acad. Brasil. Ci¸ encias., **82**, 513–520, (2010).

[22] Pogány, T. K., Saboor, A. and Provost, S. *The Marshall-Olkin exponential Weibull distribution*, Hacettepe Journal of Mathematics and Statistics, **44** (6), 1579–1594, (2015).

[23] Silva, G.O., Ortega Edwin, M.M., Cordeiro, G.M. *The beta modified Weibull distribution*, Lifetime Data Anal, **16**, 409–430,(2010).

[24] Meijer, C.S., On the G-function **I**–**VIII**, Proc. Kon. Ned. Akad. Wet., **49**: 227–237, 344–356, 457–469, 632–641, 765–772, 936–943, 1063–1072, 1165–1175 , (1946).

[25] Pham, H. and Lai, CD. *On recent generalizations of the Weibull distribution*, IEEE transactions on reliability, **56** 454–458, (2007).

[26] Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I. *Integrals and Series*, vol. 1–4. Gordon and Breach Science Publishers, Amsterdam, (1986).

[27] Xie, M. and Lai, C.D. *Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function*, *Reliab Eng Syst Safe*, **52**, 87–93, (1995).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1231 – 1244

Two different shrinkage estimator classes for the shape parameter of classical Pareto distribution

Meral Ebegil*† and Senay Ozdemir‡

Abstract

In this study, biased estimators for the shape parameter of a classical Pareto distribution are proposed using two different shrinkage techniques which give a smaller mean square error than an unbiased estimator. Then these obtained biased estimators are compared with the unbiased estimator by the means of their mean square error.

Keywords: Mean Square Error, Pareto Distribution, Shape Parameter, Shrinkage Estimation.

2000 AMS Classification: 60E05,62E99,62F10

Received : 03.09.2014 *Accepted* : 13.06.2015 *Doi* : 10.15672/HJMS.20158812908

1. Introduction

Primarily descriptive parameters of the population are used to make a statistical inference about any population. Unbiased estimators are widely used for this purpose. It can be mentioned that using biased estimators have a smaller mean square error (MSE) if the unbiased estimator has a high MSE.

There have been some studies on biased but smaller MSE estimators of an unknown population parameter. Thompson [1, 2] considered a technique of shrinking best linear unbiased estimator (BLUE) by multiplying it by a shrinking factor to obtain an estimator which has a smaller MSE than that of BLUE. Other important studies about this issue are made by Metha and Srinivasan [3], Govindarajulu and Sahai [4], Das [5], Srivastava et. al. [6], Rao and Singh [7], Bhatnagar [8], Singh and Katyar [9], Singh [10], Jani [11], Kourouklis [12], Singh et. al. [13], Singh and Singh [14], Singh and Shukla [15], Singh and Saxena [16], Prakash et. al. [17], Prakash [18].

*Department of Statistics Gazi University Ankara Turkiye, Email: `mdemirel@gazi.edu.tr`
†Corresponding Author.
‡Department of Statistics Afyon Kocatepe University Afyonkarahisar Turkiye,
Email: `senayozdemir@aku.edu.tr`

Pareto distribution was first used by Pareto [19] to describe an income distribution. Rytgaard [20] studied on the maximum likelihood estimator (MLE) and the moment estimator for the shape parameter of the Pareto distribution. Furthermore, he found a minimum variance unbiased (MVU) estimator for the shape parameter of the Pareto distribution using the obtained MLE. Sing et. al. [13] proposed new shrinkage estimators for scale parameter of the Pareto distribution using the MLE and the unbiased estimator. Then they compared the proposed estimators with the MLE and the unbiased estimator by the means of their MSE. Prakash et. al. [17] obtained that some test estimators for the scale parameter of a classical Pareto distribution are considered when a prior point guess value of the shape parameter is available. Then they showed that their proposed biased test estimators were better than other estimators through a squared error loss function. Prakash [18] derived some shrinkage test estimators and the Bayes estimators for the shape parameter of the Pareto distribution under the general entropy loss function.

In this study, two different estimator classes are obtained for the shape parameter of the Pareto distribution. These estimator classes are compared with the unbiased estimator by the means of their MSE. After that, it is tried to find out in which case obtained estimator classes are better than the unbiased estimator.

2. Shrinkage Estimators Classes

Jani [11] and Singh and Singh [14] proposed two different shrinkage estimator classes for scale parameters of exponential and normal distribution.

First, Jani [11] proposed a shrinkage estimator class for the scale parameter of the exponential distribution is given as

$$T_{(p)} = \theta_0[1 + k(\theta_0/\widehat{\theta})^p] \tag{2.1}$$

where θ_0 is a priori value of θ parameter, k is a shrinking factor minimizing MSE value, p is a nonzero real number and $\widehat{\theta}$ is the unbiased estimator of θ parameter.

Second, Singh and Singh [14] studied on the estimation problem of population variance σ^2 by adapting the estimation class defined equation (2.1) to a normal population. This estimation class is given as the following:

$$\widetilde{\sigma}^2_{(p)} = \sigma_0^2\left[1 + w\left(\frac{s^2}{\sigma_0^2}\right)^p\right] \tag{2.2}$$

where σ_0^2 is a prior value of σ^2 parameter, w is a shrinking factor minimizing MSE value, p is a nonzero real number and s^2 is the unbiased estimator of σ^2 parameter.

The biased estimators, which have a smaller MSE than the unbiased estimator for the shape parameter of Pareto distribution, are obtained using the estimator classes defined in equation (2.1) and equation (2.2).

3. The Obtained Estimators for the Shape Parameter of the Pareto Distribution and Their Properties

In this section, the shape parameter of the Pareto distribution's MVU estimator, which is proposed by Rytgaard [20], is introduced. Then the biased estimator classes, which have smaller MSE than the unbiased estimator, is obtained using various shrinking factors and MSE values of these estimators are calculated.

Let's consider, X is a Pareto distributed random variable. The probability density function (pdf) is as in equation (3.1)

$$f_X(x) = \begin{cases} (\beta\alpha^{\beta})/x^{(\beta+1)} & ; x > \alpha \\ 0 & ; x \leq \alpha \end{cases} \tag{3.1}$$

where α is the scale parameter, β is the shape parameter.

If X random variable has the pdf defined in equation (3.1), the MLE for the shape parameter of the Pareto distribution is

$$\widehat{\beta} = \frac{n}{\sum_{i=1}^{n} \ln \frac{x_i}{\alpha}}. \tag{3.2}$$

Using equation (3.2) estimator, Rytgaard [20] obtained an unbiased estimator which is defined as

$$\tilde{\beta} = \frac{n-1}{n}\widehat{\beta}. \tag{3.3}$$

It can be found that the expected value of this estimator is

$E[\tilde{\beta}] = \beta$.

Variance of $\tilde{\beta}$ estimator is

$Var(\tilde{\beta}) = \frac{1}{(n-2)}\beta^2$ where $E[\tilde{\beta}^2] = \frac{(n-1)}{(n-2)}\beta^2$.

3.1. Corollary. *The shrinkage estimator class for the shape parameter of Pareto distribution, which is obtained by help of equation(2.1), given as*

$$\beta^{\star}_{(p)} = \beta_0 + (\tilde{\beta} - \beta_0)k_{(p)} \tag{3.4}$$

where

$$k_{(p)} = (n-1)^p \frac{(n+p-1)!}{(n+2p-1)!} \tag{3.5}$$

and p is a nonzero real number. MSE of $\beta^{\star}_{(p)}$ estimator class is

$$MSE(\beta^{\star}_{(p)}) = \beta^2 \left[\frac{k^2_{(p)}}{(n-2)} + (k_{(p)} - 1)^2 (1-\lambda)^2 \right] \tag{3.6}$$

where $\lambda = \beta_0/\beta$. Furthermore bias of $\beta^{\star}_{(p)}$ estimators class given by

$$Bias(\beta^{\star}_{(p)}) = (1 - k_{(p)})(\beta_0 - \beta). \tag{3.7}$$

Proof. The shrinkage estimator class for the shape parameter of Pareto distribution is described as

$$\beta^{\star}_{(p)} = \beta_0 \left[1 + k \left(\frac{\beta_0}{\widetilde{\beta}} \right)^p \right] \tag{3.8}$$

which is obtained by means of equation (2.1).

$E\left[\left(\frac{1}{\beta}\right)^{jp}\right] = K_{1(jp)}(1/\beta)^{jp}\ , (j=1,2)$
and

$$K_{1(jp)} = (n-1)^{-jp}\frac{(n+jp-1)!}{(n-1)!}$$

functions are used to calculate the MSE of $\beta^{\star}_{(p)}$ estimator class. As it is known the value of $MSE\left(\beta^{\star}_{(p)}\right)$ is

$$(3.9)\qquad MSE\left(\beta^{\star}_{(p)}\right) = E\left[\beta^{\star}_{(p)} - \beta\right]^2.$$

If required information is written in equation (3.9) where $\lambda = \beta_0/\beta$, the MSE value is obtained as

$$MSE(\beta^{\star}_{(p)}) = \beta^2\left[k^2(\lambda)^{2(1+p)}K_{1(2p)} + 2k(\lambda)^{(1+p)}(\lambda-1)\,K_{1(p)} + (\lambda-1)^2\right].$$

Differentiating this equation with respect to k and setting the derivative equal to zero, we find

$$(3.10)\qquad k = (\lambda)^{-p}\left(\frac{1}{\lambda}-1\right)K_{1(p)}/K_{1(2p)}$$

which minimizes the MSE value. If the required values are inserted into equation (3.10), the k value is obtained as given in equation (3.11);

$$(3.11)\qquad k = \left(\frac{\beta-\beta_0}{\beta_0}\right)\left(\frac{\beta}{\beta_0}\right)^p(n-1)^p\frac{(n+p-1)!}{(n+2p-1)!}$$

The shrinking parameter k is obtained as a function of β parameter. In practice, it is impossible to attain parameter β. Therefore the unknown parameters in equation (3.11) are replaced by their unbiased estimators. So an estimator for k is obtained as

$$\hat{k} = \left(\frac{\widetilde{\beta}-\beta_0}{\beta_0}\right)\left(\frac{\beta}{\beta_0}\right)^p(n-1)^p\frac{(n+p-1)!}{(n+2p-1)!}.$$

On conclusion, when necessary adjustment is made, the estimator class for the shape parameter of Pareto distribution is obtained as

$$\beta^{\star}_{(p)} = \beta_0 + \left(\widetilde{\beta}-\beta_0\right)k_{(p)}$$

where $k_{(p)} = (n-1)^p\dfrac{(n+p-1)!}{(n+2p-1)!}$.Thus, the MSE value of $\beta^{\star}_{(p)}$ is obtained as

$$(3.12)\qquad MSE\left(\beta^{\star}_{(p)}\right) = \frac{k^2_{(p)}\beta^2}{(n-2)} + \left(k_{(p)}-1\right)^2(\beta-\beta_0)^2$$

by making necessary adjustment in equation (3.9). If $\lambda = \beta_0/\beta$ is written on its place, equation (3.12) is written as

$$(3.13)\qquad MSE\left(\beta^{\star}_{(p)}\right) = \beta^2\left[\frac{k^2_{(p)}}{(n-2)} + \left(k_{(p)}-1\right)^2(1-\lambda)^2\right].$$

The bias of $\beta^{\star}_{(p)}$ estimators class is obtained as

$$Bias\left(\beta^{\star}_{(p)}\right)=E\left(\beta^{\star}_{(p)}\right)-\beta=\left(1-k_{(p)}\right)\left(\beta_0-\beta\right)\ .$$

Thus the proof is completed. □

The relative efficiency of $\beta^{\star}_{(p)}$ estimator class with respect to $\widetilde{\beta}$ estimator is calculated by means of

$$\frac{MSE\left(\beta^{\star}_{(p)}\right)}{Var\left(\widetilde{\beta}\right)}=k^2_{(p)}+(n-2)\left(k_{(p)}-1\right)^2(1-\lambda)^2\ . \tag{3.14}$$

If equation (3.14) is smaller than 1, it is clear that $MSE(\beta^{\star}_{(p)})<Var(\widetilde{\beta})$.

Case Study 1: Consider that p=1. By using equation(3.4) and equation(3.5) an estimator is obtain as

$$\beta^{\star}_{(1)}=\beta_0+\left(\widetilde{\beta}-\beta_0\right)\frac{(n-1)}{(n+1)}\ .$$

The MSE value of this estimator is

$$MSE\left(\beta^{\star}_{(1)}\right)=\beta^2\left[\frac{(n-1)^2}{(n-2)\,(n+1)^2}+\left(\frac{(n-1)}{(n+1)}-1\right)^2(1-\lambda)^2\right].$$

The relative efficiency of $\beta^{\star}_{(1)}$ estimator with respect to $\widetilde{\beta}$ estimator is

$$\begin{aligned}\frac{MSE(\beta^{\star}_{(1)})}{Var(\widetilde{\beta})} &= \frac{(n-1)^2}{(n+1)^2}+(n-2)\left(\frac{(n-1)}{(n+1)}-1\right)^2(1-\lambda)^2\\ &= \frac{(n-1)^2}{(n+1)^2}+\frac{4\,(n-2)}{(n+1)^2}(1-\lambda)^2.\end{aligned}$$

It is clear that $\beta^{\star}_{(1)}$ estimator is better than $\widetilde{\beta}$ estimator if $\frac{MSE\left(\beta^{\star}_{(1)}\right)}{Var\left(\widetilde{\beta}\right)}<1$ inequality is true. Thus

$$\frac{(n-1)^2}{(n+1)^2}+\frac{4\,(n-2)}{(n+1)^2}(1-\lambda)^2<1 \tag{3.15}$$

inequality can be written. If the necessary adjustment is made in equation (3.15), it is obtained that

$$0<\lambda<1+\left(\frac{n}{n-2}\right)^{1/2}.$$

In case this inequality is true, it can be said that $\beta^{\star}_{(1)}$ estimator is better than $\widetilde{\beta}$ estimator. Further, when n is very large (i.e. $n\to\infty$)

$$(1-\lambda)^2<\frac{n}{n-2}$$

inequality reduces to $0<\lambda<2$.

Case Study 2: Consider that $p=2$. By using equation(3.4) and equation(3.5) an estimator is obtained as

$$\beta^{\star}_{(2)} = \beta_0 + \left(\widetilde{\beta} - \beta_0\right) \frac{(n-1)^2}{(n+1)(n+2)}.$$

The MSE value of this estimator is

$$MSE\left(\beta^{\star}_{(2)}\right) = \beta^2 \left[\tfrac{(n-1)^4}{(n+1)^2(n+2)^2(n-2)} + \left(\tfrac{(n-1)^2}{(n+1)(n+2)} - 1\right)^2 (1-\lambda)^2\right].$$

The relative efficiency of $\beta^{\star}_{(2)}$ estimator with respect to $\widetilde{\beta}$ estimator is

$$\begin{aligned}\frac{MSE(\beta^{\star}_{(2)})}{Var(\widetilde{\beta})} &= \left[\frac{(n-1)^2}{(n+1)(n+2)}\right]^2 + (n-2)\left[\frac{(n-1)^2}{(n+1)(n+2)} - 1\right]^2 (1-\lambda)^2 \\ &= \frac{(n-1)^4}{(n+1)^2(n+2)^2} + \frac{(n-2)(5n+1)^2}{(n+1)^2(n+2)^2}(1-\lambda)^2.\end{aligned}$$

It is clear that $\beta^{\star}_{(2)}$ estimator is better than $\widetilde{\beta}$ estimator if $\frac{MSE\left(\beta^{\star}_{(2)}\right)}{Var\left(\widetilde{\beta}\right)} < 1$ inequality is true. Thus

$$\frac{(n-1)^4}{(n+1)^2(n+2)^2} + \frac{(n-2)(5n+1)^2}{(n+1)^2(n+2)^2}(1-\lambda)^2 < 1 \tag{3.16}$$

inequality can be written. If the necessary adjustment is made in equation (3.16), it is obtained that

$$1 - \left(\frac{2n^2+n+3}{(n-2\ )(5n+1)}\right)^{1/2} < \lambda < 1 + \left(\frac{2n^2+n+3}{(n-2\ )(5n+1)}\right)^{1/2}.$$

In case this inequality is true, it can be said that $\beta^{\star}_{(2)}$ estimator is better than $\widetilde{\beta}$ estimator. Furthermore, when n is very large (i.e. $n \to \infty$)

$$(1-\lambda)^2 < \frac{2n^2+n+3}{(n-2\ )(5n+1)}$$

inequality reduces to $0.37 < \lambda < 1.63$.

3.2. Corollary. *The shrinkage estimator class for the shape parameter of Pareto distribution, which is obtained by means of equation(2.2), is given as*

$$\beta^{*}_{(p)} \ = \beta_0 + \left(\widetilde{\beta} - \beta_0\right) w_{(p)} \tag{3.17}$$

where

$$w_{(p)} = (n-1)^{-p} \frac{(n-p-1)!}{(n-2p-1)!}. \tag{3.18}$$

of $\beta^{*}_{(p)}$ estimator class is

$$MSE\left(\beta^{*}_{(p)}\right) = \beta^2 \left[\tfrac{w^2_{(p)}}{(n-2)} + \left(w_{(p)} - 1\right)^2 (1-\lambda)^2\right]$$

where $\lambda = \beta_0/\beta$.Furthermore bias of $\beta^{*}_{(p)}$estimator class is given by

$$Bias\left(\beta^{*}_{(p)}\right) = \left(1 - w_{(p)}\right)(\beta_0 - \beta)\ \ . \tag{3.19}$$

Proof. The shrinkage estimator class for the shape parameter of Pareto distribution is described as

$$\beta^*_{(p)} = \beta_0 \left[1 + w \left(\frac{\widetilde{\beta}}{\beta_0} \right)^p \right]$$

which is obtained by means of equation(2.2).

$$E\left(\widetilde{\beta}^{jp}\right) = K_{2(jp)}(\beta)^{jp} \ , (j = 1, 2)$$

and

$$K_{2(jp)} = (n-1)^{jp} \frac{(n-jp-1)!}{(n-1)!}$$

functions are used to calculate the MSE of $\beta^*_{(p)}$ estimator class. The value of $MSE\left(\beta^*_{(p)}\right)$ is

$$MSE(\beta^*_{(p)} = E\big[\beta^*_{(p)} - \beta\big]^2 \tag{3.20}$$

If necessary information is written in equation (3.20) where $\lambda = \frac{\beta_0}{\beta}$, the MSE value is obtained as

$$MSE\left(\beta^*_{(p)}\right) = \beta^2 \left[w^2 (\lambda)^{2(1-p)} K_{2(2p)} + 2w(\lambda)^{1-p} (\lambda - 1) K_{2(p)} + (\lambda - 1)^2 \right].$$

Differentiating this equation with respect to w and setting the derivate equal to zero, we find

$$w = \left(\frac{1}{\lambda} - 1 \right) (\lambda)^p \left(\frac{K_{2(p)}}{K_{2(2p)}} \right) \tag{3.21}$$

which is a constant minimizing the MSE value. If necessary information is written in equality which is introduced equation (3.21) w is obtained as follows:

$$w = \left(\frac{\beta - \beta_0}{\beta_0} \right) \left(\frac{\beta_0}{\beta} \right)^p (n-1)^{-p} \frac{(n-p-1)!}{(n-2p-1)!}. \tag{3.22}$$

The shrinking parameter k is obtained as a function of βparameter. In practice, it is impossible to attain parameter β. Therefore the unknown parameter in equation (3.22) is replaced by its unbiased estimator. So an estimator for w is obtained as

$$\hat{w} = \left(\frac{\widetilde{\beta} - \beta_0}{\beta_0} \right) \left(\frac{\beta_0}{\widetilde{\beta}} \right)^p (n-1)^{-p} \frac{(n-p-1)!}{(n-2p-1)!}.$$

On conclusion, when necessary adjustment is made, the estimator class for the shape parameter of Pareto distribution is obtained as

$$\beta^*_{(p)} \ = \beta_0 + \left(\widetilde{\beta} - \beta_0 \right) w_{(p)}$$

where $w_{(p)} = (n-1)^{-p} \frac{(n-p-1)!}{(n-2p-1)!}$ .Thus, the MSE value of $\beta^*_{(p)}$ is obtained as

$$MSE\left(\beta^*_{(p)}\right) = \frac{w^2_{(p)} \beta^2}{(n-2)} + \left(w_{(p)} - 1\right)^2 (\beta - \beta_0)^2 \tag{3.23}$$

by making necessary adjustment in equation (3.20). If $\lambda = \beta_0/\beta$ is written on its place, equation (3.23) is written as

$$MSE\left(\beta^*_{(p)}\right) = \beta^2 \left[\frac{w^2_{(p)}}{(n-2)} + \left(w_{(p)} - 1\right)^2 (1-\lambda)^2 \right].$$

Furthermore the bias of $\beta^*_{(p)}$estimator class can be obtained as

$$Bias\left(\beta^*_{(p)}\right) = E\left(\beta^*_{(p)}\right) - \beta = \left(1 - w_{(p)}\right)\left(\beta_0 - \beta\right) \ .$$

Thus the proof is completed. □

The relative efficiency of $\beta^*_{(p)}$ estimator class with respect to $\widetilde{\beta}$ estimator is calculated by means of

$$\frac{MSE\left(\beta^*_{(p)}\right)}{Var\left(\widetilde{\beta}\right)} = w^2_{(p)} + (n-2)\left(w_{(p)} - 1\right)^2 (1-\lambda)^2. \tag{3.24}$$

If equation (3.24) is smaller than 1, it is clear that $MSE\left(\beta^*_{(p)}\right) < Var\left(\widetilde{\beta}\right)$.

Case Study 3: Consider that p=1. An estimator is obtained as

$$\beta^*_{(1)} = \beta_0 + \left(\widetilde{\beta} - \beta_0\right) \frac{(n-2)}{(n-1)}$$

by using equation (3.17) and equation (3.18). The MSE value of this estimator is

$$MSE\left(\beta^*_{(1)}\right) = \beta^2 \left[\frac{(n-2)^2}{(n-2)(n-1)^2} + \left(\frac{(n-2)}{(n-1)} - 1 \right)^2 (1-\lambda)^2 \right].$$

The relative efficiency of $\beta^*_{(1)}$ estimator with respect to $\widetilde{\beta}$ estimator is

$$\begin{aligned} \frac{MSE(\beta^*_{(1)})}{Var(\widetilde{\beta})} &= \frac{(n-2)^2}{(n-1)^2} + (n-2)\left[\frac{(n-2)}{(n-1)} - 1\right]^2 (1-\lambda)^2 \\ &= \frac{(n-2)^2}{(n-1)^2} + \frac{(n-2)}{(n-1)^2}(1-\lambda)^2. \end{aligned}$$

It is clear that $\beta^*_{(1)}$ estimator is better than $\widetilde{\beta}$ estimator if$\frac{MSE\left(\beta^*_{(1)}\right)}{Var\left(\widetilde{\beta}\right)} < 1$ inequality is true. Thus

$$\frac{(n-2)^2}{(n-1)^2} + \frac{(n-2)}{(n-1)^2}(1-\lambda)^2 < 1$$

inequality can be written. If the necessary adjustment is made in above inequality, it is obtained that

$$0 < \lambda < 1 + \left(\frac{2n-3}{n-2}\right)^{1/2}.$$

In case this inequality is true, it can be said that $\beta^*_{(1)}$estimator is better than $\widetilde{\beta}$ estimator. Further, when n is very large (i.e. $n \to \infty$)

$$(1-\lambda)^2 < \frac{2n-3}{n-2}$$

inequality reduces to $0 < \lambda < 2.41$.

Case Study 4: Consider that $p = 2$. By using equation (3.17) and equation (3.18) an estimator is obtained as

$$\beta^*_{(2)} = \beta_0 + \left(\widetilde{\beta} - \beta_0\right) \frac{(n-3)(n-4)}{(n-1)^2}.$$

The MSE value of this estimator is

$$MSE\left(\beta^*_{(2)}\right) = \beta^2 \left[\tfrac{(n-3)^2(n-4)^2}{(n-2)(n-1)^4} + \left(\tfrac{(n-3)(n-4)}{(n-1)^2} - 1\right)^2 (1-\lambda)^2\right].$$

The relative efficiency of $\beta^*_{(2)}$ estimator with respect to $\widetilde{\beta}$ estimator is

$$\begin{aligned} \frac{MSE(\beta^*_{(2)})}{Var(\widetilde{\beta})} &= \frac{(n-3)^2(n-4)^2}{(n-1)^4} + (n-2)\left(\frac{(n-3)(n-4)}{(n-1)^2} - 1\right)^2 (1-\lambda)^2 \\ &= \frac{(n-3)^2(n-4)^2}{(n-1)^4} + \frac{(n-2)(5n-11)^2}{(n-1)^2}. \end{aligned}$$

It is clear that $\beta^*_{(2)}$ estimator is better than $\widetilde{\beta}$ estimator if $\frac{MSE\left(\beta^*_{(2)}\right)}{Var\left(\widetilde{\beta}\right)} < 1$ inequality is true. Thus

$$\frac{(n-3)^2(n-4)^2}{(n-1)^4} + \frac{(n-2)(5n-11)^2}{(n-1)^2}(1-\lambda)^2 < 1$$

inequality can be written. If the necessary adjustment is made in above inequality, it is obtained that

$$1 - \left(\frac{2n^2 - 9n + 13}{(n-2)(5n-11)}\right)^{\frac{1}{2}} < \lambda < 1 + \left(\frac{2n^2 - 9n + 13}{(n-2)(5n-11)}\right)^{\frac{1}{2}}.$$

In case this inequality is true, it can be said that $\beta^*_{(2)}$estimator is better than $\widetilde{\beta}$ estimator. Further, when n is very large (i.e. $n \to \infty$)

$$(1-\lambda)^2 < \frac{2n^2 - 9n + 13}{(n-2)(5n-11)}$$

inequality reduces to $0.37 < \lambda < 1.63$.

Note: It can be seen that the estimator class proposed by Jani [11] is directly related with that of Singh and Singh [14] for the shape parameter of the Pareto distribution. This relationship is expressed as $k_{(p)} = w_{(-p)}$.

4. Comparisons of the estimators

Here, the relative efficiency of the obtained estimator classes with respect to the unbiased estimator for the shape parameter of the Pareto distribution is calculated using different values of n, p and λ. The handled λ values are selected by considering the efficiency range for large n values in case studies.

The relative efficiency of the estimator class introduced in equation (3.4) with respect to the estimator given in equation (3.3) is calculated for the different value of n, p and λ by the help of equation (3.14). These calculated values are summarized in Table 1.

Table 1. The relative efficiency of the estimator class proposed equation (3.4) with respect to estimator given by equation (3.3)

		Sample Size n				
λ	**Estimator**	**5**	**10**	**15**	**25**	**50**
0.125	$\beta^{\star}_{(-1)}$	0.8657	0.9130	0.9357	0.9490	0.9749
	$\beta^{\star}_{(-1/2)}$	0.9738	0.9831	0.9875	0.9901	0.9951
	$\beta^{\star}_{(1/2)}$	0.9022	0.9354	0.9518	0.9615	0.9809
	$\beta^{\star}_{(1)}$	0.8719	0.9211	0.9436	0.9563	0.9796
	$\beta^{\star}_{(3/2)}$	1.1139	1.1369	1.1304	1.1190	1.0761
	$\beta^{\star}_{(2)}$	1.6853	1.6965	1.6407	1.5788	1.3687
	$\beta^{\star}_{(5/2)}$	2.4890	2.5955	2.5228	2.4111	1.9533
0.50	$\beta^{\star}_{(-1)}$	0.8148	0.8788	0.9100	0.9284	0.9646
	$\beta^{\star}_{(-1/2)}$	0.9730	0.9826	0.9871	0.9898	0.9950
	$\beta^{\star}_{(1/2)}$	0.8858	0.9236	0.9426	0.9541	0.9770
	$\beta^{\star}_{(1)}$	0.7355	0.8164	0.8594	0.8861	0.9416
	$\beta^{\star}_{(3/2)}$	0.6667	0.7647	0.8182	0.8519	0.9231
	$\beta^{\star}_{(2)}$	0.7319	0.8302	0.8785	0.9065	0.9581
	$\beta^{\star}_{(5/2)}$	0.9107	1.0270	1.0697	1.0858	1.0823
1.00	$\beta^{\star}_{(-1)}$	0.7901	0.8622	0.8975	0.9184	0.9596
	$\beta^{\star}_{(-1/2)}$	0.9726	0.9823	0.9869	0.9896	0.9949
	$\beta^{\star}_{(1/2)}$	0.8778	0.9179	0.9382	0.9504	0.9751
	$\beta^{\star}_{(1)}$	0.6694	0.7656	0.8186	0.8521	0.9231
	$\beta^{\star}_{(3/2)}$	0.4499	0.5842	0.6668	0.7224	0.8490
	$\beta^{\star}_{(2)}$	0.2696	0.4103	0.5090	0.5805	0.7590
	$\beta^{\star}_{(5/2)}$	0.1455	0.2665	0.3652	0.4432	0.6600
1.50	$\beta^{\star}_{(-1)}$	0.8148	0.8788	0.9100	0.9284	0.9646
	$\beta^{\star}_{(-1/2)}$	0.9730	0.9826	0.9871	0.9898	0.9950
	$\beta^{\star}_{(1/2)}$	0.8858	0.9236	0.9426	0.9541	0.9770
	$\beta^{\star}_{(1)}$	0.7355	0.8164	0.8594	0.8861	0.9416
	$\beta^{\star}_{(3/2)}$	0.6667	0.7647	0.8182	0.8519	0.9231
	$\beta^{\star}_{(2)}$	0.7319	0.8302	0.8785	0.9065	0.9581
	$\beta^{\star}_{(5/2)}$	0.9107	1.0270	1.0697	1.0858	1.0823
2.50	$\beta^{\star}_{(-1)}$	1.0123	1.0115	1.0097	1.0082	1.0046
	$\beta^{\star}_{(-1/2)}$	0.9760	0.9846	0.9887	0.9910	0.9956
	$\beta^{\star}_{(1/2)}$	0.9494	0.9693	0.9781	0.9830	0.9920
	$\beta^{\star}_{(1)}$	1.2645	1.2227	1.1859	1.1583	1.0892
	$\beta^{\star}_{(3/2)}$	2.4013	2.2085	2.0290	1.8880	1.5164
	$\beta^{\star}_{(2)}$	4.4301	4.1901	3.8347	3.5142	2.5509
	$\beta^{\star}_{(5/2)}$	7.0327	7.1109	6.7060	6.2265	4.4606

Table 1 shows that $\beta^{\star}_{(-1/2)}$ and $\beta^{\star}_{(1/2)}$ estimators are better than the unbiased estimators for all values of λ. Further when $0.50 \leq \lambda \leq 1.50$, the all proposed biased estimators are better than the unbiased estimators. Hence the efficiency of the proposed biased estimator class with respect to the unbiased estimator decreases as λ values differ from 1. Besides increased p values cause a decrease in efficiency of the proposed biased estimator class with respect to the unbiased estimator.

Table 2. The relative efficiency of the estimator class proposed equation (3.17) with respect to estimator given by equation (3.3)

		Sample Size n				
λ	**Estimator**	**5**	**10**	**15**	**25**	**50**
0.125	$\beta^*_{(-1)}$	0.8719	0.9211	0.9436	0.9563	0.9796
	$\beta^*_{(-1/2)}$	0.9022	0.9354	0.9518	0.9615	0.9809
	$\beta^*_{(1/2)}$	0.9738	0.9831	0.9875	0.9901	0.9951
	$\beta^*_{(1)}$	0.8657	0.9130	0.9357	0.9490	0.9749
	$\beta^*_{(3/2)}$	0.9948	0.9987	0.9997	1.0000	1.0003
	$\beta^*_{(2)}$	1.6888	1.5148	1.4053	1.3331	1.1750
	$\beta^*_{(5/2)}$	2.9280	2.6133	2.3354	2.1289	1.6252
0.50	$\beta^*_{(-1)}$	0.7355	0.8164	0.8594	0.8861	0.9416
	$\beta^*_{(-1/2)}$	0.8858	0.9236	0.9426	0.9541	0.9770
	$\beta^*_{(1/2)}$	0.9730	0.9826	0.9871	0.9898	0.9950
	$\beta^*_{(1)}$	0.8148	0.8788	0.9100	0.9284	0.9646
	$\beta^*_{(3/2)}$	0.6758	0.7768	0.8302	0.8630	0.9304
	$\beta^*_{(2)}$	0.7325	0.8001	0.8412	0.8686	0.9297
	$\beta^*_{(5/2)}$	1.0254	1.0314	1.0250	1.0197	1.0083
1.00	$\beta^*_{(-1)}$	0.6694	0.7656	0.8186	0.8521	0.9231
	$\beta^*_{(-1/2)}$	0.8778	0.9179	0.9382	0.9504	0.9751
	$\beta^*_{(1/2)}$	0.9726	0.9823	0.9869	0.9896	0.9949
	$\beta^*_{(1)}$	0.7901	0.8622	0.8975	0.9184	0.9596
	$\beta^*_{(3/2)}$	0.5212	0.6692	0.7480	0.7966	0.8966
	$\beta^*_{(2)}$	0.2689	0.4536	0.5677	0.6433	0.8108
	$\beta^*_{(5/2)}$	0.1030	0.2644	0.3897	0.4819	0.7092
1.50	$\beta^*_{(-1)}$	0.7355	0.8164	0.8594	0.8861	0.9416
	$\beta^*_{(-1/2)}$	0.8858	0.9236	0.9426	0.9541	0.9770
	$\beta^*_{(1/2)}$	0.9730	0.9826	0.9871	0.9898	0.9950
	$\beta^*_{(1)}$	0.8148	0.8788	0.9100	0.9284	0.9646
	$\beta^*_{(3/2)}$	0.6758	0.7768	0.8302	0.8630	0.9304
	$\beta^*_{(2)}$	0.7325	0.8001	0.8412	0.8686	0.9297
	$\beta^*_{(5/2)}$	1.0254	1.0314	1.0250	1.0197	1.0083
2.50	$\beta^*_{(-1)}$	1.2645	1.2227	1.1859	1.1583	1.0892
	$\beta^*_{(-1/2)}$	0.9494	0.9693	0.9781	0.9830	0.9920
	$\beta^*_{(1/2)}$	0.9760	0.9846	0.9887	0.9910	0.9956
	$\beta^*_{(1)}$	1.0123	1.0115	1.0097	1.0082	1.0046
	$\beta^*_{(3/2)}$	1.9130	1.6374	1.4877	1.3946	1.2014
	$\beta^*_{(2)}$	4.4417	3.5723	3.0293	2.6704	1.8810
	$\beta^*_{(5/2)}$	8.4051	7.1672	6.1078	5.3222	3.4012

Similarly, the relative efficiency of the estimator class proposed in equation (3.17) with respect to estimator given in equation (3.3) is calculated for different values of n, p and λ with the help of equation (3.24). These calculated values are given in Table 2.

Table 2 shows that $\beta^*_{(-1/2)}$ and $\beta^*_{(1/2)}$ estimators are better than the unbiased estimators for all λ values. Furthermore, when $0.50 \leq \lambda \leq 1.50$, the all proposed biased estimators better than the unbiased estimators. But the efficiency of the proposed biased estimator class with respect to the unbiased estimator decrease as λ values differ

Table 3. The relative biases of equation (3.4) and equation (3.17) estimators for different n and p values

p	Sample Size n				
	10	**15**	**20**	**25**	**50**
-1	0.6111	0.5714	0.5526	0.5417	0.5204
-1/2	0.2185	0.2120	0.2089	0.2070	0.2035
1/2	4.5769	4.7175	4.7880	4.8303	4.9151
1	1.6364	1.7500	1.8095	1.8462	1.9216
3/2	1.1841	1.2952	1.3569	1.3961	1.4799
2	0.9985	1.1009	1.1623	1.2030	1.2940
5/2	0.9108	0.9958	1.0530	1.0931	1.1882

from 1. In addition increased p values cause a decreased efficiency of the proposed biased estimator class with respect to the unbiased estimator. Moreover, when the estimators given in Table 1 are compared to the estimators given in Table 2, it is observed that the efficiency range of the estimator class introduced in equation (3.17) with respect to the estimator given in equation (3.3) is larger than that of the estimator class introduced in equation (3.4) with respect to estimator given in equation (3.3).

In addition to the MSE criteria, bias has an important role in comparison of estimators. A relative bias can be calculated by dividing equation (3.7) to equation (3.19). The relative bias is given in equation (4.1).

$$\frac{Bias(\beta^{\star}_{(p)})}{Bias(\beta^{*}_{(p)})} = \frac{1-k_{(p)}}{1-w_{(p)}} \,. \tag{4.1}$$

The relative bias values are calculated by means of equation (4.1) for different n and p values and given in Table 3.

In Table 3, it is seen that the biases of $\beta^{\star}_{(p)}$ estimators are smaller than those of $\beta^{*}_{(p)}$ estimators when p has a negative value. Furthermore, it can be mentioned that $Bias(\beta^{\star}_{(p)})/Bias(\beta^{*}_{(p)})$ values decrease when there is an increase on positive values of p. However, it can be noted that$\beta^{*}_{(p)}$ estimators have smaller bias than $\beta^{\star}_{(p)}$estimators if p is near 1.

5. Simulation Study

In this section, we generated a data set for the Pareto distribution with the shape parameter $\beta = 5$ and the scale parameter $\alpha = 1$. The scale parameter α was taken as 1 because the same results were obtained from experiments for $\alpha = 1,\ 1.5,\ 2, \ldots$. The shape parameter should be greater than 2 so that the variance of a data set from the Pareto distribution could be calculated. Also Thompson [1,2] used the proportion 1/5 between two descriptive parameters of the normal distribution in his study. The relative efficiency of the obtained estimator classes with respect to the unbiased estimator for the shape parameter of the Pareto distribution is calculated using different values of n, p and λ. First we calculated the MSE values to obtain the relative efficiency. These MSE values were calculated by the means of Monte Carlo Simulation study where the number of iterations was 25000. We obtained relative efficiencies similar to that of previous section. The simulation study results which are given in Table 4 support to the theoretical results.

Table 4. The relative efficiency of the estimator class proposed equation(3.4) and equation(3.17) with respect to estimator given by equation (3.3)

		Sample Size n		
λ	**Estimator**	**5**	**15**	**50**
0.50	$\beta^{\star}_{(-1)}$	0.7797	0.9339	0.9808
	$\beta^{*}_{(-1)}$	0.7124	0.8857	0.9633
	$\beta^{\star}_{(-1/2)}$	0.9715	0.9916	0.9976
	$\beta^{*}_{(-1/2)}$	0.8841	0.9609	0.9882
	$\beta^{\star}_{(1)}$	0.7123	0.8857	0.9633
	$\beta^{*}_{(1)}$	0.7796	0.9339	0.9808
1.00	$\beta^{\star}_{(-1)}$	0.5625	0.7901	0.9596
	$\beta^{*}_{(-1)}$	0.4444	0.6694	0.9231
	$\beta^{\star}_{(-1/2)}$	0.9396	0.9726	0.9949
	$\beta^{*}_{(-1/2)}$	0.7610	0.8778	0.9751
	$\beta^{\star}_{(1)}$	0.4444	0.6694	0.9231
	$\beta^{*}_{(1)}$	0.5625	0.7901	0.9596
2.50	$\beta^{\star}_{(-1)}$	0.1242	0.6646	0.8973
	$\beta^{*}_{(-1)}$	0.0195	0.4575	0.8075
	$\beta^{\star}_{(-1/2)}$	0.8471	0.9545	0.9869
	$\beta^{*}_{(-1/2)}$	0.4475	0.7948	0.9363
	$\beta^{\star}_{(1)}$	0.0195	0.4575	0.8075
	$\beta^{*}_{(1)}$	0.1242	0.6646	0.8973

6. Conclusion and Suggestions

When the biased estimators give smaller MSE than unbiased estimators, the biased estimators can be preferred to the unbiased estimators. In this study, considering this case, two different biased estimator classes are proposed. These estimators are generated by minimizing MSE.

In section 4, the cases in which the biased estimators have smaller MSE than the unbiased estimator are assessed. When $0.50 \leq \lambda \leq 1.50$, the biased estimator class which is given in equation(3.4) is better than the unbiased estimator. However the efficiency of the proposed biased estimator class with respect to the unbiased estimator decreases as the λ values differ from 1. Increased p values cause a decrease in efficiency of the proposed biased estimator class with respect to the unbiased estimator. Similarly, when $0.50 \leq \lambda \leq 1.50$, the biased estimator class given in equation (3.17) is better than the unbiased estimators. However the efficiency of the proposed biased estimator class with respect to the unbiased estimator decreases as λ values differ from 1. Further increased p values cause decrease in efficiency of the proposed biased estimator class with respect to the unbiased estimator. When the relative efficiency values given in Table 1 and Table 2 are considered, it is shown that the both biased estimators classes have almost the same efficiency range. Besides, if both biased estimator classes are considered as an efficient range, it is observed that the efficiencies of biased estimators with respect to biased estimators decrease when n increases. In addition to the relative efficiency values in both tables, the efficiency range of the estimator class introduced in equation (3.17) is greater than that of the estimator class given in equation (3.4) as shown in case study 1 and 3. It is observed that the estimator class given in equation (3.4) is more efficient than the others when p is a negative real number, while the estimator class given equation (3.17) is more efficient than others when p is a positive real number.

In conclusion, it is possible to obtain estimators that give a smaller MSE than the unbiased estimator for the shape parameter of Pareto distribution using the estimator class given in equation (3.4) if p is a negative number near zero, while it is reasonable to use the estimator class given in equation (3.17) if p is a positive number near zero, when λ values are near 1.

Acknowledgement
Authors are grateful to the referees for the valuable comments and suggestions they made in order to improve this paper.

References

[1] Thompson, J.R. *Some shrinkage techniques for estimating the mean*, Jour. Amer. Statist. Assoc. **63**, 113-122, 1968.
[2] Thompson, J.R. *Accuracy borrowing in the estimation of the mean by shrinkage to an interval*, Jour. Amer. Statist. Assoc. **63** , 953-963, 1968.
[3] Mehta, J.S. and Srinivasan, S.R. *Estimation of the mean by shrinkage to a point*, Jour. Amer. Statist. Assoc. **66** (233), 86-90, 1971.
[4] Govindarajulu, Z. and Sahai, H. *Estimation of parameters of a normal distribution with known coefficient of variation*, Reports of Statistical Application Research Union of Japanese Scientists and Engineers **91** , 85-98, 1972.
[5] Das, B. *Estimation of μ^2 in normal population*, Cal. Statist. Assoc. Bull **24**, 135-140, 1975.
[6] Srivastava, V.K., Dwivedi, T.D. and Bhatnagar, S. *Estimation of the square of mean in normal population*, Statistica, anno XL **4** , 455-466,1980.
[7] Rao, V.N. and Singh, J. *A note on estimation of μ^2 in normal density*, Jour. Ind. Soc. Agril. Statist. **34** (1), 82-84, 1982.
[8] Bhatnagar, S. *On the use of population variance in estimating mean*, Jour. Ind. Soc. Agril. Statist. **38** , 403-409, 1986.
[9] Singh, H.P. and Katyar, N.P. *A generalized class of estimators for common parameters of two normal distribution with known coefficient of variation*, Jour. Ind. Soc. Agril. Statist. **40** (2), 127-149, 1988.
[10] Singh, H.P. *Estimation of parameters in normal parent*, Jour. Ind. Soc. Agril. Statist **XL11** (1), 98-10, 1990.
[11] Jani, P.N. *A class of shrinkage estimators for the scale parameter of the exponential distribution*, IEEE Trans. Reliability **40** , 68-70, 1991.
[12] Kourouklis, S. *Estimation in the two-parameter exponential distribution with prior information*, IEEE Trans. Reliability **43** (3), 446-450, 1994.
[13] Singh, H.P. and Singh, R.*A class of shrinkage estimators for the variance of a normal population*, Microelectronand Reliability **37** (5), 863-867, 1997.
[14] Singh, H.P. and Shukla S. K. *A family of shrinkage estimators for the square of mean in normal distribution*, Statistical Papers **44** , 433-442, 2003.
[15] Singh, H.P. and Saxena S. *A class of shrinkage estimators for variance of a normal population*, Brazilian Journal of Probability and Statistics **17** , 41-56, 2003.
[16] Pareto, V.*Cours'd Economia Politique*, Roge and Cie, Lassanne and Paris , 1897.
[17] Rytgaard, M. *Estimation in Pareto Distribution*, ASTIN Bulletin **20** (2), 201-216, 1990.
[18] Singh, D.C., Singh P.and Singh, P.R. *Shrunken estimators for the scale parameter of classical Pareto Distribution*, Microelectronand Reliability **36** (3), 435-439, 1996.
[19] Prakash, G., Singh D.C. and Singh R.D. *Some test estimators for the scale parameter of classical Pareto Distribution*, Journal of Statistical Research **40** (2), 41-54, 2006.
[20] Prakash, G. *Some estimators for the Pareto Distribution*, Journal of Scientific Research **1** (2), 236-247, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1245 – 1273

Spatial decision making under determinism vs. uncertainty: A comparative multi-level approach to preference mapping

Aygün Erdoğan*† and Paul D. Zwick‡

Abstract

The aim of this study is to highlight the importance of uncertainty assessments in GIS-based multi-attribute land-use decision making. To this end, and based on the basic premise that *uncertainty makes a difference,* it makes use of an existing deterministic goal-driven and hierarchical GIS-based land-use conflict model known as LUCIS (Land-Use Conflict Identification Strategy), the aim of which is to create a land-use conflict map between agricultural, urban and ecologically sensitive land-use preferences for future planning scenarios. Being confined to its agricultural preference (overall goal) mapping, the newly developed uncertainty models and maps are compared with their corresponding deterministic models and maps at each level of the LUCIS hierarchy. The comparative models are applied to the case of Hillsborough County in Florida, which is characterized by a high level of conflict between the three land uses. Different levels of differences in terms of pattern/shape/form and the degree of agricultural land-use suitability are identified and assessed at all levels of the hierarchy.

Keywords: Multi-attribute decision making (MADM/MCDM), Land-Use Conflict Identification Strategy (LUCIS), determinism, uncertainty, probability, fuzzy logic.

2000 AMS Classification: 62C99

Received : 05.08.2014 *Accepted* : 22.06.2015 *Doi* : 10.15672/HJMS.20157712084

*Department of Urban and Regional Planning, Karadeniz Technical University; Shimberg Center for Housing Studies, University of Florida, Email: aygun@ktu.edu.tr

†Corresponding Author.

‡Department of Urban and Regional Planning; School of Landscape Architecture and Planning, College of Design, Construction and Planning, University of Florida, Email: pdzwick@ufl.edu

1. Introduction

The limitations associated with the classical Boolean logic representation of spatial data in standard geographic information systems (GIS) [41; 6; 1], which is "crisp, deterministic, and precise in nature" [1:143], has resulted in the integration of multi-attribute decision making (MADM) techniques (referred to in general as multi-criteria decision making – MCDM – in MADM literature) with GIS [29]. This approach facilitates a wide range of analytical procedures [7], and has gained increasing interest among modelers over the last two decades, based on its ability to assess uncertainty in spatial MCDM process.

GIS-based or spatial MADM is based on the discrete representation of spatial data, generally in the form of a hierarchical structure [28]. Unlike the multi-objective process of MCDM [28; 15], as in all multi-attribute decision making approaches this process involves the definition of objectives, the choice of criteria for measuring these objectives and their standardization, the criteria weighting that reflects the decision-makers' preferences, and an aggregation of the weighted standardized criterion values, allowing the alternatives to be ranked, after which the best alternative will be selected [29; 30; 27].

1.1. Uncertainty analysis in land-use planning and environmental management in spatial MCDM literature, and context of the current study. The uncertainties in the decision-making process related to planning or environment-related problems, including land-use suitability, are distinguishable in three dimensions, that is, (1) location, (2) level, and (3) nature of the uncertainty [30]. In their review of some basic works (see [45; 36; 46]) Mosadeghi et al. [30] suggest a linkage between uncertainty analysis in MCDM and the dimensions of uncertainty with respect to uncertainty in environmental decision making (Figure 1). As can be seen in the figure, uncertainties that are stochastic in nature are found in the context and model structure, and are related to the decision makers' preferences and knowledge of the MCDM process, while epistemic uncertainties are found in the context, modeling technique and input, and are related to model uncertainty. By definition, stochastic uncertainty, which is inherent in the context of natural, behavioral, social, economic, and cultural systems, is random in nature and cannot be eliminated [18; 30]. On the other hand, epistemic uncertainties are a result of imperfect or incomplete knowledge, and can be reduced through empirical efforts and high-quality data, monitoring and longer time series [18; 30; 32].

The following list explains the sources of uncertainty found in modeling that may be dealt with in an uncertainty analysis in which stochastic uncertainties are excluded. Uncertainty in the final result may originate from any of these stages [41], or may be found in one or more of the different stages of the spatial (GIS-based) MCDM process that may propagate in the final result [32]. As is common in many works [41; 18; 12; 11; 13; 40; 30; 32; 27], these stages of the modeling process, which are characterized by assessable (i.e., epistemic) uncertainty, can be listed as in the following with reference to the locational dimension presented in Figure 1.

1. Selecting a particular/appropriate model (model structure);
2. Setting or defining the problems, goals and/or objectives (model structure);
3. Identifying appropriate attribute/criteria and/or parameters (model structure);
4. Obtaining high-quality data with minimal measurement and data processing (context and input) or algorithm (model technique) errors;
5. Decision making to obtain standardized criterion maps (context and model technique);
6. Decision making for assigning of weights (model structure); and
7. Interpretation of the final results (context and model technique).

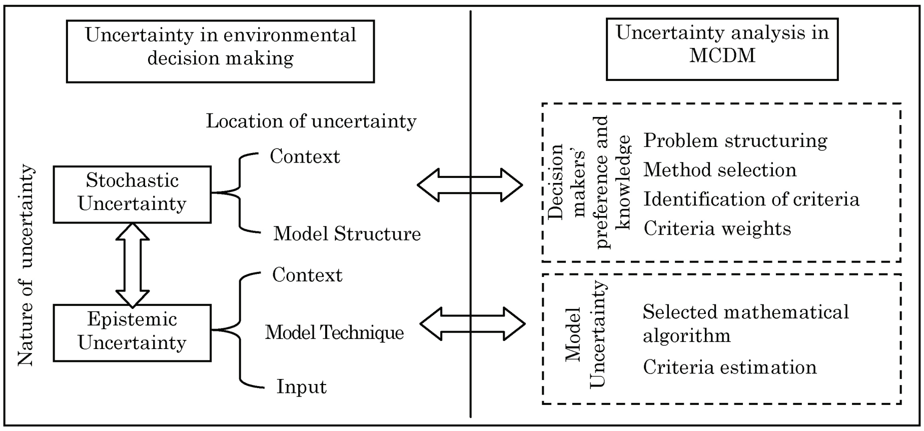


Figure 1. Linkage between uncertainty terminology in environmental decision-making science and MCDM
Source: [30:1104]

Although the number of studies that focus on uncertainty assessments in MCDM are increasing in number, they are still considered insufficient by many scholars who concentrate on the requirement for the proper expression of uncertainty in GIS-based works (see e.g., [14; 35; 40]). The shortfall, specifically, is in the quantification of uncertainty in decision making and policy assessment concerning land-use planning [32] and for environmental processes [18].

The level of resolution to the problem of uncertainty in the above listed stages of the modeling process in land-use or environmental decision making differs in existing literature. That is, while in some studies uncertainty is dealt with to a greater extent in terms of both the number of works and the variety of techniques used, others are subject to less attention by the modelers. For example, although the number of works that consider uncertainty in relation to the selection of the model, goal/objectives, criteria/parameters (stages 1 to 3) above is very limited [29; 11], those that are related to stage 4 on data quality and processing is relatively high (see e.g., [2; 26; 39]). That said, it is also known that in MCDM methods, the input data is generally assumed to be error free (see e.g., [41]), precise and accurate [29]. The majority of spatial MCDM literature focuses on stages 5 and 6 [16], and to a much lesser degree, on stage 7, which deal with decision making in terms of criteria standardization and weight assignment, and results interpretation, respectively. In literature, different MCDM techniques for dealing with uncertainty, especially in stages 5 and 6, have been developed since the first introduction of this decision-making process into the fields of economics and finance in the 1960s [30]. These multi-attribute (multi-criteria) evaluation methods include the weighted linear combination (WLC), and as an extension to its limitations, the ordered weighted averaging (OWA), as well as other additive techniques, such as multi-attribute value/utility theory (MAVT/MAUT) and analytic hierarchy process (AHP). Some WLC-variant decision rules are also included, such as ideal point methods (e.g., Technique for Order Preference by Similarity to Ideal Solution-TOPSIS) and concordance methods (e.g., Elimination et Choice Translating Reality-ELECTRE, Preference Ranking Organization Method for Enrichment Evaluations-PROMETHEE), and also some other methods that utilize theories of Fuzzy sets, Random sets and Game [28; 41; 29; 16; 30].

Literature contains a number of works that discuss the similarities and differences between uncertainty analysis (UA) and sensitivity analysis (SA); and based on these, it can be stated that while some authors claim there is no distinct difference between the two concepts and that they may be used interchangeably [35; 16; 30; 31], others consider them to be separate (see e.g., [28; 27;], but still emphasize the need of their integrated use. As Ligmann-Zielinska and Jankowski [27] point out, UA is used to quantify the variability of outcomes in a multi-criteria evaluation, given the model input uncertainty, whereas SA is used to identify which criteria or criteria weights are most responsible for this variability.

In spatial MCDM, works on land-use suitability or environmental management uncertainty are dealt with using different methods, based on different theoretical backgrounds, assumptions and different levels/types of data requirement. With reference to some basic works [25; 8; 28; 41; 29; 35; 16; 11; 30; 31] a summary table charting these uncertainty/sensitivity analysis methods, in addition to those that are deterministic, is presented, with respect to their modeling type/underlying theory, typology, uncertainty handling, method of criterion map combination, level of objectiveness and ease of communication to the decision makers (Table 1).

The purpose of an uncertainty analysis in decision making is to determine the risk in choosing a particular alternative [11]. Based on the above-listed basic works, it can be stated that in turning the uncertainty into 'risk', in addition to either data-driven traditional *(a priori)* probabilistic (e.g., logistic regression and Monte Carlo simulation), data and knowledge-driven conditional *(a posteriori)* probabilistic (e.g., Bayesian network) and their extensions (e.g., Dempster-Shafer Belief functions) or artificial intelligence (e.g., neural network and fuzzy sets) methods, there are many other approaches, including analytical error propagation, one-at-a-time (OAT), indicator-based (distance-based) analysis, variance-based analysis, methods using random sets theory and game theory (Table 1).

In spatial MCDM literature, which deals mainly with subjects of land-use suitability in land-use planning and environmental management, uncertainty is handled mainly within the 5^{th} and 6^{th} stages of the modeling process described earlier.

In environmental GIS-based MCDM studies, Falk et al. [18] assess the uncertainty estimates of the outcomes of a deterministic environmental model (Revised Universal Soil Equation-RUSLE), along with its input parameters; while Store and Kangas [41] integrate expert knowledge with a spatial multi-criteria evaluation to model GIS habitat suitability. As a resolution to the classical Boolean representation of GIS in uncertainty modeling, and to make empirical data cost savings, Store and Kangas [41] utilize expert knowledge that is based on the theoretical background of MAUT in habitat suitability. For cost saving purposes, Castrignanò et al. [10] opted for multivariate geostatistics in GIS, utilizing ancillary less-expensive information to improve the estimate uncertainty of a soil quality index. Facing the same GIS representation problem, Avdagic et al. [6] and Reshmidevi et al. [37] developed a methodology to integrate a Mamdani-type fuzzy inference rule base in GIS in land valorization for land-use planning and land suitability for particular crops, respectively. In addition, Reshmidevi et al. [37] used the local knowledge of farmers and experts, and compared two different aggregation methods: WLC and Yager's aggregation. Based on the same GIS limitation, but criticizing the integration of Mamdani-type fuzzy logic in GIS, Adhikari and Li [1] utilize a Sugeno-type fuzzy inference. Similar to Falk et al. [18], who utilized Bayesian melding in a cell-based GIS environment, O'Brien et al. [35] developed a tool called CaNaSTA (Crop Niche Selection in Tropical Agriculture) to define site suitability for particular crops and forages using sparse and uncertain data based on Bayesian modeling. In their tool, called the Catchment Evaluation Decision Support System (CEDSS), which enables the explicit

Table 1. Deterministic vs. uncertainty or sensitivity analysis methods in multi-attribute modeling that utilize GIS and other spatial analysis software in land-use suitability or environmental management

Deterministic / Uncertainty or Sensitivity analysis methods[1]	Modeling type / "Underlying theory"	Typology	Uncertainty handling	Method of criterion map combination	Objectiveness / Communication
Boolean logic (operations)	Deterministic "Determinism" Traditional set theory	The most traditional	No uncertainty assumed Probability is either 0 or 1 Set membership value is either 0 or 1	No weighting (simple overlay of 0-1 maps)	Easy communication to decision makers
Binary evidence -extension to Boolean logic				WLC	
Index overlay -extension to binary evidence				WLC/OWA/MAVT/ AHP/ideal point methods/ concordance methods	
Logistic regression	Data-driven probabilistic "Probability theory"	Traditional probability (*a priori*)	Uncertainty due to limitation in knowledge (epistemic) or randomness in occurrence of an event (stochastic) Based on probability density, probability distribution Probability is between 0 and 1	Methods listed in the leftmost column used for combining the criterion maps. However, in case they are used for criterion map estimation then Probabilistic additive weighting/OWA/ MAUT/ AHP/ideal point methods/ concordance methods is/are used (In addition, MAUT is also used in standardizing criterion maps)	Being data-driven and a priori, more objective Relatively complicated
Generalized Linear and Generalized Additive Models -extension of logistic regression					
Monte Carlo Simulation					
Bayesian network	Both data-driven and knowledge-driven probabilistic "Bayesian theory"	Conditional (Bayesian) probability (*a posteriori*)	Based on *a priori* probability and knowledge-base a posteriori probability is obtained with the principle of excluded middle		Being knowledge-driven, and to a certain extent, being a posteriori, more subjective Complicated
Dempster-Shafer Belief functions	Knowledge-driven probabilistic "Dempster-Shafer Belief theory"	Extension of Bayesian probability	Makes a distinction between probability and ignorance removing the assumption of excluded middle		
Classification and regression trees -based on decision trees	Data-driven for robust results but allow knowledge-driven assessment for deterministic or probabilistic rule base	Artificial intelligence[2]	Tolerant of imprecision, ambiguity, vagueness, uncertainty		Not necessarily more accurate but "more informed" decisions "Black box" to the decision makers
Neural network					
Cellular automata					
Fuzzy logic (operations)	Fuzzy sets theory "Possibility theory"		Uncertainty due to imprecision of knowledge or the ambiguity of an event, i.e., to which degree an event occurs Set membership value is between 0 and 1	Fuzzy additive weighting/ Fuzzy MIN/Fuzzy MAX/OWA/AHP/ ideal point methods/ concordance methods is/are used	

[1] Other basic uncertainty or sensitivity analysis methods not detailed here are analytical error propagation, one-at-a-time (OAT), indicator-based (distance-based) analysis, variance-based analysis, methods using random sets theory and game theory.
[2] Evolutionary (genetic) algorithms is a multi-objective decision making (MODM) method that utilizes artificial intelligence, and so is not included in the table.

Source: Compiled from the explanations found in [25; 8; 28; 41; 29; 35; 16; 11; 30; 31]

visual exploration of uncertainties in decision making resulting from both weights and attribute (criterion) values in GIS-based catchment management, Chen et al. [11] utilize an indicator (distance)-based method facilitated by an OAT approach. On the other hand, Ligmann-Zielinska and Jankowski [27] use a Monte Carlo simulation in addition to a variance-based analysis in an uncertainty analysis in their UA-SA integrated methodology aimed at defining habitat suitability for a wetland plant.

More transparent graphical display facilities of GIS, such as the work by Chen et al. [11], have taken a novel approach, visualizing the uncertainties in criterion weighting based especially on the AHP method, and thus its pairwise comparisons. In this respect, to evaluate epistemic uncertainties in coastal land-use planning decisions Mosadeghi et al. [31] examine the sensitivity of AHP weighting decisions to input uncertainties, and to this end, combine the conventional UA with the visualization capability of GIS and the Monte Carlo simulation algorithm. Similarly, Chen et al. [12] developed a GIS-based AHP-SA tool that utilizes the OAT method to assess the behavior and limitations of a GIS-based irrigated cropping land-use suitability model. The tool provides access to an interactive range of user-defined simulations to evaluate the dependency of the model output on the weights of the input parameters, identifying the criteria that are sensitive to weight changes. In further developing their work (AHP-SA), Chen et al. [13] developed the AHP-SA2 to increase the tool's efficiency, while also improving its flexibility and enhancing its visualization capability to analyze the weight sensitivity resulting from both direct and indirect weight changes using the OAT technique. Likewise, based on the subjectivity limitation of AHP, Ahmad et al. [3] developed a new technique called the "Objective Spatial Analytic Hierarchy Process (OSAHP)", combining AHP with regression modeling to identify potential agroforestry areas using GIS. With the aim of sustainable development and consensus building, and considering the uncertainties in the land-use planning process, Soltani et al. [40] utilize a GIS-based urban land-use model combined with UA. In their GIS-based MCDM they used AHP, sensitivity analysis, Monte Carlo simulation and probability classification methods, and made use of the visual spatial representation of the results for different stages of the decision-making process under different conditions.

As in the above-mentioned literature, this study deals mainly with the uncertainty in the 5^{th} and 6^{th} stages (decision making on criteria standardization and weight assignment) of the spatial MCDM modeling process listed earlier, and with the 7^{th} stage to the extent of discussing the possibility of different results based on different interpretations of the results of the modeling.

In doing this, rather than carrying out classical sensitivity analysis procedures on criterion values and weights, the intention is to examine the differences between the results of a deterministic approach and an uncertainty approach using standardized criterion maps at lower levels of a hierarchical GIS-based multi-criteria model, and those of weighted and aggregated maps at higher levels. By assessing the differences at each level (multi levels) in the two modeling approaches, and between their equivalent overall goal (preference) maps, this study aims to show that *uncertainty makes a difference* in the ranking and the spatial pattern of the alternatives in land-use decision making, and presents empirical proof of the importance of uncertainty assessment in spatial multi-criteria modeling.

In this respect, the study does not deal with the question of uncertainty in terms of the potentially subjective decisions given by the decision makers, in this case, the two modelers. In other words, the study does not make a sensitivity analysis of the criterion values and weighting of the two models, but rather shows that the deterministic results should not be seen as the only solution set with a particular ranking and spatial pattern of alternatives in land-use suitability, and reveals that they are subject to change under different conditions of decision making, which is characterized by uncertainty.

As mentioned earlier, although there is an increasing number of works on uncertainty assessment, related especially to the 5^{th} and 6^{th} stages of spatial multi-criteria modeling, there has to date been no one-to-one comparison of the deterministic and uncertainty maps at each level of an MCDM land-use suitability model in a GIS environment.

With this study, two main types of uncertainty method, being probability and fuzzy set theories, in addition to MAUT (Table 1), were used to obtain standardized criterion maps at the lowest levels of the hierarchical structure of the existing deterministic model. Then, a weighting process was carried out, which included trade-offs at levels under the goal level and entropy at the goal level compared to existing model's AHP at all levels of hierarchy with two exceptions (i.e., for one lower level sub-objective and for goals). After weighting, each criterion map at the lower levels was aggregated at the higher levels to obtain the preference (overall goal) map for a particular land-use type via either modeling approach (deterministic vs. uncertainty). In these stages, each map pair from either of the modeling approaches at each level of the hierarchy was compared for the case study area, being Hillsborough County in the state of Florida.

The deterministic model used in this study is the Land-Use Conflict Identification Strategy (LUCIS), the structure of which is described in brief in the following section.

1.2. Deterministic spatial multi-attribute land-use modeling: Land-Use Conflict Identification Strategy (LUCIS). The Land-Use Conflict Identification Strategy (LUCIS) is a deterministic MADM process and "a goal-driven GIS model that produces a spatial representation of probable patterns of future land use" [9:9]. In order to assess the conflicts between the three main land-use types (agricultural, urban, and ecologically sensitive) and possible future land-use patterns, models are established to obtain preference maps related to each of these land uses (Figure 2). Even though the complete LUCIS deals with conflict identification based on three different land uses, and in total involves a 6^{th} level at the top of the hierarchical structure, the scope of this study is limited up to 5^{th} level, and to the agricultural land use (Figure 2). In this respect, the uncertainty maps obtained in this study, like their corresponding deterministic equivalents from existing models, consist of the overall goal map, referred to as the preference map hereafter, at the top of the hierarchical structure, followed by maps charting the goals, objectives, sub-objectives and lower level sub-objectives at the lower levels.

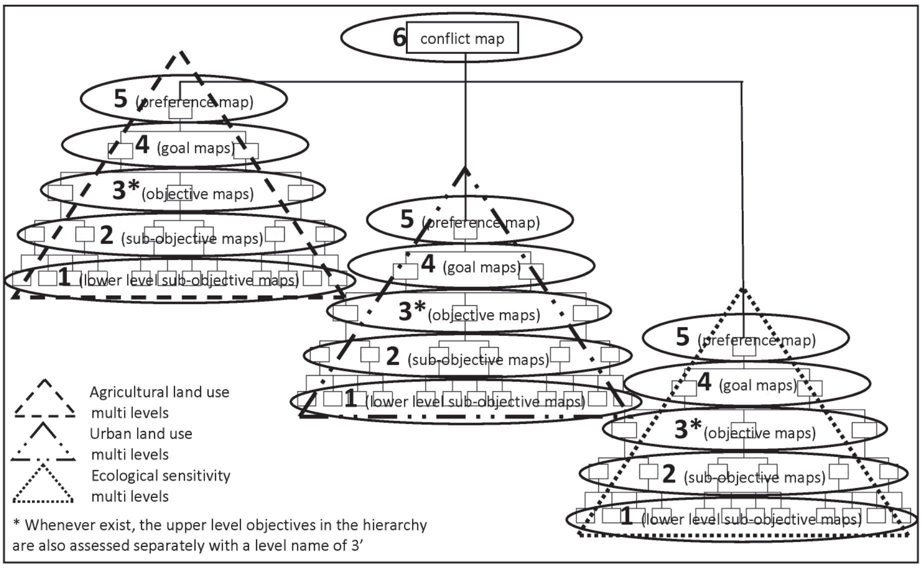


Figure 2. Symbolic representation of multi-level LUCIS hierarchies (study covers the levels concerning agricultural land use on the left, the preference map being at the top)
Source: Adapted from [9:231,233,236]

The related numbering, naming and a short description of the LUCIS hierarchical levels for the agricultural land-use preference map seen on the left part of Figure 2 is

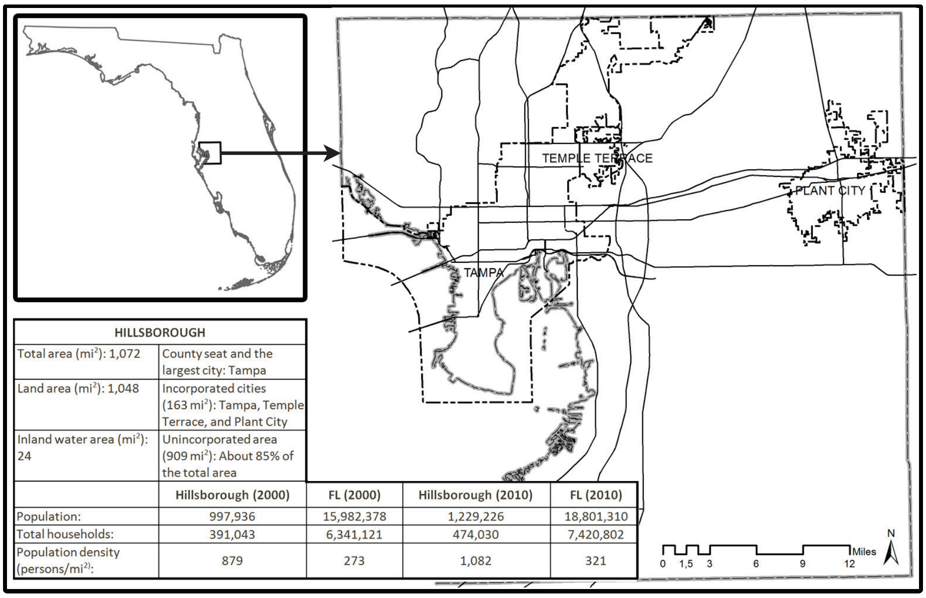
Table 2. Numbering, naming, and a short description of the LUCIS hierarchical levels for the agricultural land-use preference map

Level 4 Goal maps	Level 3' Upper level objective maps*	Level 3 Objective maps	Level 2 Sub-objective maps	Level 1 Lower level sub-objective maps
Row crops land suitability (1)	-	Physical suitability (11)	Soils suitability (111)	a:Grass; b:Strawberries; c:Corn; d:Sugarcane; e:Cabbage; f:Peppers; g:Soybeans; h:Snapbeans; i:Watermelons; j:Peanuts; k:Cucumbers
			Land-use suitability (112)	-
		Proximity suitability (12)	Local markets proximity (122)	a:City population; b:Row crops distance
			Major roads proximity (123)	-
		Land value suitability (13)	-	-
Livestock suitability (2)	High-intensity livestock suitability (2A)	Physical suitability (21)	Land-use suitability (211) Distance to open water resources (212) Aquifer recharge suitability (213) Soils suitability (214) Distance to existing urban areas (215)	- - - - -
		Proximity suitability (22)	Local markets proximity (221) Major roads proximity (223)	- -
		Land value suitability (25)	-	-
	Low-intensity livestock suitability (2B)	Physical suitability (23)	Land-use suitability (231) Distance to open water resources (232) Aquifer recharge suitability (233) Soils suitability (234)	- - - -
		Proximity suitability (24)	Local markets proximity (241) Major roads proximity (243)	- -
		Land value suitability (26)	-	-
Specialty farming suitability (3)	-	Physical suitability (31)	Land-use suitability (311) Distance to open water resources (312) Aquifer recharge suitability (313) Soils suitability (314)	- - - -
		Proximity suitability (32)	Proximity to processing plants (321) Major roads proximity (323)	- -
		Land value suitability (33)	-	-
Nursery suitability (4)	-	Physical suitability (41)	Land-use suitability (411) Parcel size suitability (412)	- -
		Proximity suitability (42)	Local markets proximity (421) Major roads proximity (423)	- -
		Land value suitability (43)	-	-
Timber suitability (5)	-	Physical suitability (51)	Land-use suitability (511) Aquifer recharge suitability (513) Soils suitability (514) Parcel size suitability (515)	- - - -
		Proximity suitability (52)	Local markets proximity (521) Major roads proximity (522)	- -
		Land value suitability (53)	-	-

presented in Table 2, in which all of the goals and objectives at all levels are phrased in such a way that they are tried to be maximized in the decision-making process. As is clearly apparent in Table 2, the LUCIS agricultural land-use hierarchical levels, each of which is in fact a GIS map layer, follow a naming convention that is composed of an alphanumeric code for each different level. For example, Level 4: Goal map 1, Level 3: Objective 11, Level 2: Sub-objective 111 and Level 1: Lower Level Sub-objective criterion maps under sub-objective 111 are named respectively with codes ag1; ag1o11; ag1o11o111; and criterion maps under sub-objective ag1o11so111, which are named with a letter a–k to ensure ease in following, and since the maps at these levels are only a few in number.

2. Study area and data

Hillsborough County is located on the west coast of central Florida (Figure 3). It has total of surface area of 1,072 square miles (1,048 sq mi of land and 24 sq mi of inland water). Tampa is the County seat and the largest city in Hillsborough, in which there are two more municipal cities: Temple Terrace and Plant City [23]. It is a rapidly urbanizing county [44] with a population increase of 23.2 percent (from 997,936 to 1,229,226) and a population density increase from 879 to 1082 persons/sq mi between 2000 and 2010 [43]. The rapid and continuous urban development, which has been mainly in the form new suburban construction, especially into the more rural, unincorporated part of the county [23], has caused both the environmental degradation of natural resources, such as soil erosion and compaction, deforestation and disturbance to aquifers [44], and a decrease in valuable agricultural lands, which makes up one of the most important production capacities in the state total [38].



HILLSBOROUGH	
Total area (mi^2): 1,072	County seat and the largest city: Tampa
Land area (mi^2): 1,048	Incorporated cities (163 mi^2): Tampa, Temple Terrace, and Plant City
Inland water area (mi^2): 24	Unincorporated area (909 mi^2): About 85% of the total area

	Hillsborough (2000)	FL (2000)	Hillsborough (2010)	FL (2010)
Population:	997,936	15,982,378	1,229,226	18,801,310
Total households:	391,043	6,341,121	474,030	7,420,802
Population density (persons/mi^2):	879	273	1,082	321

Figure 3. The study area, Hillsborough County in the state of Florida Source: Map data compiled from [19]; Tabular data compiled from [23;43]

The strong competition with an essentially high level of decision-making uncertainty among the urban, agricultural and natural land uses in Hillsborough County was the main reason for the selection of this area for a study of the impact of uncertainty on a deterministic multi-criteria land-use modeling, aiming to identify land-use conflicts (LUCIS) among the three land uses.

Aside from the annual agricultural sales of Hillsborough County, obtained from Census of Agriculture data, and the 'Critical Lands and Waters Identification Project (CLIP): Version 2.0 data [21], all other data used in the study at both the county and state level were obtained from the 'Florida Geographic Data Library' (FGDL) website [19], as the source of the most recent available data at the time of writing.

In this study, all the models for the deterministic approach were built and run using ArcGIS® software. The same software was used also for the uncertainty approach, although for some models, additional software was needed, such as, spreadsheet environment (MS Excel®) and spatial data analysis (CrimeStat®).

3. Methodology and application

In this section, the methodology and its applications to the study area will be explained in three subsequent stages. The first stage includes the development of uncertainty models for LUCIS and the comparisons with their deterministic equivalents in terms of the standardization of criterion maps (each different GIS layer) at the different hierarchical levels (levels 1, 2 and 3) prior to any weighting being applied. The second stage involves a comparison of the decision rules of the two different approaches (deterministic vs. uncertainty) for combining the criterion maps under each relevant level of hierarchy (levels 1, 2, 3, 3' and 4). In the final stage, a comparison is made for the preference maps of the two modeling approaches (level 5). The results of the two modelings of these three stages, considering all hierarchical levels of LUCIS (up to the 5^{th}), are explained in Section 4.

3.1. Comparison of newly developed uncertainty models and their existing deterministic equivalents in criteria standardization (levels 1, 2 and 3). The criteria standardization in uncertainty modeling was carried out using seven different groups of methods, each applied to a different group of maps prior to any weighting (i.e., the maps have no other sub-level maps) (Table 2). The seven groups of methods are listed in Table 3 according to the groups of criterion maps (GIS layers) to which they were applied, which are referred using their alphanumeric names described earlier.

In general, the GIS-based uncertainty models in criteria standardization were developed with reference to the characteristics of the decision variable: whenever they are numeric, the uncertainty is assumed to be a result of limited information related to the decision-making process in a particular spatial system and dealt with traditional probability [25; 4; 28] (Table 1), contrary to the unit probability of an alternative in the deterministic DM process [20; 22]. However, if the variables are categorical, and imply that the uncertainty is a result of the imprecision or ambiguity of the information or, in other words, if the variables are linguistic or fuzzy, the fuzzy set membership methods [28] are used to obtain the criterion maps. Both of these two types of maps are then compared with those obtained from the deterministic variables with binary, discrete or continuous values at each level of the hierarchy.

In the former type of variables, probabilistic maps are obtained with discrete, continuous or mixed variable values, and the transformation processes are based on probability density or cumulative probability density functions, in which most of the maps can be considered to be data-driven, based on objective probabilistic methods (Table 1) using relative frequency (or area) distributions. The only exceptions to this are the two lower level sub-objectives handled by MAUT, in which the derivation of utility functions includes the assessment of the decision maker's expected utility. The remaining assessments of uncertainty involve the use of fuzzy logic (Table 1) by means of linguistic variables.

In Table 4 below the detailed methodology applied to the seven different groups of criterion maps are explained in terms of both the deterministic and uncertainty approaches.

3.2. Comparison of decision rules in criteria aggregation and weighting in the deterministic and uncertainty models (levels 1, 2, 3, 3' and 4). In the deterministic modeling, the decision rule for combining the criterion maps at each weighting

Table 3. Seven groups of uncertainty methods for criteria standardization, and the criterion maps (GIS layers) to which they were applied

Methods	Uncertainty method type	Hierarchical level	Criterion map name	Objective (in terms of minimization or maximization)
Method 1	Expected utilities based on frequencies multiplied by a particular value (yield)	Level 1	Lower Level Sub-objective under Sub-objective ag1o11so111	- row crops-physical-soils
Method 2	Utility functions and utility function multiplied by a particular value (probability of standard deviation of the prediction)	Level 1	Lower Level Sub-objective under Sub-objective ag1o12so122	- row crops-proximity-local markets
Method 3	Fuzzy set membership (and fuzzy overlay) based on expert knowledge and spatial MCDM literature	Level 2	Sub-objective ag1o11so112 Sub-objective ag4o41so411 Sub-objective ag5o51so511 Sub-objective ag2o21so213 Sub-objective ag2o23so233 Sub-objective ag3o31so313 Sub-objective ag5o51so513 Sub-objective ag2o21so214 Sub-objective ag2o23so234 Sub-objective ag3o31so314 Sub-objective ag5o51so514 Sub-objective ag4o41so412 Sub-objective ag5o51so515	- row crops-physical-land-use - nursery-physical-land-use - timber-physical-land-use - livestock-high-intensity livestock physical- aquifer recharge - livestock-low-intensity livestock physical-aquifer recharge - specialty farming-physical-aquifer recharge - timber-physical-aquifer recharge - livestock-high-intensity livestock physical-soils - livestock-low-intensity livestock physical-soils - specialty farming-physical-soils - timber-physical-soils - nursery-physical-parcel size - timber-physical-parcel size
Method 4	Fuzzy set membership based on the mean and standard deviations of already grouped data with respect to their fuzzy membership values, based on expert knowledge and spatial MCDM literature	Level 2	Sub-objective ag1o12so123 Sub-objective ag4o42so421 Sub-objective ag5o52so521 Sub-objective ag4o42so423 Sub-objective ag5o52so522	- row crops-proximity-roads - nursery-proximity-local markets - timber-proximity-local markets - nursery-proximity-roads - timber-proximity-roads
Method 5	Functional transformation of probabilities based on areas	Level 2	Sub-objective ag2o21so211 Sub-objective ag2o23so231 Sub-objective ag3o31so311	-livestock-high-intensity livestock physical-land-use -livestock-low-intensity livestock physical-land-use -specialty farming-physical-land-use
Method 6	Fuzzy set membership based on the mean and standard deviations of already grouped data with respect to their functional transformation of probabilities, based on areas	Level 2	Sub-objective ag2o21so212 Sub-objective ag2o23so232 Sub-objective ag3o31so312 Sub-objective ag2o21so215 Sub-objective ag2o22so221 Sub-objective ag2o24so241 Sub-objective ag3o32so321 Sub-objective ag2o22so223 Sub-objective ag2o24so243 Sub-objective ag3o32so323	-livestock-high-intensity livestock physical-open water -livestock-low-intensity livestock physical-open water -specialty farming-physical-open water -livestock-high-intensity livestock physical-existing urban -livestock-high-intensity livestock proximity-local markets -livestock-low-intensity livestock proximity-local markets -specialty farming-proximity-processing plants -livestock-high-intensity livestock proximity-roads -livestock-low-intensity livestock proximity-roads -specialty farming-proximity-roads
Method 7	Fuzzy set membership based on enumeration derived from spatial k-means clustering and non-spatial mean and standard deviations	Level 3	Objective ag1o13 Objective ag2o25 Objective ag2o26 Objective ag3o33 Objective ag4o43 Objective ag5o53	-row crops-land value -livestock-high-intensity livestock-land value -livestock-low-intensity livestock-land value -specialty farming-land value -nursery-land value -timber-land value

Table 4. Detailed deterministic and uncertainty methodology applied to the seven different groups of criterion maps

<table>
<tr><th></th><th>Level</th><th>Map name</th><th>Aim</th><th>Deterministic models</th><th>Uncertainty models</th></tr>
<tr><td>Method 1</td><td>Level 1 maps</td><td>(a-k: different types of crops) under so111</td><td>maximize soil suitability for each crop type</td><td>Score assignment by linearly increasing values between 1and 9 to either individual or classified increasing crop yield amounts</td><td>Expected utility estimation for each row crop type by the number of pixels (i.e., area) of each row crop type, multiplied by the yield value of that crop, and divided by the total of these products
(spreadsheet used for floating point rasters, conditional map algebra operation in GIS used for value assignment)</td></tr>
<tr><td>Method 2</td><td>Level 1 maps</td><td>(a:city population; b:row crops distance) under so122</td><td>maximize proximity to local markets for row crops (cities' population and row crop areas)</td><td>- Results from a deterministic interpolation method (inverse distance weighting – IDW) on the cities of the county with non-zero population were used for reclassifying the raster
- An Euclidian distance map of row crop areas used for reclassification based upon the mean and 1/4 standard deviation distances found in the zonal statistics table for row crop areas [9]</td><td>- Cities' populations (including neighbor counties) interpolation by a geostatistical process of kriging that provided a prediction and its variance raster. Prediction surface is used with an estimated utility function by using indifference method for standardization. This 0-1 range prediction map was multiplied by the probability of square root of the variance raster to give higher weights to the values having less errors and vice versa.(1)
- Row crops distance standardized utility scores was estimated by application of a utility function to the raw scores obtained by Euclidian distances.(1)</td></tr>
<tr><td rowspan="4">Method 3</td><td rowspan="4">Level 2 maps</td><td>so112, so411, so511</td><td>maximize agricultural land-use suitability in terms of land cover, soils and parcels</td><td rowspan="4">Expert knowledge and spatial MCDM literature used in assigning the deterministic values of either 1 and 9 or all or some of the values between 1 and 9. The higher and the highest suitability (9) values were given to the existing and higher potential areas or to the lower criticality areas for the respective five agricultural goals, while the lower and the lowest (1) suitability values were given to areas that have a reverse impact on suitability</td><td rowspan="4">Conversion of deterministic assignments (1-9) into linguistic rankings (1-very low and 9-very high) to use fuzzy large or small transformation in GIS to obtain different levels of standardized suitability scores between 0 and 1.(2) Contrary to final combination method max cell statistics operation in the deterministic approach, of land cover and soils map for so112 map a fuzzy OR overlay was used here, however, similar to deterministic approach a focal statistics was used for so213, so233, so313 and so513 maps.</td></tr>
<tr><td>so412, so515</td><td>minimize and maximize parcel size for nursery and timber respectively</td></tr>
<tr><td>so214, so234, so314, so514</td><td>maximize drainage capacity of soils</td></tr>
<tr><td>so213, so233, so313, so513</td><td>maximize disturbance to aquifers</td></tr>
<tr><td rowspan="2">Method 4</td><td rowspan="2">Level 2 maps</td><td>so123, so423, so522</td><td>maximize proximity to roads for row crops, nursery and timber</td><td rowspan="2">First, Euclidian distance raster maps were created from major roads and from local market objects that were considered to be the median center of vacant lands for nursery (so421) and lumber yd/mill for timber (so521).
Then, a reclassification was made [9] based upon the mean and 1/4 standard deviation distances found in the zonal statistics for the selection of a set of objects related to each of the respective sub-objectives: row crop areas for so123, plant nursery for so421 and so423, and timber for so521 and so522.</td><td rowspan="2">Uncertainty models for these criterion maps were derived from fuzzy set memberships large transformation function in GIS by grouping of previous land-use sub-objective maps (so112, so411 and so511).(3)</td></tr>
<tr><td>so421, so521</td><td>maximize proximity to local markets for nursery and timber</td></tr>
<tr><td rowspan="2">Method 5</td><td rowspan="2">Level 2 maps</td><td>so211, so231</td><td>maximize land-use suitability for high- and low-intensity livestock</td><td rowspan="2">Expert knowledge and MCDM literature used in assigning the deterministic suitability scores from 1 to 9 to the selected and rasterized objects on the related fields out of the parcel data</td><td rowspan="2">Based on computation of probabilities and the functional transformation of the probabilities for the areas found to have been given a suitability score greater than 1 in the respective deterministic models. (4)</td></tr>
<tr><td>so311</td><td>maximize land-use suitability for specialty farming</td></tr>
</table>

Table 4. (cont.)

Method 6	Level 2 maps	so212, so232, so312	maximize distance to open water areas for high-intensity livestock; and wetlands and open water areas for low-intensity livestock and specialty farming	- The same methodology used in the method 4 above based on the Euclidian distance raster maps obtained by selected objects that the distance is either aimed to be maximized or minimized. - The distance maps for the major roads were the same as prepared for so123, so423 and so522. - For the zonal statistics table a set of selections of the objects, which had a suitability score of 9 from so211 for so212, so215, so221 and so223; from so231 for so232 and so243; from so311 for so312; a selection of miscellaneous agriculture or pasture parcels for so241; and a selection of orchard/citrus for so321 and so323 were used. - Linearly increasing or decreasing suitability scores between 1 and 9 was determined by whether the goal is the maximization or minimization of distances from the respective sources	The same Euclidian distance maps were used as a base in the corresponding models with a similar way of uncertainty assessment as in method 4 above. However, there were differences in terms of - number of mean-standard deviation pairs of zonal statistic raster maps - number of groups of selections from raster maps - the way that linguistic hedges were ordered (interfering in this case) - the way that constant rasters were created The models ended with either fuzzy small (for so212, so232, so312, so215) or large (for so221, so241, so321, so223, so243, so323) transformation functions with the default mid-point and spread values. (5)
		so215	maximize distance to existing urban areas for high-intensity livestock		
		so221, so241 so321	maximize proximity to packing plants and food processing parcels for high- and low-intensity livestock and specialty farming, in addition to four main restaurants in the county for high-intensity livestock		
		so223, so243, so323	maximize proximity to roads for high- and low-intensity livestock and specialty farming		
Method 7	Level 3 maps	o13, o25, o26, o33, o43, o53	maximize land value suitability for row crops, high- and low-intensity livestock, specialty farming, nursery, timber	Just values (market value) per acre for a set of selected parcels, which were greater and equal to 1 acre, for o13, o25, o26, o33 and o53 and 0.2 acre for o43 to exclude the sliver areas, included 'crops' and 'pasture'; 'dairies/feedlots', 'packing plants', 'poultry/bees/fish'; 'pasture', 'vacant acreage', 'miscellaneous agriculture'; 'orchard/citrus'; timber'; and 'plant nursery', respectively. Then, the mean and standard deviation of the just value/acre was used to update the vector data and then to reclassify its (1-9 value) raster form later [9]. Score (1) was given to parcels with 'header' and 'note' information as their land-use description.	The uncertainty in obtaining the utilities for suitability of the land values was handled by fuzzy set membership functions based on both alternatives' (i.e., parcels') spatial and non-spatial aspects. (6)

(1) See Appendix 1 for details of uncertainty method 2
(2) See Appendix 2 for details of uncertainty method 3
(3) See Appendix 3 for details of uncertainty method 4
(4) See Appendix 4 for details of uncertainty method 5
(5) See Appendix 5 for details of uncertainty method 6
(6) See Appendix 6 for details of uncertainty method 7

level (1, 2, 3, 3' and 4) is the weighted summation of the standardized map scores using Equation 3.1.

$$A_i = \sum_j w_j x_{ij} \tag{3.1}$$

In this equation, x_{ij} is the score of the i^{th} alternative with respect to the j^{th} attribute (criterion), and the weight w_j is a normalized weight, so that $\sum_j w_j = 1$ [28].

Similar to this, in the uncertainty approach the weighted summation turned out to be of the linear utility function [33], where the scores are replaced by utilities [28] (Equation 3.2).

$$U_i = \sum_j w_j u_{ij} \tag{3.2}$$

In the deterministic models, the combining methods also involved some other operations, including conditional map algebra or cell statistics, whereby the related land-use layers or urban land-use layers were used as constraint maps. In these operations, the existence of urban uses were given the minimum suitability at the final level (for the result of level 3 for goals 1, 3, 4, 5 and level 3' for goal 2), or the agriculture-related land uses were given maximum suitability or maximum cell statistic at each level at which they were utilized (at level 2 for goal 1, for the result of level 2 and level 3' for goal 2, for the result of level 2 for goal 3). In the uncertainty approach, the only additional method used after weighting was a transformation using Equation A.1.2 in Appendix 1 to obtain the final so111 map. In this approach, the constraint mapping for existing urban and suburban land uses was made only once on the final preference map.

In the deterministic modeling, all of the priority weights were obtained from the AHP method carried out with the community and experts of a similar county, with only two exceptions. These included the use of information obtained from the annual agricultural sales of Hillsborough County in determining the weights for each row crop type at level 1 to obtain the so111 at level 2, and the weights for five different goals at level 4 to obtain the preference map at level 5.

After the row crops weighting at level 1, the objectives weighting at level 3 under goal 1, and after goals weighting at level 4, the deterministic approach used Equation 3.3 to transform the suitability scores to a range of 1 to 9. The comparison of the final preference map with the one obtained from the uncertainty approach was made on the final untransformed map.

$$(X'_{ij}) = \frac{(X_{ij} - X_{j\,\text{old}}^{\,\text{min}})(X_{j\,\text{new}}^{\,\text{max}} - X_{j\,\text{new}}^{\,\text{min}})}{(X_{j\,\text{old}}^{\,\text{max}} - X_{j\,\text{old}}^{\,\text{min}})} + 1 \tag{3.3}$$

In Equation 3.3, X'_{ij} is the transformed standardized score for the i^{th} alternative of the j^{th} attribute (criterion), X_{ij} is the raw standardized score, and $X_{j\,\text{old}}^{\,\text{min}}$ and $X_{j\,\text{new}}^{\,\text{min}}$ and $X_{j\,\text{old}}^{\,\text{max}}$ and $X_{j\,\text{new}}^{\,\text{max}}$ are the minimum and maximum scores for the j^{th} attribute before and after transformation, respectively.

In the uncertainty approach, to assess the decision maker's (here, the modeler) preference uncertainty on the priority weights at levels 1, 2, 3 and 3' for all goals, with the aim of maximizing agricultural suitability, a direct weighting estimation method – a trade-off – was utilized with consistency checks [33].

For the weighting of the goals themselves (at level 4), a mixed methodology was used to assign weights based on their size in terms of acreage, just (market) value and annual sales. This raised a question of how to weight these weights for different criteria. For this purpose, and to assess the uncertainty in this process, the concept of entropy was utilized by applying a series of formulations to the decision matrix (see [24:52-56]), consisting of goals versus their weightings, based on the three different data sets.

3.3. Comparing the final stages in the two modelings to obtain the agricultural preference maps, and the comparison of these two maps (level 5). The deterministic and uncertainty approaches resulted in their own agricultural preference maps after a weighted sum operation (Equations 3.1 and 3.2) on the goal maps. These maps were finalized by merging them with the constraint map data relating to existing urban-suburban land uses, which were assigned values of 1 and 0 – the minimum standardized scores – in the deterministic and uncertainty models, respectively. However, the comparison of preference maps also involved the exclusion of these areas from their final forms, since they covered the areas that were given and were not a result of the either of the modelings.

4. Results and discussion

In the following sections the results obtained from the deterministic and uncertainty modeling approaches are presented in the order in which they were compared in terms of their methodology and application, as explained in Section 3.

4.1. Comparison of results of the criteria standardization obtained from the two modeling approaches (levels 1, 2 and 3).

Method 1: The results of the maximization of soil suitability for each crop type from the deterministic and uncertainty approaches were found to be different in terms of the level of suitability assigned to areas of similar shapes, in that the latter approach (in this case, the probabilistic one) considered not only yield values, but also their occurrences in space. Since the frequencies have a much greater influence in the multiplication than yield values, the larger areas assumed higher utilities for soil suitability, even though they had lower yield values. When considering long-term land-use planning, this result can be seen as a positive impact on the preservation of large row crop areas, despite their low yields.

Method 2: The uncertainty method (in this case, the probabilistic one) adopted in these two level-1 criterion maps, required subjective evaluations of the decision makers (here, the modeler) by means of utility functions that result from the indifference technique (see Appendix 1). This and the other differences in data processing (such as kriging and additional processes on its results as opposed to IDW in the deterministic approach) yielded highly different results in terms of patterns and the levels of suitability for the cities' population map. In contrast, the deterministic model's linear value assignment for the Euclidian distance map and the non-linear utility function's utility assignment in the uncertainty approach produced rather similar results in terms of the relative placement of higher values to alternatives closer to row crop areas (for an illustrative comparison of the results of the two approaches having different and similar patterns and/or suitability scores, refer to Figure 4 in Section 4.2).

Method 3: The resultant maps from the two approaches were found to be similar in terms of patterns, although the levels of suitability that they reflected were found to be different to the extent that their raw data value ranges were either different (as in so213, so233, so313 and so513) or as a natural result of nonlinear fuzzy membership functions (as in so112, so214, so234, so314, so514 and so412) (see Appendix 2). The remaining group of sub-objective criterion maps (so411, so511 and so515) displayed similarities both in terms of their patterns, and in their level of suitability, as an essential result of two discrete groupings of the same selections from the raw data.

Method 4: The resultant maps from the two approaches were found to be different, which resulted from the uncertainty approach's assessment of major roads or local markets proximities, based on the two-group categorization of the study area (see Appendix 3). In addition to the variations between different levels for the land-use suitability groups, the results showed also internal variations within each group. For each related distance map, the uncertainty approach provided different series of suitability levels for each different parcel of the highly suitable land uses, and for the areas having lower land-use suitability based on a constant mean and standard deviation. On the other hand, the deterministic criterion maps were distinguished by small quarter standard deviation increments around the most suitable area distance buffer, as defined by the mean zonal distance of the existing/ most suitable land for each respective agricultural goal, i.e., goal 1 (row crops), goal 4 (nursery) and goal 5 (timber).

Method 5: When the results of the two approaches were compared, a more significant variety was observed on the 0-1 uncertainty maps than on the 1-9 deterministic maps. The probabilities computed in the former maps allowed the assignment of utilities for land-use suitability with respect to their occurrences in space for the three different agricultural activities (high- and low-intensity livestock, and specialty farming) that were evaluated (see Appendix 4). The transformation of smaller probabilities to higher utilities for high-intensity livestock (so211) and specialty farming (so311), and of higher probabilities to higher utilities for low-intensity livestock (so231), were based on the increasing and decreasing revenues per unit area for the respective agricultural activities.

Method 6: The evaluations of the results of the two approaches were found to be quite similar to those made in the Section Method 4, although differences existed in the higher level of variety in the uncertainty maps. This was due to the grouping of the respective land-use maps into three rather than two, in which the study area was divided into areas of high-, moderate- and low-level land-use potential. Another difference was found in the interfering fuzzy ranking in models so212, so232 and so312, which were set in such a way that the nearer and then the nearest areas to the water resources were left to be given the least suitable ranking in each of the three groups of land-use potentials, i.e., high- and low-intensity livestock and specialty farming (see Appendix 5). Finally, in contrast to the only proximity maximization problems handled in the Method 4 models, both approaches dealt with both the aims of maximization of proximity and distance (i.e., minimization of proximity) on the Euclidian distance maps.

Method 7: In comparing the results from the two approaches, although at first look, the non-spatial component of the resultant corresponding objective criterion maps from the uncertainty approach seems to resemble the deterministic maps, the final uncertainty maps were found to have different patterns and levels of suitability. This was due to a variety of factors, including (1) the existence of their spatial components, (2) the overall fuzzy hedge ordering in each of the components after the enumeration process carried out for both types, and (3) the respective fuzzy set membership values (see Appendix 6).

4.2. Comparison of results for criteria aggregation after weighting from the two modeling approaches (levels 1, 2, 3, 3' and 4). The results of the two approaches after any weighting process at levels of 1, 2, 3, 3' and 4 turned out to be different from each other, to the extent that their component maps are different. The level of differences with respect to the same alternatives (pixels) between the two groups of results at the same level can be categorized into four groups, such that they have either:

1. very different patterns/shapes/forms and different levels of suitability;
2. partially different patterns/shapes/forms and different levels of suitability;
3. similar patterns/shapes/forms and different levels of suitability; or
4. similar patterns/shapes/forms and similar levels of suitability.

Each of the above-listed groups of aggregated weighted map result differences are illustrated by some of the level 2 and level 3 results in Figure 4's 1a-4a (deterministic) vs. 1b-4b (uncertainty) sections.

4.3. Comparison and interpretation of agricultural preference maps from the two modeling approaches (level 5). For a comparison of the preference maps (overall goal) obtained from the two modeling approaches at level 5 of the hierarchical structure of LUCIS, the z-scores of each pair of maps (including and excluding the existing urban-suburban areas) and the z-score differences were computed. The maps, their distributions and the summary statistics of these comparisons are shown in Figure 5.

When the first case was evaluated in terms of its z-scores, the deterministic result was found to vary between -1.153 and 1.783, and the uncertainty between -1.157 and 1.978 (Figure 5). However, when the given urban-suburban areas were excluded from

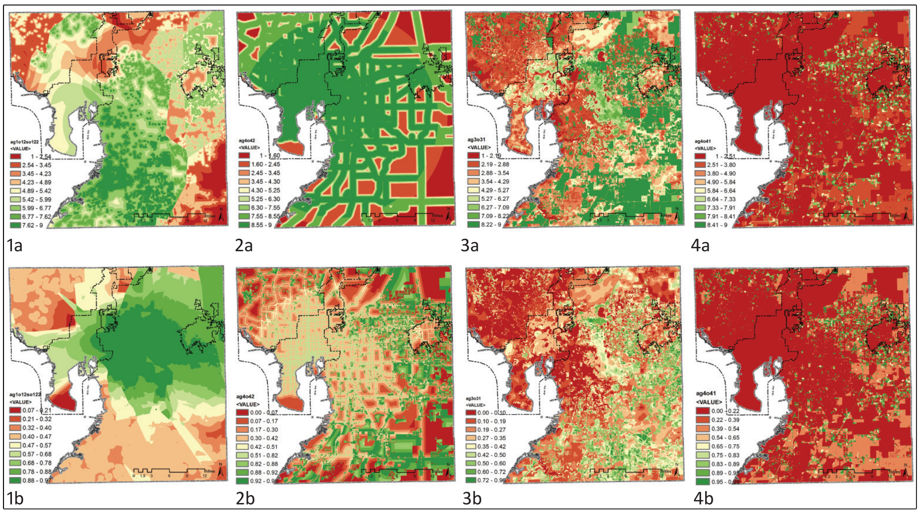


Figure 4. Deterministic (1a, 2a, 3a, 4a) and uncertainty (1b, 2b, 3b, 4b) aggregated maps for sub-objective 122 (1a and 1b), objective 42 (2a and 2b), objective 31 (3a and 3b), and objective 41 (4a and 4b)

the analysis, which composed the modes (i.e., the most repeated land-use type) for both distributions (see the 1^{st} and 3^{rd} graph in the 2^{nd} row of Figure 5), the minimum values of the maps increased to -0.825 and -0.946, respectively. In the second case, the graph of the deterministic result revealed a bi-modal distribution, with one near to its mean (0.742), and the other towards the end of its lower tail (at about 1.53). Accordingly, it suggested a data spread that cannot be attributed to a normal distribution (see the 2^{nd} graph in the 2^{nd} row of Figure 5); however, looking at the graph of the uncertainty result (see the 4^{th} graph in the 2^{nd} row of Figure 5), it is seen that it was more or less normally distributed about its own mean (0.744). The main difference between the two results was observed in the uncertainty result filling the gap between the two modes of the deterministic approach. This comparison can be illustrated by overlaying the two graphs after converting them to the same scale, after which the difference can be seen in the light grey tone frequency distribution in the 2^{nd} graph on the bottom row of Figure 5. It can also be seen in this graph that following the exclusion of unsuitable areas from the analysis, a substantial part of all alternatives (pixels) in both results is observed on the positive side of the z-score distribution.

The z-score difference maps for the two cases (i.e., including and excluding urban and sub-urban areas) was found to vary between a minimum of -1.629 and a maximum of 1.263, suggesting a non-statistically significant difference between the two results in a one-to-one comparison of each pixel (alternative) at a 95 percent confidence interval (see the summary statistics in the 3^{rd} and 4^{th} rows of Figure 5). Moreover, in the second case, when the given urban-suburban areas were excluded, as would be expected, the distribution of the difference map was found to be approximating a normal distribution around a mean value, which was very close to zero (-0.0088) (see the 1^{st} graph and the summary statistics on the bottom row of Figure 5). Accordingly, based on the comparison of agricultural preference maps in terms of their z-score pixel values, it can be stated that the newly developed uncertainty models did not result in a significant difference over the existing deterministic models of LUCIS.

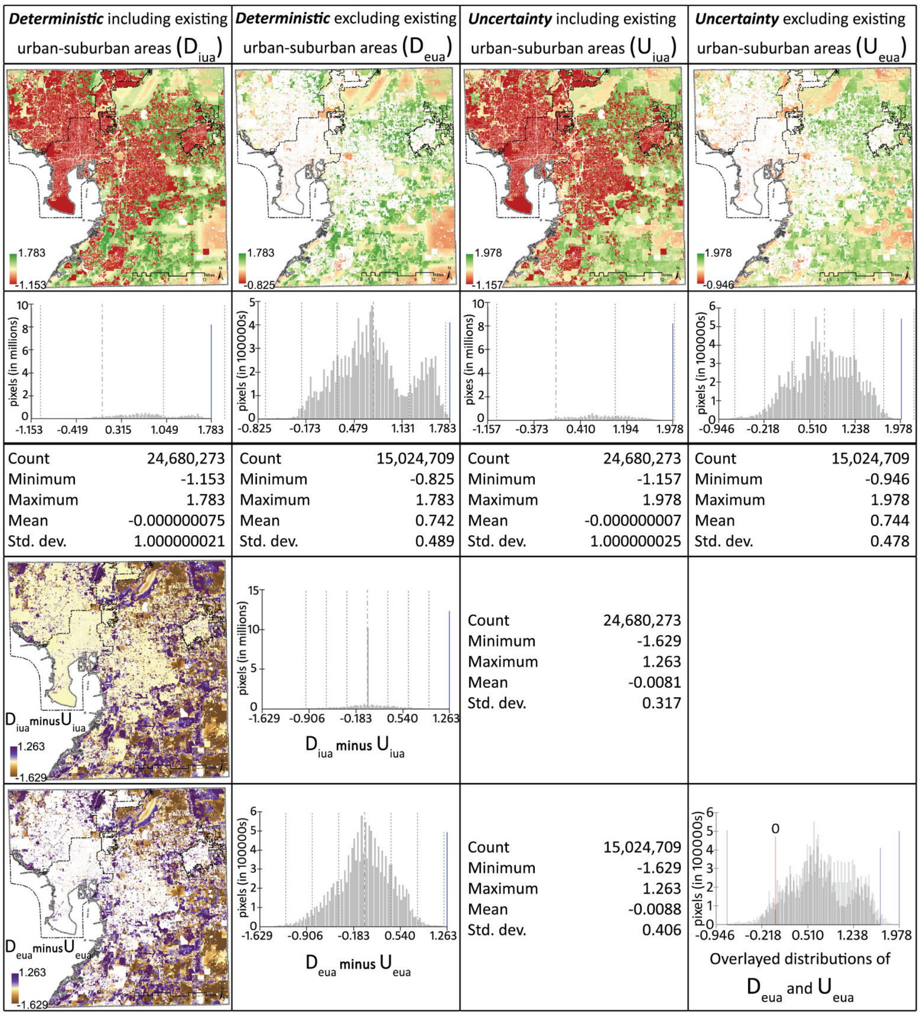


Figure 5. Z-score agricultural preference maps of deterministic and uncertainty approaches and their z-score difference maps, including and excluding the existing urban-suburban areas, distributions and summary statistics of maps

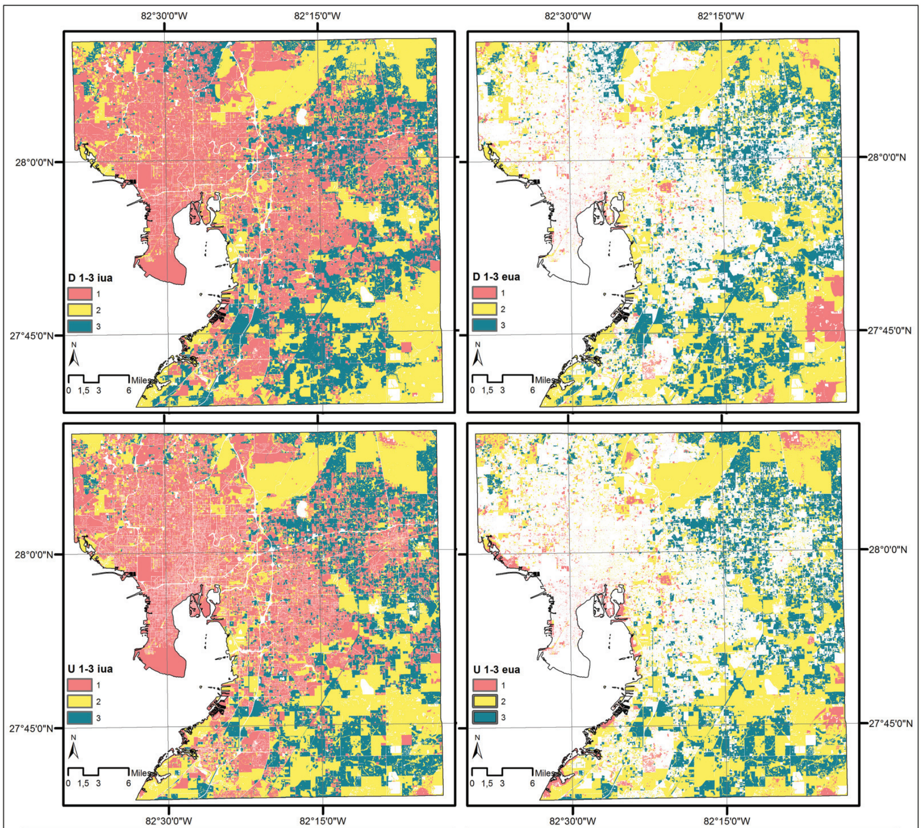
On the other hand, as stated earlier, by means of three different land-use preference maps (agricultural, urban and ecologically sensitive), the ultimate aim in LUCIS modeling is to achieve a land-use conflict map (Figure 2), and based on this, to develop possible future land-use scenarios. The first step in the conflict analysis requires the three preference maps to be collapsed into three classes, in which each map is differentiated by low, moderate and high levels of preferences [9]. Therefore, to be evaluated as a base map in the conflict analysis, the two agricultural preference maps from deterministic and uncertainty modeling were also compared after being collapsed into three equal interval rank groups, labeled 1 for low, 2 for moderate and 3 for high preferences, i.e., their agricultural land-use suitability. The results of these analyses for both cases, i.e., including and excluding the given urban-suburban areas, are shown in Figure 6.

In the first case, as expected, the total number of alternatives (pixels) on the two preference maps was found to have a correspondence level as high as 83.72 percent, about 40 percent of which was a result of the same given urban-suburban areas having the same preference level of 1 (see the table on the left in Figure 6). Accordingly, the Cohen's Kappa, which is a measure of agreement between the two ordered preference groups [34] of the two maps, was found to be 0.75 (Figure 6). Since the used data was the population itself, its significance was not assessed. Nevertheless, the results for the second case suggested a higher level of difference between the two maps. When the existing urban-suburban areas were excluded, the total difference in one level of preference from 1 to 2 or 2 to 1, and from 2 to 3 or 3 to 2, increased by almost two times, i.e., from 16.28 percent to 32.04 percent (see the two tables in Figure 6). In addition, although negligible, the difference in two levels of preference from 1 to 3 or 3 to 1 increased to 0.18 percent from 0.00058 percent, which was the result of only one category of the collapsed map having a value of 1 in the deterministic and 3 in the uncertainty components. That is, in the second case, the collapsed map had a newly emerged category for two levels of preference difference with a value of 3 from the deterministic component and 1 from the uncertainty map. As a result of the second case analysis, as expected, the Cohen's Kappa value decreased to 0.39 (Figure 6), which suggested only a moderate level of agreement between the two ordered preference groups of the two maps rather than a strong one [34].

5. Conclusion

Recognizing the need for studies relating to the proper expression of uncertainty in GIS-based multi-criteria in land-use planning, this study has concentrated on epistemic uncertainties, concerning particularly the last three stages of the spatial multi-criteria modeling process commonly defined in spatial MCDM literature, being decision making on criteria standardization, criteria weighting and the interpretation of the final results. In general, the uncertainty associated with criteria standardization and weighting processes is assessed by way of classical error propagation or sensitivity analyses, which measure the impact of the errors found in, or perturbations made to the criterion values and their weights on the outputs in terms of the suitability ranking of alternatives. Instead of utilizing these indirect methods of uncertainty assessment at the final output level in decision making [28], this study set out with the main premise that *uncertainty makes a difference* in terms of both the pattern and level of suitability of the alternatives at each hierarchical level of multi-criteria land-use planning. In doing this, no consideration was given to how "objective" or "sensitive" the decisions were, and by whom they were taken in the decision-making process, whether individual modelers, a group of experts with different backgrounds – such as planners [42] –, community participants [9; 17] and/or politicians.

To this end, the study tried to show the importance of determining the risk in choosing a particular alternative [11] in land-use planning, and for this purpose it made use of LUCIS (Land-Use Conflict Identification Strategy), which is a deterministic GIS-based multi-criteria decision process, and compared it with a newly developed equivalent uncertainty modeling at each level of the hierarchical structure. Although the ultimate aim of LUCIS is to represent the probable patterns of future land use based on a conflict map obtained from the overlaying of low, moderate and high levels of preferences or suitability



Determinstic 1-2-3 map including existing urban-suburban areas ($D_{1\text{-}3\ iua}$)	Count 24,680,273 Mean 1.852 Std. dev. 0.795	***Deterministic*** 1-2-3 map excluding existing urban-suburban areas ($D_{1\text{-}3\ eua}$)	Count 15,024,709 Mean 2.256 Std. dev. 0.582
Uncertainty 1-2-3 map including existing urban-suburban areas ($U_{1\text{-}3\ iua}$)	Count 24,680,273 Mean 1.816 Std. dev. 0.778	***Uncertainty*** 1-2-3 map excluding existing urban-suburban areas ($U_{1\text{-}3\ eua}$)	Count 15,024,709 Mean 2.258 Std. dev. 0.574

	$U_{1\text{-}3\ iua}$ (1)	$U_{1\text{-}3\ iua}$ (2)	$U_{1\text{-}3\ iua}$ (3)	**Total**
$D_{1\text{-}3\ iua}$ (1)	9844203 39.89%	47095 0.19%	142 0.00058%	**9891440** **40.08%**
$D_{1\text{-}3\ iua}$ (2)	299110 1.21%	6729315 27.27%	1522482 6.17%	**8550907** **34.65%**
$D_{1\text{-}3\ iua}$ (3)	0 0.00%	2150388 8.71%	4087538 16.56%	**6237926** **25.27%**
Total	**10143313** **41.10%**	**8926798** **36.17%**	**5610162** **22.73%**	**24680273** **100.00%**

Cohen's Kappa = 0.75

	$U_{1\text{-}3\ eua}$ (1)	$U_{1\text{-}3\ eua}$ (2)	$U_{1\text{-}3\ eua}$ (3)	**Total**
$D_{1\text{-}3\ eua}$ (1)	556759 3.71%	528431 3.52%	25158 0.17%	**1110348** **7.39%**
$D_{1\text{-}3\ eua}$ (2)	479183 3.19%	6604299 43.96%	1869867 12.45%	**8953349** **59.59%**
$D_{1\text{-}3\ eua}$ (3)	887 0.01%	1935323 12.88%	3024802 20.13%	**4961012** **33.02%**
Total	**1036829** **6.90%**	**9068053** **60.35%**	**4919827** **32.74%**	**15024709** **100.00%**

Cohen's Kappa = 0.39

Figure 6. Agricultural preference maps and distributions of three-class equal interval z-score agricultural preference maps, including and excluding the existing urban and suburban areas, their cross-tabulations and Cohen's Kappa values

for the three land uses (agricultural, urban, and ecologically sensitive) at the 6^{th} level of the hierarchy, the scope of this study was limited to the agricultural preference (overall goal) map at the 5^{th} level, starting from the maps in the 1^{st} level, corresponding to lower level sub-objectives.

The two modeling approaches were applied to the case of Hillsborough County in the state of Florida, which is characterized by heavy urbanization and an urban footprint [44] that continues to expand into the valuable natural and agricultural areas. The comparison of the methodologies and results of the two modeling were made in three stages of the analysis: (1) in criteria standardization, prior to the application of any weighting at levels 1, 2 and 3; (2) in criteria weighting and aggregation at levels 1, 2, 3, 3' and 4; and (3) in obtaining the preference maps and the interpretation of these maps at level 5.

The first stage at which uncertainty is assessed by means of probability, fuzzy sets and multi-attribute utility theories under seven different groupings of the unweighted criterion maps of the model revealed:

- different suitability levels and more variability in the alternatives for similar physical boundaries (method 1 and method 5, respectively);
- different suitability levels with similar patterns (part of method 3);
- different suitability levels with different patterns (part of method 2, method 4, method 7) with more variability (method 6); and
- similar suitability levels with similar patterns (part of method 2, part of method 3).

Similarly, the comparisons of the maps at the aggregation levels after the weighting which were handled by Analytic Hierarchy Process in the deterministic modeling and using the trade-off method, except for the weights of goal maps in the uncertainty modeling, were found to have differentiating levels of differences in terms of pattern/shape/form and the degree of land-use suitability.

In the final stage of the analysis, which addresses directly the agricultural land-use preferences in the decision-making process, a moderate level of difference was identified between the two approaches when the given urban-suburban areas are excluded from the analysis and when the agricultural preference map is collapsed into three different levels of preference (low, moderate and high), which is a critical, and in fact an uncertain, process in defining and interpreting the results of modeling. This process needs special attention when the preference maps results are not utilized on the basis of individual alternatives (pixels), but rather on the basis of data that is collapsed into only a few broad categories. The main difference in these broad categories was reflected in the change between the moderate and high levels of suitability between the two approaches in about 13 percent of the alternatives in either direction, that is from moderate to high and vice versa, and their locations in the southeast and north east parts of the county. The use of different algorithms for modeling uncertainty in decision making in the standardization of criteria values and criteria weighting would have given rise to a different set of solutions in terms of the ranking and spatial pattern of agricultural land-use suitability. This study has aimed to show this possibility, and to clarify that the unique solution set obtained through a deterministic approach should not be considered as the only one, and also that uncertainty assessments are an indispensable part of land-use planning, since they make a difference. This point should be considered when engaged in informed decision and policy making to allocate limited land resources to their most appropriate land uses in future, being aware of the limitations and assumptions of the utilized modeling.

Acknowledgement

This study was carried by Aygün Erdoğan (PhD) between Feb 1, 2012 and Feb 1, 2013 while she was a visiting research scholar in the Dept. of Urban and Regional Planning, College of Design, Construction and Planning, University of Florida. The authors owe great thanks to Bill O'Dell, the Director of the Schimberg Center for Housing Studies in the University of Florida, for providing the office space, hardware and software that allowed them to carry out the study. Appreciation also goes to the two researchers from the Center, Abdulnaser Arafat (PhD) and Liz Thompson, for their valuable comments on the study and for their help in evaluating some of the data, and also to the anonymous reviewers for their constructive comments and suggestions.

Appendices

A.1. Details of uncertainty method 2. The probability raster of the standard deviation of prediction for "cities' population" was obtained by applying a probability distribution (exponential) function observed for its distribution given in Equation A.1.1.

$$\text{(A.1.1)} \quad f(x;\lambda) = \lambda e^{-\lambda x}$$

In this equation, the value of λ, which is a scale parameter, is estimated by calculating the observed mean nearest neighbor distance of the used cities' distribution. The final criterion map was obtained by applying a cumulative exponential distribution with the formula given in Equation A.1.2 on the multiplied raster.

$$\text{(A.1.2)} \quad f(x;\lambda) = 1 - e^{-\lambda x}$$

The value of λ, which is now a rate parameter and is the reciprocal of the scale parameter, found by dividing 1 by a denominator that was assumed to be the mean of the distribution of the weighted cities' population map obtained by Equation A.1.1.

Figure A.1.1 below shows the estimated utility function obtained in the spreadsheet environment applied on the "row crops" Euclidian distance raster.

Outcome	Utility (Probability)
24213	0
15000	0.0625
10000	0.125
7000	0.25
5000	0.5
2500	0.75
500	0.875
0	1

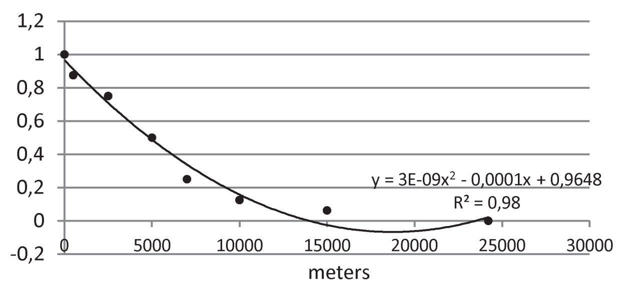


Figure A.1.1. Utility scores and curve estimated through the indifference technique for distance to row crop areas to obtain the row crops distance criterion map

A.2. Details of uncertainty method 3. Fuzzy large or small transformation functions with their default mid-point and spread values were used, whereby the larger and smaller input values are more likely to be a member of the set, respectively [5]. The only two exceptions to the use of mid-point default values were the use of the mean of the distribution of the raster values as for so213, so233, so313 and so513, and the mean of parcels having a size of equal or greater than 10 acres for so412. After the rasterization of this map, the values for Nodata (null) pixels were computed using a conditional map algebra, assigning them a value through the multiplication of the number of cells by 100 to find their area in square meters, which was then converted to acres. Similarly, in the model for the timber parcel size sub-objective (so515), a conditional map algebra was run so that the null pixels had a value of 1 in contrast to other selected and rasterized pixel values of 9 before the fuzzy membership operation. Moreover, in the aquifer recharge models, a conditional map algebra was run in which the null values (originally water surfaces) were set to a membership value of 0. A final additional operation in the row crops land-use model (so112) was a fuzzy OR overlay on the fuzzy membership maps.

A.3. Details of uncertainty method 4. The models followed the course of actions below.

1. Two zonal statistics raster maps were obtained for the regions having a fuzzy set membership value of $0.5 < x \leq 1$ (higher level of suitability from so112, so411 and so511): one for the mean, and the one for the standard deviation, based on the five respective Euclidian distance maps obtained for the deterministic approach.
2. To assign the utilities of distances to major roads and/or local markets for the three goals' land-use values with a suitability level of $0.5 < x \leq 1$, two conditional map algebra were operated respectively on each of these statistical raster maps described above, and on the respective Euclidian distance maps. As a result, two new raster maps were obtained showing the rankings seen in Table A.3.1.

Table A.3.1. Ranks assigned to conditional rasters

Conditional Raster 1	
Euclidian distances having a value lower than mean – 3 standard deviations	8
Euclidian distances having a value lower than mean – 2 standard deviations	7
Euclidian distances having a value lower than mean – 1 standard deviation	6
Euclidian distances having a value lower than mean	5
Otherwise	4
Conditional Raster 2	
Euclidian distances having a value higher than mean + 3 standard deviations	1
Euclidian distances having a value higher than mean + 2 standard deviations	2
Euclidian distances having a value higher than mean + 1 standard deviation	3
Otherwise	4

3. For the areas with a value of 4 in the 2^{nd} conditional raster, the values of the 1^{st} conditional raster were computed (otherwise the values of the 1^{st} were taken) on a new raster that combined the two. The assignments in the first and second conditional raster maps and the combined raster are illustrated on a normal curve in Figure A.3.1.
4. In a similar way, the regions with a fuzzy set membership value of $0 \leq x \leq 0.5$ (lower level of suitability from so112, so411 and so511) were reclassified with the same range of 1-8 by means of the constant rasters. These rasters were created using the mean values of the mean and standard deviation zonal statistics maps of the complementary areas (i.e., where $0.5 < x \leq 1$).

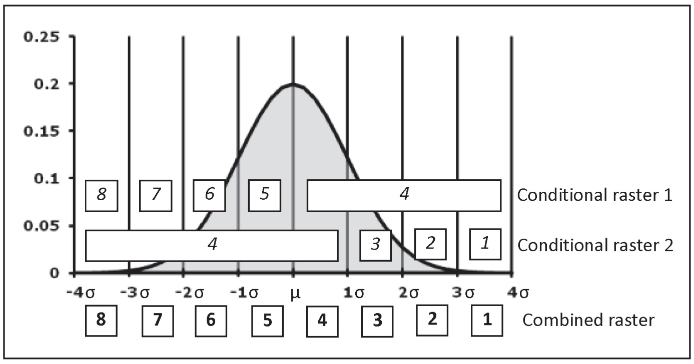


Figure A.3.1. The rank assignments for row crops, nursery and timber land-use suitability levels of $0.5 < x \leq 1$ with respect to their mean and standard deviations found from their distances to major roads and/or local markets

5. The final ranking of the utilities for suitability, which gave higher priority to the more suitable areas, were used in a fuzzy set membership operation after the reclassification of the combined conditional maps for higher-to-lower suitability rank groups, one after the other. That is, from 1-7 to 8-15 (for so123, so423 and so522) or from 1-8 to 9-16 (for so421 and so521), and merging the two resultant combined rasters. This was achieved through a maximum cell statistic operation to obtain raster maps having values of 1-15 or 1-16 in different rankings for the distances. Finally, the resultant maps for the five sub-objectives were obtained through a fuzzy large transformation function with default mid-point and spread values.

A.4. Details of uncertainty method 5. In the uncertainty models, the objects with suitability scores greater than 1 in the deterministic models were selected. Since the high-intensity livestock (so211) and specialty farming (so311) activities were carried out mainly on smaller farmland areas and low-intensity livestock (so231) on larger farmland areas, their probabilities were computed from the area of each selected object divided by the total area of all the selected objects. For the former two sub-objectives (so211 and so311), the probability values with a value of 0 at a 1/1,000,000 precision level were assumed to be slivers, and were thus excluded from any further analysis. This was to prevent them from having higher utilities for suitability based on the subsequent transformation of their probabilities. The probabilities were then recalculated in the same way, with the results processed on a spreadsheet, after which a transformation function of a logarithm of base 0.000001 was carried out, resulting in a maximum utility of 1 for the smallest probabilities for so211 and so311. The results of these functional transformations were merged with the original vector data in the model, and raster maps were created based on these utilities by way of a polygon-to-raster operation, followed by the assignment of 0 to any pixels having null values by a map algebra operation. In the model for low-intensity livestock land use (so231), after obtaining a raster map of the computed probabilities in the first step, the cumulative exponential distribution function given in Equation A.1.2 was applied. The value of the rate parameter of λ was found by dividing 1 by the mean of the probability distribution, which was assumed to be the scale parameter. The model for so231 was then finalized by the assignment of 0 to any pixels having null values.

A.5. Details of uncertainty method 6. Although the uncertainty models utilized the same Euclidian distance maps as a base in the corresponding deterministic models, their results were different due to the uncertainties in these models, which were assessed in a similar way to that explained in Method 4. The difference here was the evaluation of three rather than two mean-standard deviation pairs of zonal statistic raster maps (total of 6 rasters) by means of the three different selections. These were based on the pixels from the resultant land-use suitability maps from so211, so231 and so311 for sub-objective groups of (1) so212, so215, so221 and so223 related to high-intensity livestock activities; (2) so232, so241 and so243 related to low-intensity livestock activities; and (3) so312, so321 and so323 related to specialty farming, respectively. In addition, instead of two groups, the selection of three groups from the raster maps here included the selection of alternatives (x) having utility levels based on the functional transformations of the probabilities found for the respective land-use parcels, which were $0.5 < x \leq 1$; $0 < x \leq 0.5$ and x=0. Another difference was found in the interfering linguistic hedges (as in so212, so232 and so312) for these groups of probabilities (an example is given in Figure A.5.1), rather than their one-after-the-other ordering (as in so215, so221, so241, so321, so223, so243 and so323). Moreover, the constant rasters were created using the mean values of the mean and standard deviation zonal statistics maps of the complementary land-use probability groups having a value of $0 < x \leq 0.5$ for models so212, so312, so215, so221, so321, so223, so323, and by the one having a value of $0.5 < x \leq 1$ for models so232, so241 and so243. Finally, the models ended with a fuzzy small (for so212, so232, so312, so215) or large (for so221, so241, so321, so223, so243, so323) transformation function with default mid-point and spread values.

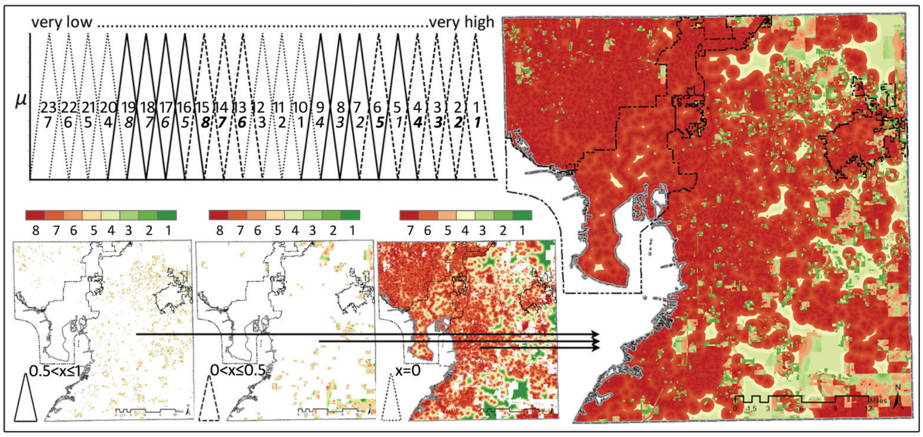


Figure A.5.1. Simplified fuzzy representations and combined rasters, and the resultant fuzzy ranked sub-objective 212 criterion map

A.6. Details of uncertainty method 7. In order to handle any uncertainties in the fuzzy set membership functions based on both the spatial and non-spatial aspects of the alternatives, i.e., parcels, the uncertainty models involved the following steps:

1. To obtain the spatial component of the model, the parcel objects having the land uses that were selected in the deterministic model were selected based on the same sliver assumption criteria for each respective objective.
2. The X-Y coordinates of the selected data centroids were computed in the GIS, and the data was inputted into the spatial data analysis software in order to run a K-means clustering routine, for which the separation parameter was set as 5. Since the main clustering regions were observed to be 3 for objectives 13, 26, 33 and 43, and 2 for objectives 25 and 53, the K-location values were set as 3 and 2 for the respective objectives.
3. The first and second standard deviation ellipses of the computed three or two respective K-means clusters were visualized in the GIS, and their parameters were used to compute the third standard deviation ellipses through a table to ellipse operation.
4. The model continued with dissolving, erasing and merging operations (and geometry repairment operations when needed to remove slivers etc.) to obtain combined concentrated zones of three standard deviation ellipses with no self-intersecting areas. Subsequently, these areas were rasterized and reclassified with respect to their standard deviation ellipse numbers and as Nodata around the third ellipses to be combined with the non-spatial component of the model.
5. As for the non-spatial component, the parcel objects having descriptions other than 'header' and 'note' were selected, and a raster layer was obtained from these objects based on just value per acre field. This raster was reclassified with the listed ranks below for the 3 K-means cluster objectives of 13, 26, 33 and 43, and without rank 4 for the 2 K-means cluster objectives of 25 and 53.
   - 0 and mean $(\overline{x})$ as rank 1;
   - $(\overline{x})$ and $(\overline{x})+$ one standard deviation (s) as rank 2;
   - $(\overline{x}) + (s)$ and $(\overline{x}) + 2(s)$ as rank 3;
   - $(\overline{x}) + 2(s)$ and a value that is more than the largest just value/acre value in the data set as rank 4;
   - Nodata as Nodata
6. Two separate map algebra tools were used in an enumeration process of the two classified standard deviation raster maps, with one based on the spatial (cluster location) properties of the selected parcels, and the other on the just value/acre values of all the parcels with respect to the mean and the standard deviation statistics of the selected parcels. The enumeration processes involved the multiplication of the first component by 10, then adding the second component to the result. These processes, their respective maps, the assignment of ranking to the enumeration results and the combined enumeration map through a mean cell statistic (max or min operations would also have given the same result) operation can be seen in Figure A.6.1 with the example for objective 13. While all other 3 K-means cluster objectives (o26, o33, o43) utilized a similar 4x4 enumeration and a 10-level ranking, 2 K-means cluster objectives (o25 and o53) used a 3x3 enumeration tables removing the 4^{th} column and 4^{th} row from the 4x4 one, and a 6-level ranking, which replaced 7 with 6 in the 4x4 table (Figure A.6.1).
7. Before carrying out the final fuzzy membership operation to obtain the final objective criterion maps, the parcel objects having descriptions of 'header' and 'note' were selected and assigned a value of 11 for objectives 13, 26, 33 and 43, and a value of 7 for objectives 25 and 53, and then merged with the final raster map obtained in the previous step already having a 1-10 or 1-6 ranking, respectively.

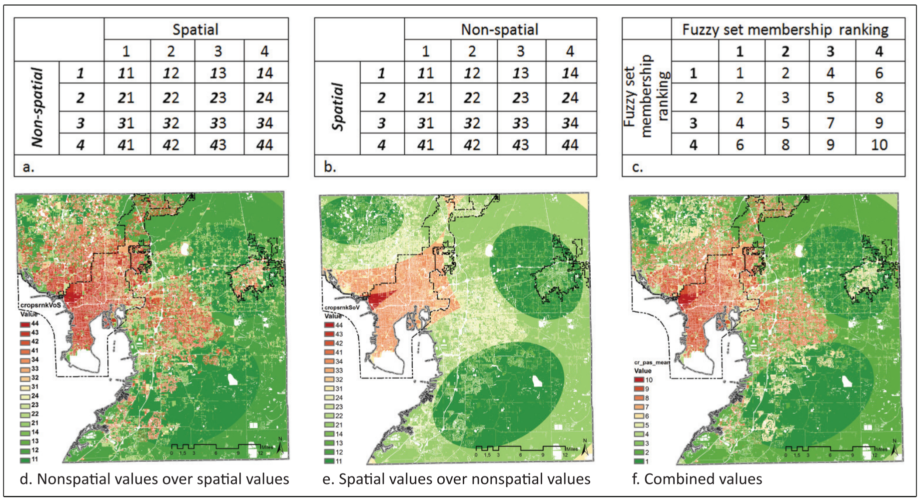


Figure A.6.1. The enumeration processes (a and b), rank assignment to the enumeration results (c) and the respective maps (d for a and e for b), including their combined cell statistic resultant map (f) for a further fuzzy set membership operation in the model for objective 13

References

[1] Adhikari, B. and Li, J. *Modelling ambiguity in urban planning*. Annals of GIS **19** (3): 143-152, 2013.

[2] Agumya, A. and Hunter, G.J. *Responding to the consequences of uncertainty in geographical data*. Int. J. Geographical Information Science **16** (5): 405-417, 2002.

[3] Ahmad, T., Rai, A. and Singh, R. *Objective spatial analytic hierarchy process for identification of potential agroforestry areas*. Model Assisted Statistics and Applications **7** (1): 65-73, 2012.

[4] Altman, D. *Fuzzy set theoretic approaches for handling imprecision in spatial analysis*. International Journal of Geographical Information Systems **8** (3): 271-289, 1994.

[5] ArcGIS Resources, *How fuzzy membership works*, http://resources.arcgis.com/en/help/main/10.1/index.html#//009z000000rz000000, accessed on 1 July 2012.

[6] Avdagic, Z., Karabegovic, A. and Ponjavic, M. *Fuzzy logic and genetic algorithm application for multi criteria land valorization in spatial planning*. in Artificial Intelligence Techniques for Computer Graphics, Studies in Computational Intelligence 159, eds. D. Plemenos, G. Miaoulis, (Springer-Verlag Berlin Heidelberg, 2009), 175-198.

[7] Baja, S., Chapman, D.M. and Dragovich, D. *Spatial based compromise programming for multiple criteria decision making in land use planning*. Environmental Modeling and Assessment **12** (3): 171-184, 2007.

[8] Bonham-Carter, G.F. *Geographic information systems for geoscientists: modelling with GIS*, Computer Methods in the Geosciences 13, (Pergamon, Ontario, 1996).

[9] Carr, M.H. and Zwick, P.D. *Smart land-use analysis: the LUCIS model* (ESRI Press, California, 2007).

[10] Castrignano‘, A., Buondonno, A., Odierna, P., Fiorentino, C. and Coppola, E. *Uncertainty assessment of a soil quality index using geostatistics*. nvironmetrics **20** (3): 298-311, 2009.

[11] Chen, H., Wood, M.D., Linstead, C., Maltby, E. *Uncertainty analysis in a GIS-based multi-criteria analysis tool for river catchment management.* Environmental Modelling & Software **26** (4), 395-405, 2011.

[12] Chen, Y., Yu, J., Khan, S. *Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation.* Environmental Modelling & Software **25** (12): 1582-1591, 2010.

[13] Chen, Y., Yu, J., Khan, S. *The spatial framework for weight sensitivity analysis in AHP-based multi-criteria decision making.* Environmental Modelling & Software **48** (10): 129-140, 2013.

[14] Couclelis, H. *The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge.* Transactions in GIS **7** (2): 165-175, 2003.

[15] Çelik, H.M. and Türk, E. *Determination of optimum environmental conservation: Using multi-criteria decision-making techniques.* European Planning Studies **19** (3): 479-499, 2011.

[16] Delgado, M.G. and Sendra, J.B. *Sensitivity Analysis in Multicriteria Spatial Decision-Making: A Review.* Human and Ecological Risk Assessment **10** (6): 1173-1187, 2004.

[17] Erdoğan, A. *Modelling of expert knowledge in Geographic Information Systems-based planning of the Tuz Lake special environmental protection area, Turkey.* Planning Practice & Research **24** (4): 435-454, 2009.

[18] Falk, M.G., Denham, R.J. and Mengersen, K.L. *Estimating Uncertainty in the Revised Universal Soil Loss Equation via Bayesian Melding.* Journal of Agricultural, Biological, and Environmental Statistics **15** (1): 20-37, 2009.

[19] FGDL, Florida Geographic Data Library, *Metadata explorer: Search and download data,* http://www.fgdl.org/metadataexplorer/explorer.jsp, accessed on 1 March 2012.

[20] Fishburn, P. C. *Utility theory for decision making* (John Wiley & Sons, New York, 1970).

[21] FNAI, Florida Natural Areas Inventory, *Critical Lands and Waters Identification Project (CLIP): Version 2.0 Technical Report-January 2012* by J. Oetting, T. Hoctor and B. Stys, http://www.fnai.org/pdf/CLIP_2_Tech_Report.pdf, accessed on 15 February 2012.

[22] French, S. *Decision theory: an introduction to the mathematics of rationality* (Ellis Horwood Limited, Chichester, West Sussex, England, 1986).

[23] HCG, Hillsborough County Government, *About the County,* http://www.hillsboroughcounty.org/index.aspx?nid=423, accessed on 1 March 2013.

[24] Hwang, C-L., and Yoon, K. *Multiple attribute decision making: methods and applications-a state-of-the-art survey* (Springer-Verlag, Berlin & New York, 1981), pp. 52-56.

[25] Kosko, B. *Fuzziness vs. Probability.* International Journal of General Systems, **17** (2-3): 211-240, 1990.

[26] Laskey, K.B., Wright, E.J., and da Costa, P.C.G., *Envisioning uncertainty in geospatial information.* International Journal of Approximate Reasoning, **51** (2): 209-223, 2010.

[27] Ligmann-Zielinska, A. and Jankowski, P. *Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation.* Environmental Modelling & Software **57** (7): 235-247, 2014.

[28] Malczewski, J. *GIS and multicriteria decision analysis* (John Wiley & Sons, New York, 1999).

[29] Malczewski, J. *GIS-based land-use suitability analysis: a critical overview.* Progress in Planning **62** (1): 3-65, 2004.

[30] Mosadeghi, R., Warnken, J., Tomlinson, R. and Mirfenderesk, H. *Uncertainty analysis in the application of multi-criteria decision-making methods in Australian strategic environmental decisions.* Journal of Environmental Planning and Management, **56** (8): 1097-1124, 2013a.

[31] Mosadeghi, R., Warnken, J., Mirfenderesk, H. and Tomlinson, R. *Spatial uncertainty analysis in coastal land use planning: a case study at Gold Coast, Australia.* Journal of Coastal Research, Special Issue 65: 1003-1008, 2013b.

[32] Murray, A.T., Wei, R. and Grubesic, T.H. *An approach for examining alternatives attributable to locational uncertainty.* Environment and Planning B: Planning and Design **41** (1): 93-109, 2014.

[33] Nijkamp, P., Rietveld, P., and Voogd, H. *Multicriteria evaluation in physical planning* (Elsevier Science Publishers, Amsterdam, 1990)

[34] Norušis, M. J. *PASW statistics 18 guide to data analysis* (Prentice Hall Ptr., Upper Saddle River, NJ, 2010).

[35] O'Brien, R., Cook, S., Peters, M. and Corner, R. A *Bayesian modeling approach to site suitability under conditions of uncertainty*, in: Proceedings of the 7^{th} International Conference on Precision Agriculture and Other Precision Resources Management (25-28 July 2004, Minneapolis, USA), ed. D.J. Mulla, (St. Paul, USA, 2004), 615-626.

[36] Refsgaard, J.C., van der Sluijs, J.P., Hojberg, A.L. and Vanrolleghem, P. *Uncertainty in the environmental modelling process-A framework and guidance.* Environmental Modelling & Software, **22** (11): 1543-1556, 2007.

[37] Reshmidevi, T.V., Eldho, T.I. and Jana, R. A *GIS-integrated fuzzy rule-based inference system for land suitability evaluation in agricultural watersheds.* Agricultural Systems, **101** (1-2): 101-109, 2009.

[38] Reynolds, J.E. *Urbanization and land use change in Florida and the south,* in: Proceedings of a Regional Workshop on Current Issues Associated with Land Values and Land Use Planning (June 2001), ed. J.C. Bergstrom (Southern Rural Development Center/Farm Foundation, 2001), 28-49.

[39] Scholz, R.W. and Lu, Y. *Uncertainty in Geographic Data on Bivariate Maps: An Examination of Visualization Preference and Decision Making.* ISPRS International Journal of Geo-Information, **3** (4): 1180-1197, 2014.

[40] Soltani, S.R., Mahiny, A.S., Monavari, S.M. and Alesheikh, A.A. *Sustainability through uncertainty management in urban land suitability assessment.* Environmental Engineering Science, **30** (4): 170-178, 2013.

[41] Store, R. and Kangas, J. *Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling.* Landscape and Urban Planning, **55** (2): 79-93, 2001.

[42] Türk, E. and Çelik, H.M. Impacts of planners' different viewpoints on optimum land-use allocation. *European Planning Studies,* **21** (12): 1937-1957, 2013.

[43] 2000 and 2010 Census data, http://www.census.gov/, accessed on 1 December 2012.

[44] van Beynen, P., Feliciano, N., North, L. and Townsend, K. *Application of a karst disturbance index in Hillsborough County, Florida.* Environmental Management, **39** (2): 261-277, 2007.

[45] Walker, W.E., Harremoës, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P. and Krayer von Krauss, M.P. *Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support.* Integrated assessment, **4** (1): 5-17, 2003.

[46] Warmink, J.J., Janssen, J.A.E.B., Booij, M.J. and Krol, M.S. *Identification and classification of uncertainties in the application of environmental models.* Environmental modelling & Software, **25** (12): 1518-1527, 2010.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1275 – 1296

Single server queueing model with Gumbel distribution using Bayesian approach

A. Jabarali* † and K. Senthamarai Kannan ‡

Abstract

Bayesian methodology is an important technique in statistics, and especially in mathematical statistics. It consists of the sample information along with the prior information available about the parameter before the sample has been observed. This paper exhibits the estimation of the parameters of queueing model with inter-arrival time and service time which follows Gumbel distribution. Bayesian procedure is applied to obtain the estimation of the model parameters and the traffic intensity of queueing model based on the informative and the non-informative prior knowledges. In this paper, the Bayesian estimates are carried out by numerically and graphically with the help of Markov Chain Monte Carlo (MCMC) simulation technique, particularly in Gibbs sampling algorithm.

Keywords: Queue, Gumbel distribution, Bayesian estimation, Gibbs sampling, Markov Chain Monte Carlo technique.

2000 AMS Classification: 60K25, 62C10, 62C12, 11K45

Received : 18.01.2015 *Accepted* : 02.07.2015 *Doi* : 10.15672/HJMS.20157512015

*Department of Statistics, Madras Christian College, Chennai-600 059, Tamil Nadu, India. Email: statjafar@gmail.com

†Corresponding Author.

‡Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli - 627 012, Tamil Nadu, India. Email: senkannan2002@gmail.com

1. Introduction

Statistical inference in queueing theory has drawn the attention of researcher in the past few decades. The problem of estimation is concerned with the parameters of the queueing process such as arrival rate, service rate and traffic intensity. It is the most important thing in the queueing systems [6]. The pioneer investigators have derived the Maximum Likelihood Estimates (MLE) for the arrival and service parameters of an M/M/1 queueing model [9] and an infinite server queueing model [4]. The hypothesis testing for the point and the interval estimations of the M/M/1 queueing model using Bayesin approaches [7] and the non-zero waiting time of the model has been disscussed by using the convential and the Bayesian approaches along with the risk factors [8]. The five different approaches has been applied for the constructed 100(1-α) % of the Confidence Interval (CI) of the intensity of the queuing system [22]. Examining the MLE and Moment Estimate (ME) of the parameters of the inter-arrival and the service time distributions of GI/G/1 queueing model are discussed [3]. Consequently, the inferential procedures are concerned with the traffic intensity of M/E_k/1 queueing model which discussed [16]. The stationary solution of MLE of Markovian two server queueing model parameters have been obtained in the case of the non-identical servers [11]. Later, the stationary solution of the MLE of the generalized form of the multi-server queueing model in the presence of the non-identical servers and some of the CI of these model parameters are obtained [28]. Meanwhile, the MLE and the Bayesian estimates of the M/M/1/1 queueing model parameters are explained and the large sample test for the model parameters are also discussed [17]. The inferential process for the parameters of the bulk service queues is derived by using the Bayesian hierarchical model approaches [1]. Recently, the single server queueing model with working vacations has considered based on MLE approaches and simulation studies are carried out by the performance measures of the model [21].

The service times and the inter-arrival times of queueing model are not followed by the exponential distribution because of the high variability is observed in the inter-arrival time and the service time, most of the times are smaller than the minor proportion of the time and this leads to the characterisation of the heavy tails not only by the exponentially distributed [25]. In this regard, serveral authors have been devoted by queueing models based on the heavy tailed distribution [13], [15], [26]. Weibull, Parato, lognormal, Burr type III, Burr type XII and Gumbel distributions are some heavy tailed behaviour distributions [19]. The Bayesian estimation for the double Pareto lognormal (dPlN) distribution which has been proposed by the model in the queueing system for the heavy-tailed phenomena [10]. The evaluation of M/G/1 queueing model with the service time as assumed to Gumbel distribution, which has been explained by numerically and graphically based on the various combinations of the arbitrary values [20]. The extended queueing model when service time distributed according to Gumbel distribution under multiple working vacations scenario and the model parameters has been estimated based on Bayesian approaches with Gibbs sampling algorithm through Markov Chain Monte Carlo (MCMC) technique [18].

This article introduces tele-traffic and insurance data and some of the unusual characteristics of these types of data which motivate some of the inter-arrival and service time model that are analyzed through heavy tailed nature, particularly in Gumbel distribution. In insurance context, the claim sizes can take on extremely large values so they can be well modeled by heavy-tailed distribution. However, one difference between the insurance data and the internet traffic data is that in the insurance context, high autocorrelations are not observed to such an extent as with the tele-traffic data and that the insurance claims processes do not exhibit burstiness so much as the tele-traffic data, which suggests that heavy-tailed, but independent distributions may be reasonable for modeling insurance claims data in many contexts [12].

This paper proposes the new queueing model when exponential times of the inter-arrival time and service time are disappeared due to unusual characteristics. Therefore, the inter-arrival times of two successive arrival of customers and service times becomes a heavy tailed. For this reason, here the inter-arrival times and service times of the system follows Gumbel distribution. No attempts are found in the literature on evaluating the queueing models under Gumbel distribution based on Bayesian approaches. Determination of Gumbel/Gumbel/1 queueing model using Bayesian approach is discussed. The posterior distribution of the queueing model is derived incorporating the natural conjugate prior and non-informative prior to the parameters of the Gumbel distribution. The objective of this paper is to analyse the traffic congestion of the Gumbel/Gumbel/1 queueing model satisfying the stability condition of the system.

The probability generating function and cumulative distribution function of the Gumbel distribution are based on the location parameter, α and the scale parameter, β, respectively,

$$f(x:\alpha,\beta)=\frac{1}{\beta}e^{-\frac{(x-\alpha)}{\beta}}e^{-e^{-\frac{(x-\alpha)}{\beta}}} \quad for \quad x\in\Re,\alpha\in\Re;\beta>0 \quad and \tag{1.1}$$

$$F(x)=e^{-e^{-\frac{(x-\alpha)}{\beta}}} \tag{1.2}$$

with the mean $\alpha+\beta\gamma$ where $\gamma=0.5722...$ is the Euler's constant.

This paper is organized into the five sections, this is being the first. Section 2 contains model descriptions. The frame work of Bayesian estimation of model parameters is presented in section 3. The computational studies for the empirical Bayesian estimates by using Gibbs sampling algorithm in MCMC technique of the model are discussed in section 4 and section 5 provides the summary and conclusion of this work.

2. Model descriptions

Consider an Gumbel/Gumbel/1 queueing model,

- The inter-arrival time of two consecutive arrival of the customers which follows Gumbel distribution (α,β) with mean inter-arrival time, $1/\lambda=1/[\alpha+\beta\zeta]$ where Euler's constant, $\zeta=$ 0.5277....
- The service time of the system is distributed according to Gumbel distribution (γ,δ) with mean service time, $1/\mu=1/[\gamma+\delta\zeta]$ where Euler's constant, $\zeta=$ 0.5277....
- The inter-arrival times and the service times are mutually independent of each other.
- The server gives the services the single stage service with First In First Out (FIFO) discipline.
- In order to learn about the congestion of the system, the inference about the parameters governing the whole system $\theta=\{\alpha,\beta,\gamma,\delta\}$ is considered.

- The queueing system is consolidated and operated for the long time which indicates that it is working at equilibrium and satisfies the ergodic condition.
- Note that, the ergodic assumption implies that the parameters can only move freely in the reduced parametric space $\Theta_e = \{\theta : \lambda < \mu, \lambda, \mu > 0\}$. Hence, the traffic intensity of the model is $\rho = \frac{\gamma+\delta\zeta}{\alpha+\beta\zeta} < 1$.

3. Estimation on Gumbel/Gumbel/1 model

The Bayesian methodology consists of the sample information along with the prior information available about the parameter before the sample has been observed. The Bayesian approach treats that the model parameters are the random variables. The suitable probability distribution is determined for the models parameters for the queueing system say $\tau(\theta)$ with reference to the prior information. The information about the parameter given by the sample $\underline{x}$ is obtained from the likelihood function, $L(\theta|\underline{x})$. A prior probability distribution that represents perfect ignorance or indifference would produce the posterior probability distribution that represents that one should need about the parameter on the basis of the evidence alone. The prior distributions can be classified into two main categories like the informative prior and non-informative prior (vague, objective, and diffuse). The informative prior expresses the previous knowledge about parameter and the non-informative prior provides the formal way of expressing ignorance of the value of the parameter over the permitted range. The efforts to construct the priors may be represented by the absence of the knowledge. They have failed because no probability distribution to represent the pure ignorance. Combining these two information, the updated information about the parameter is obtained as the posterior distribution, $\tau(\theta|\underline{x})$. The inference about the parameter, θ is drawn from this posterior distribution. More details about the Bayesian methods can be found in [2], [27].

In the Gumbel/Gumbel/1 queueing model, the n_a inter-arrival times $x_a = (x_{1_a}, x_{2_a}, ..., x_{n_a})$ are a random samples distributed according to Gumbel (α, β) and the n_s recorded service time $x_s = (x_{1_s}, x_{2_s}, ..., x_{n_s})$ constitute a random sample from Gumbel (γ, δ). The joint probability generating function of this model is

$$f(x|\theta) = \frac{1}{\beta} e^{\frac{(x_a-\alpha)}{\beta}} e^{e^{\frac{(x_a-\alpha)}{\beta}}} \frac{1}{\delta} e^{\frac{(x_s-\gamma)}{\delta}} e^{e^{\frac{(x_s-\gamma)}{\delta}}} \quad \forall \quad x_a,\ x_s > 0 \tag{3.1}$$

where $\theta = \{\lambda < \mu; \quad \alpha, \beta, \gamma \quad and \quad \delta > 0\}$.

From Eqn. 3.1, the corresponding likelihood equation are as follows,

$$L(\theta|x) = \Pi_{i=1}^{n} \frac{1}{\beta^{n_a}} e^{\frac{(x_{i_a}-\alpha)}{\beta}} e^{e^{\frac{(x_{i_a}-\alpha)}{\beta}}} \Pi_{j=1}^{n} \frac{1}{\delta^{n_s}} e^{\frac{(x_{j_s}-\gamma)}{\delta}} e^{e^{\frac{(x_{j_s}-\gamma)}{\delta}}} \tag{3.2}$$

where, $x_a = \sum_{i=1}^{n} x_{i_a}$ is the total time until the arrival of n_a customer considered in the queue and $x_s = \sum_{j=1}^{n} x_{j_s}$ is the total time taken by the server to compelete the service under consideration. Note that, the restriction in the domain of the likelihood in Eqn. 3.2 corresponding to the ergodic condition of the queueing model.

Suppose that, the inverted Gamma distribution is employed as a probability model for the inter-arrival and service parameters based on the information obtained from the history of previous process of the queues respectively. The inverted Gamma distribution is a natural conjugate prior for sampling from the gumbel distribution for the inter-arrival and service parameters. The probability density function of inverted Gamma distribution is given in Eqn. 3.3.

$$\tau(c,d) = \begin{cases} \frac{d}{\Gamma(c)} x^{-(c+1)} e^{-c/x} & \text{c,d>0; x>0} \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

In certain situations, especially in the investigation of new problems of a pioneering nature, useful prior information may not be available. In such situations, the statistician will be forced to select a prior distribution which will reflect a situation of no prior information. This led to the notion of vague or diffused or non-informative prior distributions. The parameters is continuous and can take any value in a finite interval, then one can use a continuous uniform distribution as the prior distribution for the parameter. Such prior distributions are called non-informative priors and sometimes as vague priors (see more [5], [2], [27]). Furthermore, it may be considered that the uniform distribution is a non-informative prior knowledge about the model parameters α, β, γ and δ. The probability density function of uniform distribution is

$$\tau_3(\phi) = \frac{1}{q-p}; 0 < p \le \phi \le q \tag{3.4}$$

The updated inforamtions of posterior distribution is obtained for the model parameters is given by

$$\begin{aligned} \tau_I\left(\alpha,\beta,\gamma,\delta|data\right) \quad \propto \quad & \frac{d_1 d_2 d_3 d_4}{\beta^{n_a}\delta^{n_s}\Gamma\left(c_1\right)\Gamma\left(c_2\right)\Gamma\left(c_3\right)\Gamma\left(c_4\right)}\alpha^{-(c_1+1)}\beta^{-(c_2+1)} \times \\ & \gamma^{-(c_3+1)}\delta^{-(c_4+1)}e^{-(c_1/\alpha+c_2/\beta+c_3/\gamma+c_4/\delta)} \times \\ & \Pi_{i=1}^{n} e^{\frac{(x_{i_a}-\alpha)}{\beta}} e^{e^{\frac{(x_{i_a}-\alpha)}{\beta}}} \Pi_{j=1}^{n} e^{\frac{(x_{j_s}-\gamma)}{\delta}} e^{e^{\frac{(x_{j_s}-\gamma)}{\delta}}} \end{aligned} \tag{3.5}$$

$$\tau_{NI}\left(\alpha,\beta,\gamma,\delta|data\right) \propto \frac{1}{\beta^{n_a}\delta^{n_s}} \Pi_{i=1}^{n} e^{\frac{(t_{i_a}-\alpha)}{\beta}} e^{e^{\frac{(x_{i_a}-\alpha)}{\beta}}} \Pi_{j=1}^{n} e^{\frac{(x_{j_s}-\gamma)}{\delta}} e^{e^{\frac{(x_{j_s}-\gamma)}{\delta}}} \tag{3.6}$$

Since, the posterior distributions of the informative and the non-informative prior knowledges are not attained in the closed form expression. Hence, MCMC simulation technique is more appropriate to deal with the empirical estimates of the model parameters. The empirical Bayesian estimates are computed particularly through Gibbs sampling algorithm [24] using OpenBugs software.

4. Gibbs sampling algorithm in MCMC technique

The Markov chains have significant role in Bayesian statistics because it is generally possible to construct the Markov chain in such a way that the target distribution is the joint posterior distribution of all the unknown parameters in the Bayesian model. Thus, the Markov chain Monte Carlo methods provide a way of generating samples from the joint posterior distribution in the realistic and high-dimensional Bayesian models. The Gibbs sampling algorithm is a special case of Metropolis-Hastings sampling algorithm which one particular way of constructing the transition kernel to produce the Markov chain with the desired target distribution. The Gibbs sequence converges to the stationary(equilibrium) distribution that is independent of the initial values, and by the attaining this stationary distribution is the target distribution. The step-by-step procedure in Gibbs sampling alogrithm for the proposed queueing model as follows:

(1) Set initial values $\alpha^{(0)}$, $\beta^{(0)}$, $\gamma^{(0)}$, $\delta^{(0)}$
(2) For t=1,...,T
  (a) For i=1,2,...,n
    (i) Generate $x_i^{(t)}$ from $f\left(x|\alpha^{(t-1)}, \beta^{(t-1)}, \gamma^{(t-1)}, \delta^{(t-1)}\right)$
  (b) Generate $\alpha^{(t)} \sim \tau\left(\alpha|x^{(t)}\right)$
  (c) Generate $\beta^{(t)} \sim \tau\left(\beta|x^{(t)}\right)$
  (d) Generate $\gamma^{(t)} \sim \tau\left(\gamma|x^{(t)}\right)$
  (e) Generate $\delta^{(t)} \sim \tau\left(\delta|x^{(t)}\right)$

The MCMC samples are generated through Gibbs sampling algorithm from the posterior distribution of model parameters for the given set of the informative priors Eqn. 3.3 and non-informative prior Eqn. 3.4 for obtaining the Bayes estimates of the model. The markov chain is run in OpenBugs for 10,000 number of iterations for various arbitrary values and samples.

4.1. Convergence diagnostics of MCMC. From the outputs of OpenBugs, the diagnostic checking plots for each model parameters are presented in Appendix. The Markov chain has converged in both informative and non-informative priors since it likely to be sampling from the stationary distribution and horizontal band, with no long upward or downward trends as shown in Figure [15, 16, 17, 18]. Moreover, the autocorrelation is almost negligible for all the model parameters (see Figure[19, 20, 21, 22]). Therefore, the generated samples, in each iteration from posterior densities under informative and non-informative priors are independent to each other. Further, the kernal densities of model parameters α, β, γ, δ for various samples 50, 100, 150, 200, 250 and $\alpha = 0.2, 0.3, 0.4, 0.5$, $\beta = 0.3, 0.4, 0.5, 0.6$, $\gamma = 0.1, 0.2, 0.3, 0.4$, & $\delta = 0.2, 0.3, 0.4, 0.5$ under informative and non-informative priors are displayed (see Figure[23, 24, 25, 26]) for checking the convergence of the algorithm. Also, the Monte Carlo Error (MC.E) of Gumbel/Gumbel/1 queueing model is presented Table.1 & Table.3. It is to be observed that, MC error is minimum for each estimates in model parameters.

4.2. Numerical results of Bayesian estimation. The posterior mean and 95 % credible region of Gumbel/Gumbel/1 queueing model parameters are presented in Table.1 - Table.4. Meanwhile, the empirical Bayesian estimates of traffic intensity of the model are computed from the posterior mean of corresponding parameter and it is to be observed that in Figure [1, 2] , the stable intensity level has been maintained when sample observations and values of model parameters increase in both prior informations. The congenstion level of the each model belongs to the interval of 0.5 - 0.95.

5. Summary and conclusion

In this paper, the Bayesian estimates of an Gumbel/Gumbel/1 queueing model under the informative and non- informative prior knowledges is considered. The empirical posterior mean, 95 % credible region and the diagonostic checking plots are carried out for various size of the sample observations and different sets of arbitrary values based on Gibbs sampler through the MCMC simulation technique using the OpenBugs software. From those results, the traffic intensity of the model has been increased when the model parameters of inter-arrival time and service time are increased but not in the increasing size sample observations. Meanwhile, the stable intensity level has been maintained when the sample observations and the values of model parameters are increased in both

Table 1. Empirical Bayesian estimates of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples based on informative priors

Arbitrary	*Samples*	$\hat{\alpha}$	*MC.E*	$\hat{\beta}$	*MC.E*	$\hat{\gamma}$	*MC.E*	$\hat{\delta}$	*MC.E*
$\alpha = 0.2$,	50	0.257	0.00040	0.233	0.00038	0.195	0.00026	0.212	0.00029
$\beta = 0.3$,	100	0.331	0.00030	0.268	0.00027	0.178	0.00017	0.154	0.00015
$\gamma = 0.1$,	150	0.331	0.00024	0.258	0.00018	0.218	0.00016	0.189	0.00014
$\delta = 0.2$	200	0.294	0.00019	0.250	0.00017	0.193	0.00014	0.156	0.00011
	250	0.299	0.00020	0.264	0.00014	0.183	0.00011	0.154	0.00093
$\alpha = 0.3$,	50	0.359	0.00050	0.290	0.00041	0.297	0.00045	0.338	0.00050
$\beta = 0.4$,	100	0.331	0.00029	0.268	0.00025	0.279	0.00025	0.238	0.00021
$\gamma = 0.2$,	150	0.402	0.00035	0.333	0.00025	0.318	0.00022	0.259	0.00019
$\delta = 0.3$	200	0.407	0.00024	0.345	0.00022	0.314	0.00019	0.247	0.00017
	250	0.428	0.00020	0.311	0.00018	0.298	0.00017	0.253	0.00013
$\alpha = 0.4$,	50	0.574	0.00057	0.411	0.00050	0.435	0.00054	0.587	0.00062
$\beta = 0.5$,	100	0.464	0.00038	0.379	0.00032	0.347	0.00036	0.342	0.00027
$\gamma = 0.3$,	150	0.508	0.00043	0.473	0.00033	0.386	0.00026	0.316	0.00021
$\delta = 0.4$	200	0.515	0.00038	0.446	0.00028	0.428	0.00026	0.353	0.00025
	250	0.484	0.00030	0.419	0.00025	0.428	0.00023	0.331	0.00019
$\alpha = 0.5$,	50	0.687	0.00073	0.585	0.00061	0.468	0.00069	0.440	0.00055
$\beta = 0.6$,	100	0.464	0.00035	0.379	0.00030	0.460	0.00050	0.394	0.00039
$\gamma = 0.4$,	150	0.582	0.00045	0.487	0.00038	0.560	0.00041	0.473	0.00032
$\delta = 0.5$	200	0.670	0.00039	0.551	0.00037	0.571	0.00037	0.473	0.00028
	250	0.608	0.00035	0.513	0.00023	0.541	0.00031	0.415	0.00023

Table 2. 95% Credible region of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples based on informative priors

Arbitrary Values	*Samples*	α		β		γ		δ	
		LB	UB	LB	UB	LB	UB	LB	UB
$\alpha = 0.2$,	50	0.1908	0.3266	0.1823	0.3002	0.1578	0.2697	0.1541	0.2491
$\beta = 0.3$,	100	0.2778	0.3880	0.2268	0.3164	0.1471	0.2099	0.1310	0.1811
$\gamma = 0.1$,	150	0.2887	0.3762	0.2270	0.2956	0.1872	0.2500	0.1654	0.2158
$\delta = 0.2$	200	0.2584	0.3322	0.2235	0.2802	0.1709	0.2156	0.1397	0.1749
	250	0.2653	0.3337	0.2385	0.2934	0.1632	0.2035	0.1388	0.1707
$\alpha = 0.3$,	50	0.2777	0.4459	0.2297	0.3685	0.2373	0.3763	0.2539	0.4272
$\beta = 0.4$,	100	0.2774	0.3865	0.2267	0.3166	0.2313	0.3280	0.204	0.2797
$\gamma = 0.2$,	150	0.3458	0.459	0.2932	0.381	0.2765	0.3619	0.2279	0.2958
$\delta = 0.3$	200	0.3589	0.4580	0.3071	0.3873	0.2792	0.3507	0.2217	0.2769
	250	0.3888	0.4688	0.4688	0.3444	0.2658	0.3309	0.2286	0.2800
$\alpha = 0.4$,	50	0.4551	0.6915	0.3293	0.5134	0.3503	0.5456	0.4655	0.7124
$\beta = 0.5$,	100	0.3879	0.5410	0.3242	0.4439	0.2767	0.4197	0.2900	0.4040
$\gamma = 0.3$,	150	0.4322	0.5860	0.4152	0.5422	0.3352	0.4396	0.2782	0.3627
$\delta = 0.4$	200	0.4516	0.5802	0.3999	0.5000	0.3766	0.4809	0.3140	0.3958
	250	0.4297	0.5382	0.3790	0.4637	0.3865	0.4716	0.2995	0.3665
$\alpha = 0.5$,	50	0.5258	0.8328	0.4738	0.7197	0.3503	0.5582	0.3422	0.5937
$\beta = 0.6$,	100	0.3877	0.5430	0.3238	0.4437	0.3823	0.5415	0.3358	0.4632
$\gamma = 0.4$,	150	0.5016	0.6628	0.4265	0.5559	0.4819	0.6391	0.4163	0.5381
$\delta = 0.5$	200	0.5925	0.7496	0.4936	0.6155	0.5031	0.6422	0.4224	0.5303
	250	0.5417	0.6749	0.4647	0.5683	0.4880	0.5964	0.3756	0.4589

prior informations. Hence, this model is appropriate for studying the queueing system when the inter-arrival time and service time follows heavy tailed particularly in Gumbel distribution.

Acknowledgements. The authors are grateful to the member of Editorial Board and anonymous referees who made significant suggestions for improvement of the paper. The writers express their gratitude to the UGC-SAP(DRS-I) for providing the financial support to carry out this work.

References

[1] Armero, C. and Conesa, D. *Bayesian hierarchical models in manufacturing bulk service queues*, Journal of Statistical Planning and Inference, 136, 335-354, 2006.

Table 3. Empirical Bayesian estimates of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples based on non-informative priors

Arbitrary	*Samples*	$\hat{\alpha}$	*MC.E*	$\hat{\beta}$	*MC.E*	$\hat{\gamma}$	*MC.E*	$\hat{\delta}$	*MC.E*
$\alpha = 0.2$,	50	0.251	0.00035	0.227	0.00030	0.210	0.00029	0.193	0.00027
$\beta = 0.3$,	100	0.329	0.00028	0.266	0.00024	0.177	0.00018	0.153	0.00014
$\gamma = 0.1$,	150	0.351	0.00030	0.261	0.00023	0.215	0.00018	0.186	0.00018
$\delta = 0.2$	200	0.293	0.00019	0.248	0.00016	0.192	0.00013	0.155	0.00012
	250	0.298	0.00018	0.264	0.00013	0.182	0.00010	0.153	0.00093
$\alpha = 0.3$,	50	0.357	0.00044	0.287	0.00041	0.339	0.00053	0.297	0.00043
$\beta = 0.4$,	100	0.410	0.00036	0.333	0.00032	0.278	0.00029	0.238	0.00022
$\gamma = 0.2$,	150	0.419	0.00044	0.339	0.00035	0.313	0.00028	0.262	0.00027
$\delta = 0.3$	200	0.407	0.00030	0.345	0.00022	0.313	0.00020	0.247	0.00016
	250	0.428	0.00023	0.310	0.00015	0.297	0.00021	0.252	0.00018
$\alpha = 0.4$,	50	0.590	0.00078	0.415	0.00065	0.602	0.00006	0.441	0.00057
$\beta = 0.5$,	100	0.466	0.00045	0.379	0.00037	0.347	0.00039	0.342	0.00029
$\gamma = 0.3$,	150	0.512	0.00043	0.476	0.00037	0.387	0.00027	0.317	0.00024
$\delta = 0.4$	200	0.518	0.00045	0.448	0.00030	0.428	0.00026	0.353	0.00022
	250	0.485	0.00031	0.419	0.00025	0.430	0.00022	0.332	0.00020
$\alpha = 0.5$,	50	0.750	0.00102	0.616	0.00091	0.475	0.00007	0.446	0.00069
$\beta = 0.6$,	100	0.770	0.00065	0.589	0.00057	0.462	0.00049	0.396	0.00036
$\gamma = 0.4$,	150	0.588	0.00043	0.490	0.00036	0.566	0.00051	0.475	0.00036
$\delta = 0.5$	200	0.682	0.00047	0.557	0.00038	0.575	0.00035	0.476	0.00030
	250	0.613	0.00037	0.516	0.00029	0.543	0.00030	0.415	0.00023

Table 4. 95% Credible region of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples based on non-informative priors

Arbitrary	*Samples*	α		β		γ		δ	
Values		LB	UB	LB	UB	LB	UB	LB	UB
$\alpha = 0.2$,	50	0.1846	0.3178	0.1781	0.2925	0.1546	0.2681	0.1525	0.2452
$\beta = 0.3$,	100	0.2759	0.3854	0.2268	0.3127	0.1460	0.2090	0.1307	0.1800
$\gamma = 0.1$,	150	0.2971	0.4060	0.2230	0.3063	0.1784	0.2537	0.1587	0.2212
$\delta = 0.2$	200	0.2573	0.3295	0.2216	0.2795	0.1701	0.2152	0.1391	0.1748
	250	0.2634	0.3325	0.2384	0.2925	0.1627	0.2027	0.1386	0.1702
$\alpha = 0.3$,	50	0.2725	0.4415	0.2282	0.3622	0.2516	0.4292	0.2366	0.3766
$\beta = 0.4$,	100	0.3424	0.4788	0.2822	0.3941	0.2296	0.3274	0.2033	0.2788
$\gamma = 0.2$,	150	0.3495	0.4894	0.2890	0.3989	0.2607	0.3675	0.2235	0.3093
$\delta = 0.3$	200	0.3576	0.4583	0.30751	0.3883	0.2781	0.3506	0.2217	0.2764
	250	0.3877	0.4692	0.2809	0.3430	0.2643	0.3307	0.2280	0.2795
$\alpha = 0.4$,	50	0.4634	0.7169	0.3279	0.5256	0.4744	0.7313	0.3517	0.554
$\beta = 0.5$,	100	0.3878	0.5430	0.3235	0.4451	0.2778	0.4198	0.2904	0.4056
$\gamma = 0.3$,	150	0.4332	0.5950	0.4164	0.5470	0.3343	0.4407	0.2775	0.3626
$\delta = 0.4$	200	0.4551	0.5855	0.3999	0.5042	0.3767	0.4815	0.3137	0.3966
	250	0.4314	0.5408	0.3793	0.4638	0.3864	0.4731	0.2999	0.3686
$\alpha = 0.5$,	50	0.5712	0.9297	0.4893	0.7790	0.3451	0.6087	0.3527	0.5693
$\beta = 0.6$,	100	0.6481	0.8922	0.5001	0.6953	0.3816	0.548	0.3376	0.4654
$\gamma = 0.4$,	150	0.5064	0.6710	0.4306	0.5620	0.4862	0.6497	0.4183	0.5428
$\delta = 0.5$	200	0.6010	0.7654	0.4979	0.6260	0.5065	0.6471	0.4257	0.5340
	250	0.5474	0.6807	0.4661	0.5713	0.4902	0.598	0.3761	0.4600

[2] Bansal, A. K. *Bayesian parametric inference,* Narosa Publishing House, 2007.

[3] Basava, I. V. and Prabhu, N. U. *Estimation in single server queues*, Naval Research Logistics Quart., 28(3), 475-487, 1981.

[4] Benes, V. E. A sufficient set of statistics for a simple telephone exchange model, Bell System Technology Journal, 36, 939-964, 1957.

[5] Berger, J. O. *Statistical decision theory and Bayesian analysis*, Second Edition, Springer, New York, 1980.

[6] Bhat, U. N. An introduction to queueing theory-modeling analysis in applications, Birkhauser Boston, USA, 2008.

[7] Choudhury, A. and Borthakur, A. C. *Bayesian inference and prediction in the single server markovian queue*, Metrika, 67, 371-383, 2008.

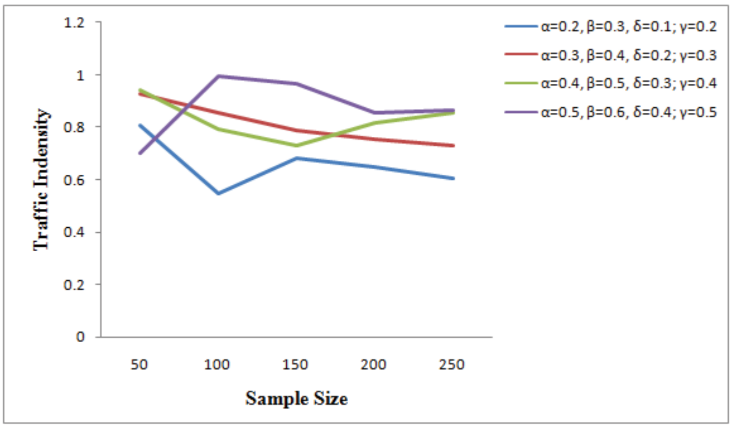


Figure 1. Traffic intensity of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples under informative prior

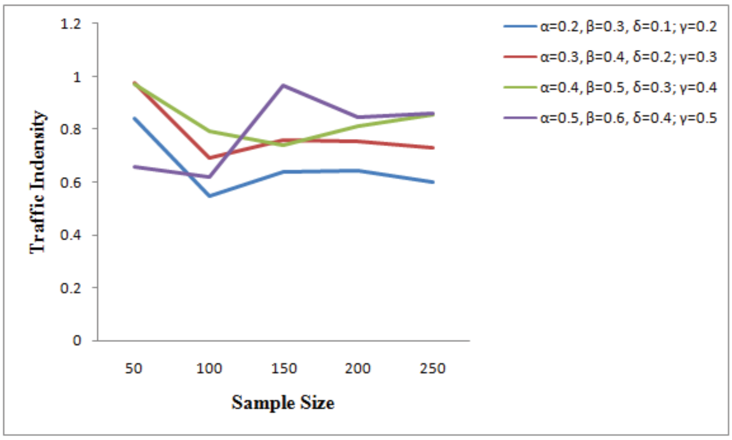


Figure 2. Traffic intensity of Gumbel/Gumbel/1 queueing model for various arbitrary values and samples under non-informative prior

[8] Chowdhury, S. and Mukherjee, S. P. *Estimation of waiting time distribution in an M/M/1 queue*, Opsearch, 48(4), 306-317, 2011.

[9] Clake, A. B. *Maximum likelihood estimates in a simple queue*, Annals of Mathematical Statistics, 28, 1036-1040, 1957.

[10] Cobo, P. R., Lillo, R. E., Wilson, S. and Wiper, M. P. *Bayesian inference for double Pareto lognormal queues*, The Annals of Applied Statistics, 4(3), 1533-1557, 2010.

[11] Dave, A. B. and Shah, Y. K. *Maximum likelihood estimates in an M/M/2 queue with heterogeneous servers*, Journal of Operational Research Society, 31, 423-436, 1980.

[12] Embrechts, P., C. Kluperlberg and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin, Heidelberg.

[13] Feldmann, A and Whit, W. *Fitting mixtures of exponentilas to long-tail distributions to analyze network performance model*, Performance Evaluation, 31, 245-279, 1998.

[14] Gross, D., Shortle, J. F., Thompson, J. M. and Harris, C. M. *Fundamentals of queueing theory*, Wiley Series in Probability and Statistics, USA, 2008.

[15] Glynn, P. W. and Whitt, W. *Heavy-traffic extreme-value limits for queue*, Operations Research Letters, 18, 107-111, 1995.

[16] Harishchandra, K. and Subba Rao, S. *A note on statistical inference about the traffic intensity parameter in $M/E_k/1$ queue*, The Indian Journal of Statistics, 50(1), 144-148, 1988.

[17] Harishchandra, K and Lakshmi, S. K. *Maximum likelihood and Bayesian estimation of the parameters of the M/M/1/1 queueing system*, Carmelight, 7, 115-124, 2010.

[18] Jabarali, A. and Kannan, K. S. *Bayesian queueing model with multiple working vacations under Gumbel distribution*, Applied Mathematical Sciences, 9(48), 2363-2369, 2015.

[19] Johnson, N. L., Kotz, S and Balakrishnan, K. *Continuous univariate distributions*, Second Edition, John Wiley and Sons, Inc., New York, 1995.

[20] Kannan, K. S. and Jabarali, A. *Queueing model with Gumbel service time distribution in teletraffic*, Proceeding of the International Conference on Mathematical Sciences, Elsevier, 767-771, 2014a.

[21] Kannan, K. S. and Jabarali, A. *Estimation on single server queueing model with working vacations*, Research and Reviews: Journal of Statistics(Special Issue on Recent Statistical Methodologies and Applications), 2, 94-98, 2014b.

[22] Ke, J. C and Chu, Y. K. *Comparison of five estimation approaches of intensity for a queueing system with short run*, Computational Statistics, 24, 567-582, 2009.

[23] Kumar, R., A. K. Srivastava and V. Kumar. *Analysis of Gumbel model for software reliability using Bayesian paradigm*, International Journal of Advanced Research in Artificial Intelligence, 1(2), 39-45, 2012.

[24] Ntzoufras. I. *Bayesian Modeling Using WinBugs*, Wiley Series in Computional Statistics, USA, 2009.

[25] Paxson, V and Floyd, S. *Wide-area traffic-the failure of poisson modeling*, IEEE Transactionsin Networking, 3, 226-244, 1995.

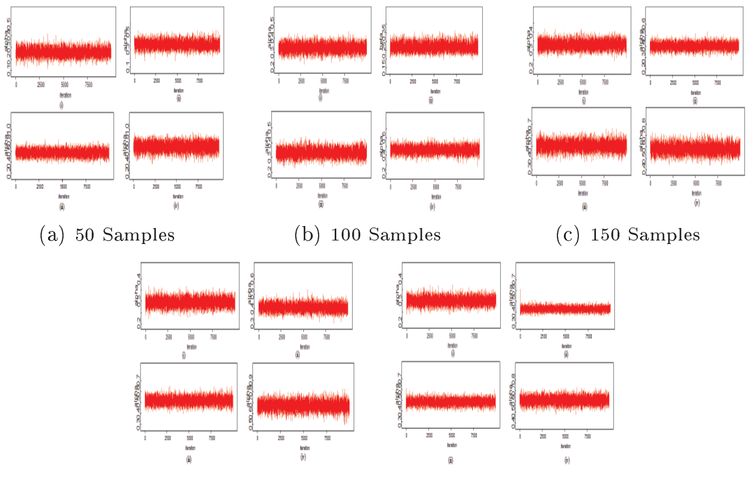
[26] Psounis, K., Molinero-fernandez, P., Prabhakarr, B. and Papadopoulos, F. *Systems with multiple servers under heavy-tailed workloads*, Performance Evaluation, 62, 456-474, 2005.

[27] Rajagopalan, M and P. Dhanavanthan *Statistical Inference*, PHI Learning Private Limited, New Delhi, 2012.

[28] Tsung-Yin Wand and Wang. *Maximum likelihood estimates and confidence intervals of an M/M/R queue with heterogeneous servers*. Mathematical Methods Operational Research, 63, 371-384, 2006.

Appendices

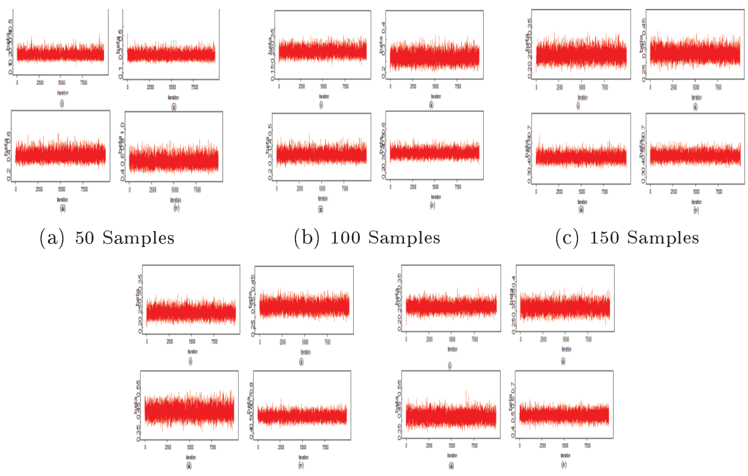
Diagnostics checking plots



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

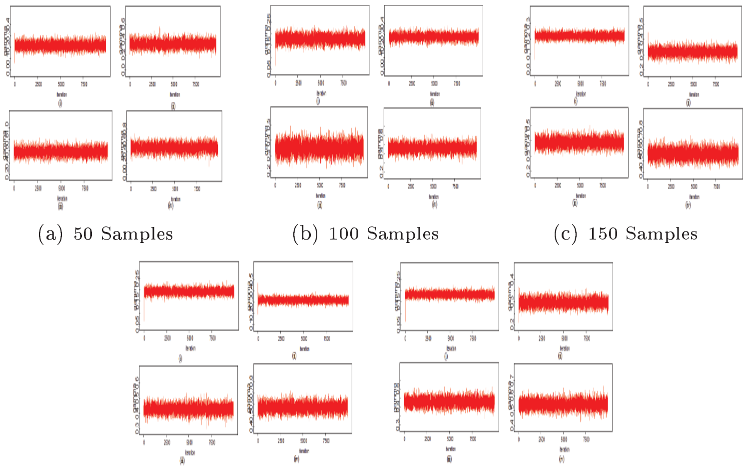
Figure 3. History plots of model parameter α under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

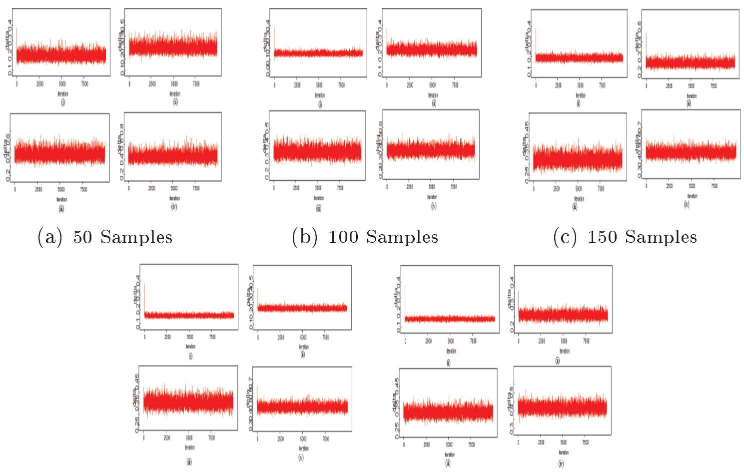
Figure 4. History plots of model parameter β under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

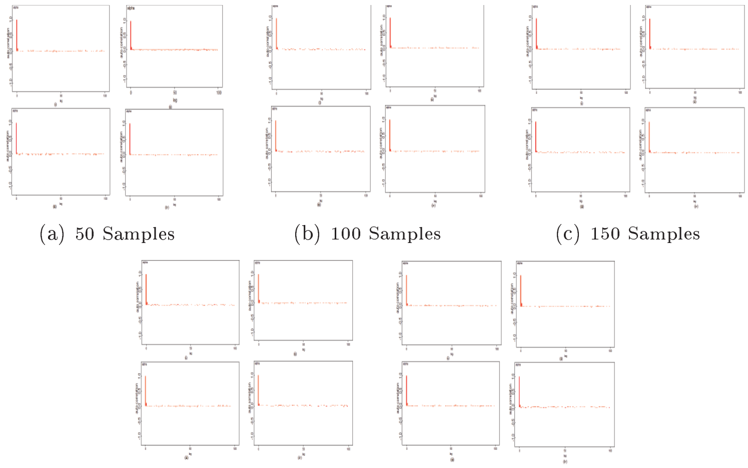
Figure 5. History plots of model parameter γ under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

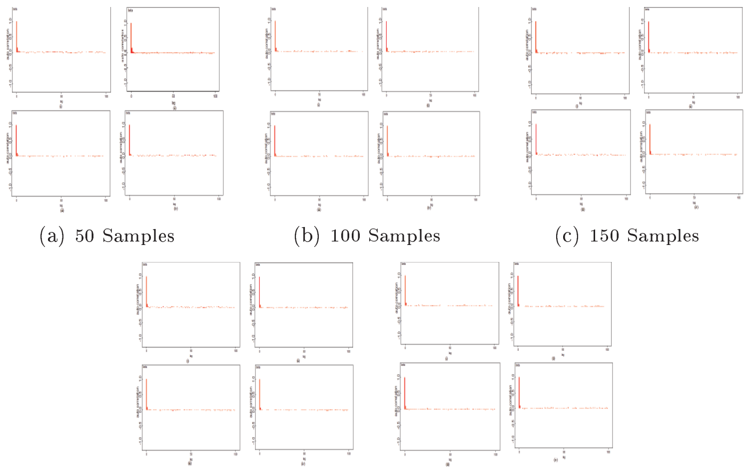
Figure 6. History plots of model parameter δ under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

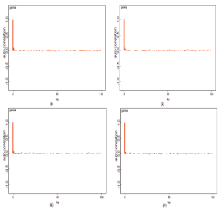
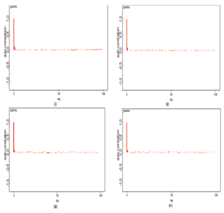
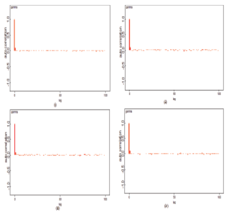
Figure 7. Auto correlation plots of model parameter α under informative prior for various samples and arbitrary values



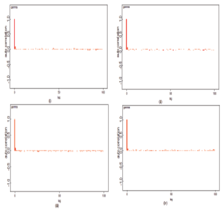
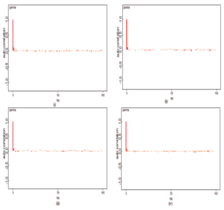
(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

Figure 8. Auto correlation plots of model parameter β under informative prior for various samples and arbitrary values

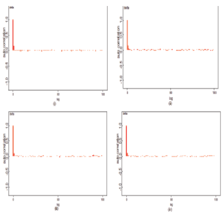
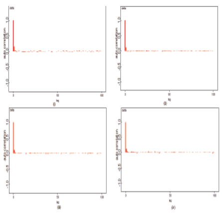
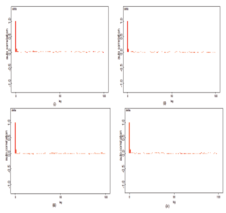




(a) 50 Samples (b) 100 Samples (c) 150 Samples

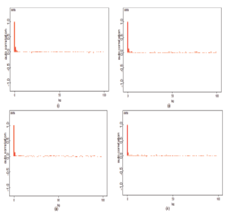
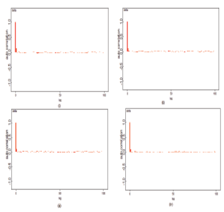



(d) 200 Samples (e) 250 Samples

Figure 9. Auto correlation plots of model parameter γ under informative prior for various samples and arbitrary values

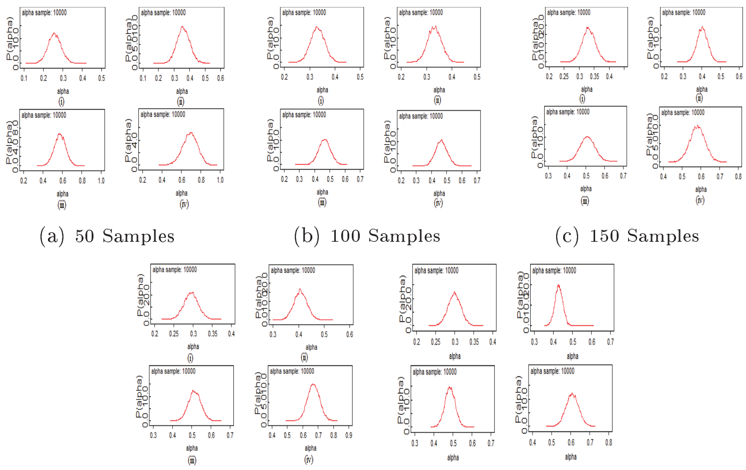




(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

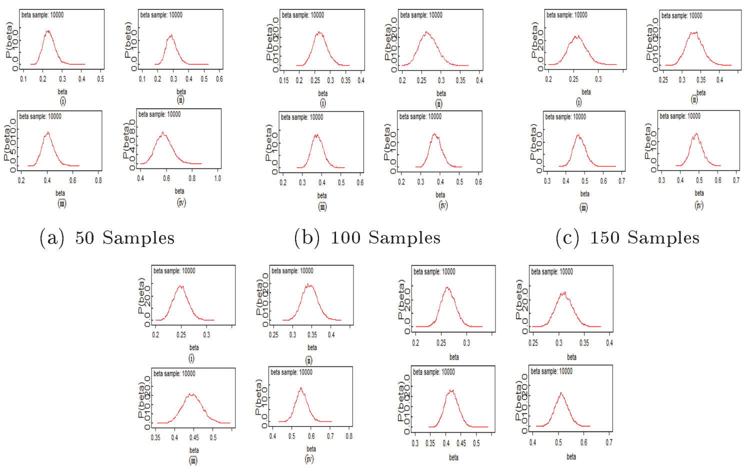
Figure 10. Auto correlation plots of model parameter δ under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

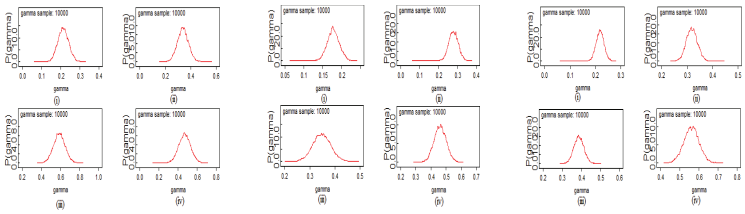
Figure 11. Kernal densities of model parameter α under informative prior for various samples and arbitrary values



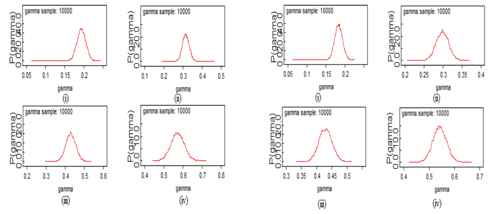
(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

Figure 12. Kernal densities of model parameter β under informative prior for various samples and arbitrary values

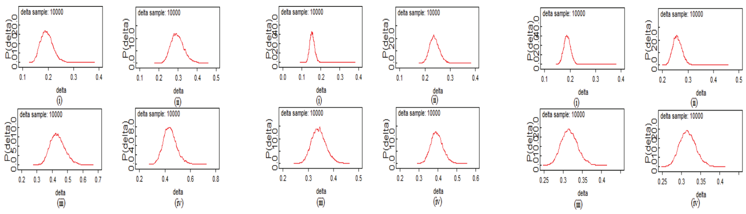


(a) 50 Samples (b) 100 Samples (c) 150 Samples

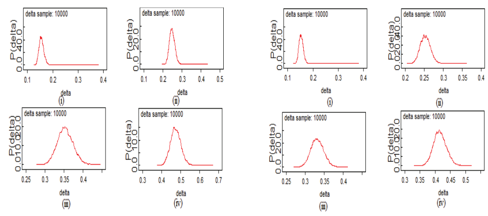


(d) 200 Samples (e) 250 Samples

Figure 13. Kernal densities of model parameter γ under informative prior for various samples and arbitrary values

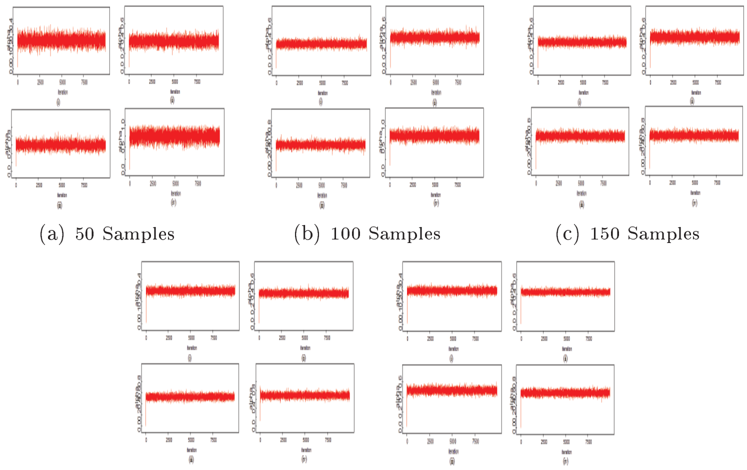


(a) 50 Samples (b) 100 Samples (c) 150 Samples



(d) 200 Samples (e) 250 Samples

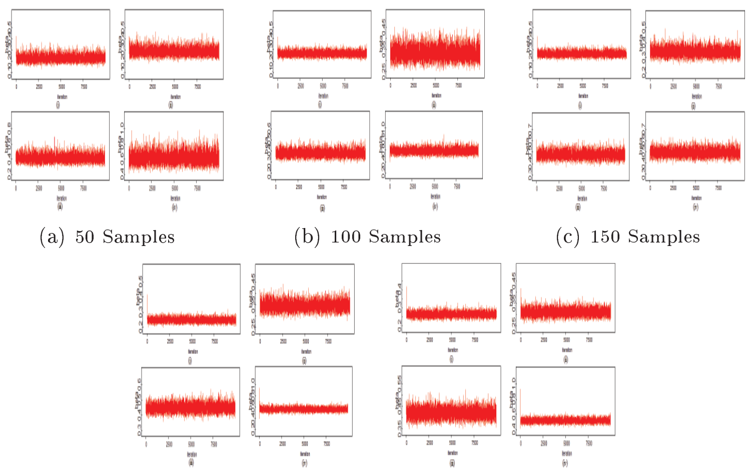
Figure 14. Kernal densities of model parameter δ under informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

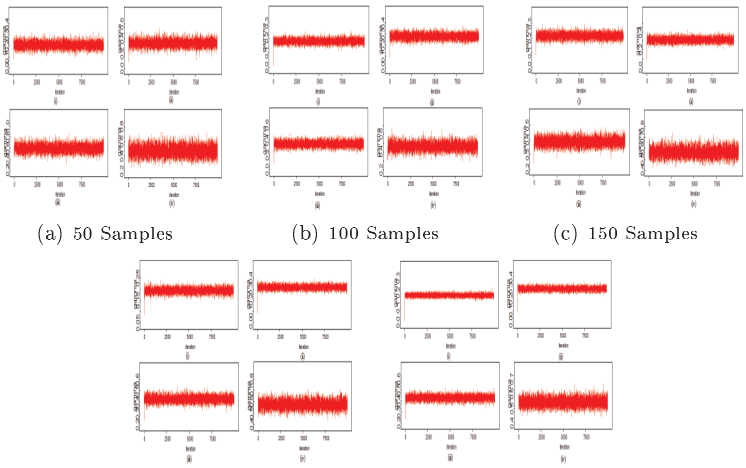
Figure 15. History plots of model parameter α under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

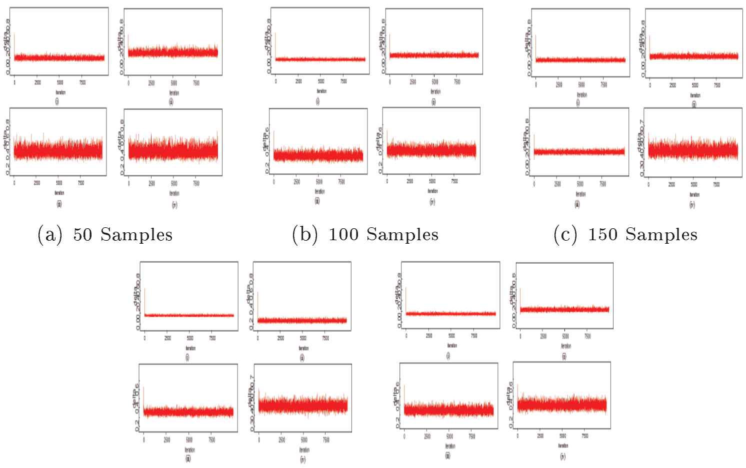
Figure 16. History plots of model parameter β under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

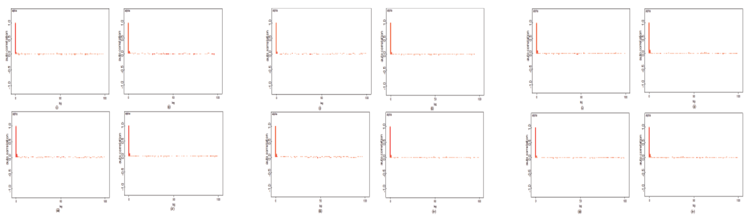
Figure 17. History plots of model parameter γ under non-informative prior for various samples and arbitrary values



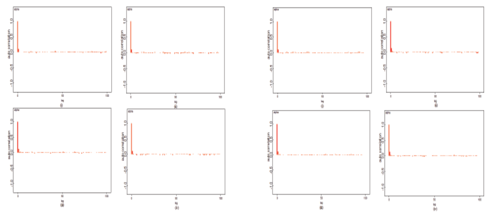
(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

Figure 18. History plots of model parameter δ under non-informative prior for various samples and arbitrary values

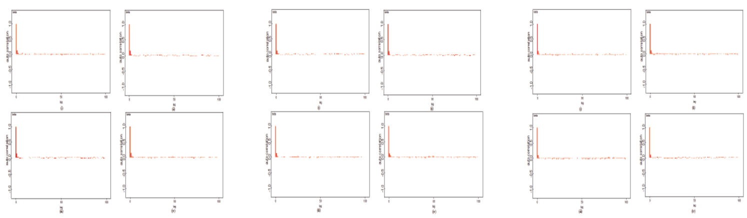


(a) 50 Samples (b) 100 Samples (c) 150 Samples

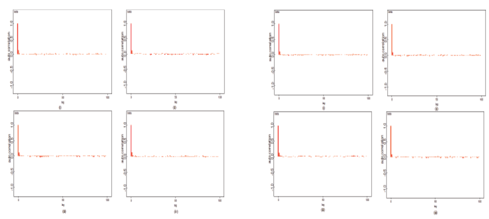


(d) 200 Samples (e) 250 Samples

Figure 19. Auto correlation plots of model parameter α under non-informative prior for various samples and arbitrary values

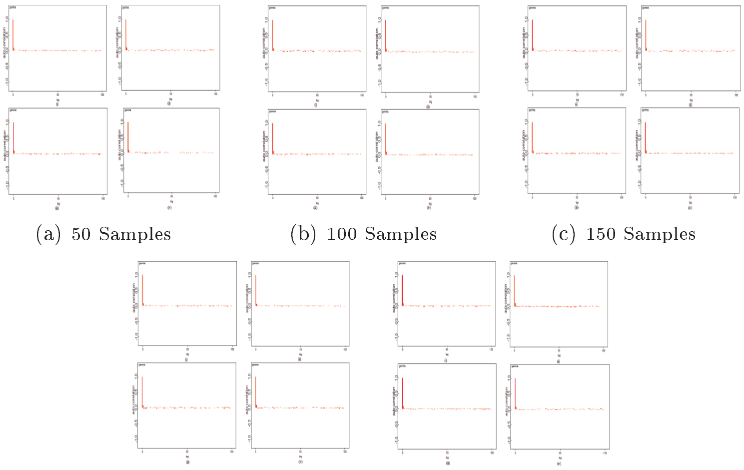


(a) 50 Samples (b) 100 Samples (c) 150 Samples



(d) 200 Samples (e) 250 Samples

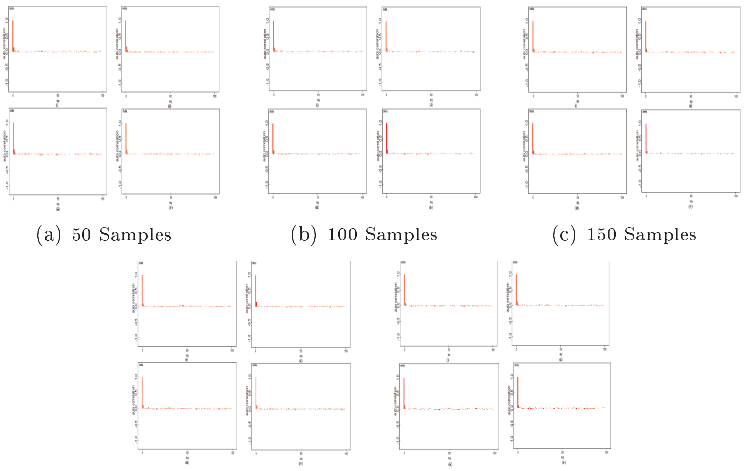
Figure 20. Auto correlation plots of model parameter β under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

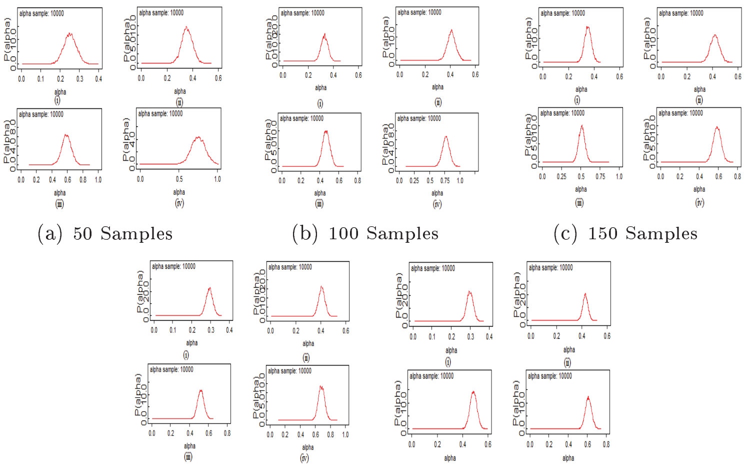
Figure 21. Auto correlation plots of model parameter γ under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

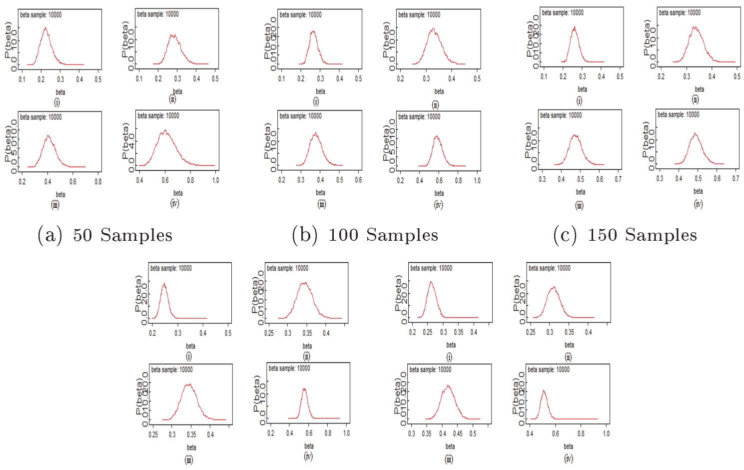
Figure 22. Auto correlation plots of model parameter δ under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

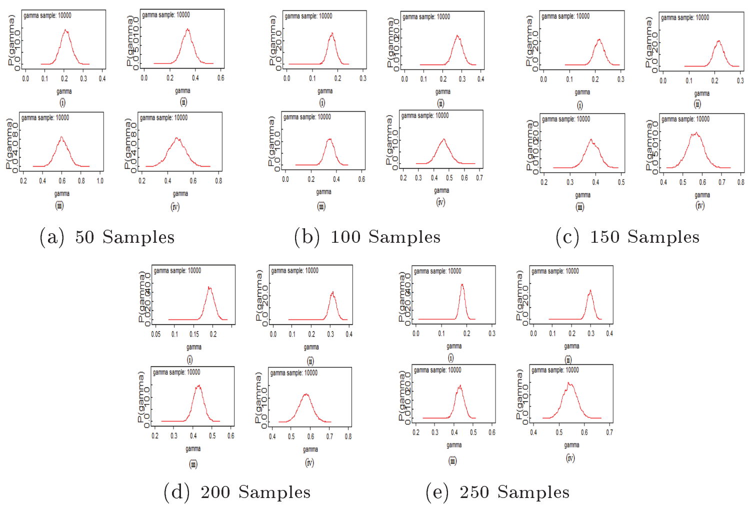
Figure 23. Kernal densities of model parameter α under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

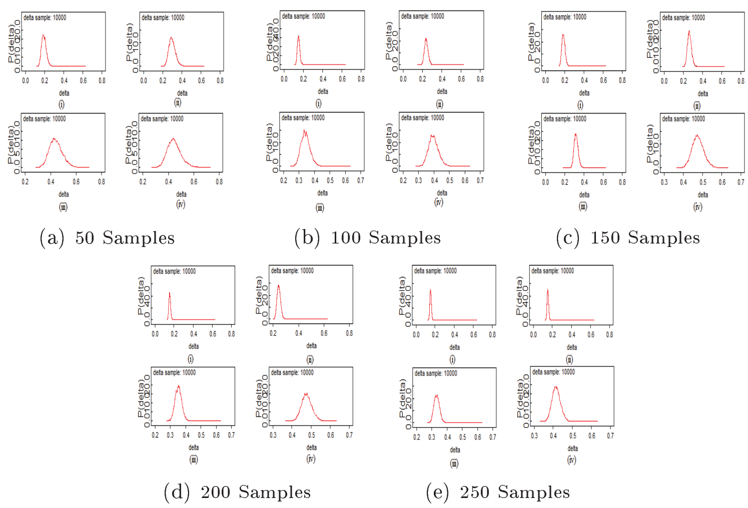
Figure 24. Kernal densities of model parameter β under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

Figure 25. Kernal densities of model parameter γ under non-informative prior for various samples and arbitrary values



(a) 50 Samples (b) 100 Samples (c) 150 Samples

(d) 200 Samples (e) 250 Samples

Figure 26. Kernal densities of model parameter δ under non-informative prior for various samples and arbitrary values

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1297 – 1306

Generalized multi-phase regression-type estimators under the effect of measuemnent error to estimate the population variance

Saadia Masood*†, Javid Shabbir‡

Abstract

In this article, we suggest some regression-type estimators for the estimation of finite population variance using multi-variate auxiliary information under multi-phase sampling schemes when measurement error (ME) contaminates the study variable. An empirical study is also carried out to judge the merits of proposed estimators.

Keywords: Multi-phase sampling, measurement error, mean square error, efficiency.

2000 AMS Classification: 62D05

Received : 28.01.2014 *Accepted* : 06.04.2015 *Doi* : 10.15672/HJMS.201612116352

1. Introduction

In sample surveys, it is customary to exploit the auxiliary information to enhance the precision of estimators. Ratio and regression estimators provide one type of example. Sometimes the sample units are chosen with probability proportionate to some measure of size based on the auxiliary variate. In all these cases it is information on just one auxiliary variate that is used for reasons of sample selection or estimation. Pretty often we take information on several variates and it may be considered important to make use of the whole of the available material to improve the precision of at least some of the key items in the survey (see Raj [10]). Isaki [7] has discussed multi-variate ratio estimators to estimate finite population variance S_y^2. Singh and Solanki ([16], [17]) and Solanki and Singh [19] proposed the procedure for variance estimation using auxiliary information under simple random sampling.

Two-phase sampling of a finite population occurs when a sample from the population is itself sampled, with the goal of determining variates in the sub-sample not already

*Department of Statistics Quaid-i-Azam University, Islamabad, Pakistan,
Email: saadia2masood@yahoo.com
†Corresponding Author: Saadia Masood.
‡Department of Statistics Quaid-i-Azam University, Islamabad, Pakistan,
Email: javidshabbir@gmail.com

available in the sample. An important example is the regression estimator for means or totals, which uses values of an auxiliary variable from the full sample to estimate the mean of a variable of interest that is available only on the subsample. Multi-phase sampling is not widely discussed in literature. Mukerjee et al. [9] considered mainly three phases. Singh [18] proposed a class of estimators for population variance under two-phase sampling, whose composition was partially defined for the single auxiliary variable. Dorfman [5] proposed regression estimator for estimation of population variance under two-phase sampling scheme. Allen et al. [1] proposed a family of estimators of population mean using multi-auxiliary information in presence of measurement errors.

In most of the statistical studies, it is one of the common believes that the data are error-free but usually in realistic circumstances this statement is not absolutely met and the data are infected by errors. The consequences made for the error free data become invalid for the measurement error situation. Some important sources of measurement error are discussed in Cochran [3]. In sampling theory, the use of suitable auxiliary information results in considerable reduction in mean square error. Shukla et al. ([11],[12]) contributed by suggesting a mean estimator as well as class(es) of factor-type estimator(s) in the presence of measurement error. Singh and Karpe [13] have paid attention towards the estimation of population mean of the study variable y using the auxiliary information in presence of measurement error. Singh and Karpe [14] considered the problem of estimation of population variance S_y^2 under the assumptions: (i) when the study variable y is measured without error and auxiliary variable x is affected by error with known error variance S_v^2, (ii) when the study variable y is affected by error with known error variance S_u^2 and the auxiliary variable x is free from error. Furthermore, under the assumption of measurement error in study variable$y$, Singh and Karpe [15] paid attention towards the estimation of finite population variance S_y^2. Bhushan et al. [2] proposed two-phase generalized class of regression-type estimators using auxiliary information. Diana and Giordan [4] have proposed a family of estimators for the population variance S_y^2 by assuming error in both variables $y$and x under the regression approach. In practical application, let a psychiatrist wants to estimate the population variance of level of pathology in certain class of patients which depends upon the thinking disturbance , aggressive attitude, number of major miss-haps in life, ect.

In literature, the work on estimation of finite population variance using multi-auxiliary variables under multi-phase sampling is lacking especially when the study variable y is assumed to be contaminated with measurement error, so the present article is one of the steps to the solution of such situation.

Proposed Set-up: In the present study, we consider the following set-up:

(1) Complete Information Case (CIC): When information on all q auxiliary variables is *known* (we use single phase sampling).
(2) Incomplete Information Case (IIC): When information on some auxiliary variables is *known*.
(3) No Information Case (NIC): When information on all q auxiliary variables is *unknown*.

In Section 2, the symbols and notations used in this article are discussed. Section 3 presents the generalized regression-type variance estimator based on complete information of the multi-auxiliary variables about population variance, when the study variable is contaminated with measurement error. Section 4 present the generalized regression-type variance estimator when population variance of few auxiliary variables is known. In

Section 5, the generalized regression-type variance estimator is proposed when the population variance of all multi-auxiliary variable is unknown. Sections 6, 7 and 8 present the efficiency comparison, numerical analysis and concluding remarks respectively.

2. Symbols and notations

Let $U = \{1, 2, ..., j, ..., N\}$ be a finite population of N distinct and identifiable units. Let y and x_i $(i = 1, 2, ..., r, r+1, ..., q)$, be the study and the q auxiliary variables respectively, taking values y_j and x_{ij} for the j-th population unit. We are interested in estimating the finite population variance $\left(S_y^2\right)$ under multi-phase sampling schemes. Specifically we assume that a preliminary large sample $n_{(1)}$ is drawn with simple random sampling without replacement (SRSWOR) from a population and information on the auxiliary variable x_1 is taken. In second phase, a relatively small sample of size $n_{(2)}$ is drawn from $n_{(1)}$ $\left(n_{(2)} < n_{(1)}\right)$ and information on both auxiliary variables x_1 and x_2 is taken. This procedure goes up to the last phase when the smallest sample of size $n_{(q+1)}$ $\left(n_{(q+1)} < n_{(q)} < .... < n_{(1)}\right)$ is drawn. At this phase, all the q auxiliary variables as well as the variable of interest y are also observed. According to assumption, the measurement error is present in the variable of interest y denoted by $y^{\otimes}$ with known variance S_u^2. Moreover, let $S_{x_i}^2$ and $s_{x_{i(l)}}^2$denote the known population variance and sample variance of the i-th auxiliary variable $(i = 1, 2, ..., r, r+1, ..., q)$ at l-th phase $(l = i, ..., q, q+1)$, respectively. We limit our numerical study to two-phase sampling using *three* auxiliary variables.

The observational or measurement errors are defined as
$u_{j(l)} = y_{j(l)}^{\otimes} - y_{j(l)}$ and $v_{ij(l)} = x_{ij(l)}^{\otimes} - x_{ij(l)}$ $(i = 1, 2, ..., r, r+1, ..., q)$,
where $u_{j(l)}$ and $v_{ij(l)}$are assumed to be stochastic with zero mean and constant variances S_u^2 and $S_{v_i}^2$. As $\bar{y}_{(l)}$ and $\bar{x}_{i(l)}$ are unbiased estimators but $s_{y_{(l)}}^2$and $s_{x_{i(l)}}^2$ are biased estimators.
Let $\left(\bar{Y}, \bar{X}_i\right)$ and $\left(S_y^2, S_{x_i}^2\right)$ be the population means and population variances of the true values of $y_{j(l)}$ and $x_{ij(l)}$ respectively with corresponding sample means $\left(\bar{y}_{(l)}, \bar{x}_{i(l)}\right)$ and sample variances $\left(s_{y_{(l)}}^2, s_{x_{i(l)}}^2\right)$ at l-th phase. We know that $\bar{y}_{(l)}^{\otimes} = \frac{1}{n_{(l)}} \sum_{j=1}^{n_{(l)}} y_{i(l)}^{\otimes}$is unbiased estimator but $s_{y_{(l)}}^{\otimes 2} = \frac{1}{n_{(l)}-1} \sum_{j=1}^{n_{(l)}} \left(y_{j(l)}^{\otimes} - \bar{y}_{(l)}^{\otimes}\right)^2$ is biased estimator of S_y^2 due to measurement error. Similarly $\bar{x}_{ij(l)}^{\otimes} = \frac{1}{n_{(l)}} \sum_{j=1}^{n_{(l)}} x_{ij(l)}^{\otimes}$ is unbiased estimator but $s_{x_i}^{\otimes 2} = \frac{1}{n_{(l)}-1} \sum_{j=1}^{n_{(l)}} \left(x_{ij(l)}^{\otimes} - \bar{x}_{i(l)}^{\otimes}\right)^2$ $(i = 1, 2, ..., r, r+1, ..., q)$ is biased estimator of $S_{x_i}^2$ due to measurement error at l-th phase.
The expected values of $s_{y_{(l)}}^{\otimes 2}$ and $s_{x_i}^{\otimes 2}$are given by
$E\left(s_{y_{(l)}}^{\otimes 2}\right) = S_y^2 + S_u^2$ and $E\left(s_{x_{i(l)}}^{\otimes 2}\right) = S_{x_i}^2 + S_{v_i}^2$,
where S_u^2 and $S_{v_j}^2$are known, then the unbiased estimators of S_y^2 and $S_{x_j}^2$are
$\hat{S}_{y_{(l)}}^2 = s_{y_{(l)}}^{\otimes 2} - S_u^2$ and $\hat{S}_{x_{i(l)}}^2 = s_{x_{i(l)}}^{\otimes 2} - S_{v_i}^2 (i = 1, 2, ..., r, r+1, ..., q)$.

To obtain the properties of proposed estimators, we use the following approximations. For l-th and $(l+1)$-th phase, we define the notations as
Let $s_{y_{(l)}}^{\otimes 2} = S_y^2 \left(1 + e_{y_{(l)}}^{\otimes}\right)$, $\quad s_{y_{(l+1)}}^{\otimes 2} = S_y^2 \left(1 + e_{y_{(l+1)}}^{\otimes}\right)$, $\quad s_{y_{(l)}}^2 = S_y^2 \left(1 + e_{y_{(l)}}\right)$,
$s_{x_{i(l)}}^{\otimes 2} = S_{x_i}^2 \left(1 + e_{x_{i(l)}}^{\otimes}\right)$, $\quad s_{x_{i(l)}}^2 = S_{x_i}^2 \left(1 + e_{x_{i(l)}}\right)$, $\quad s_{x_{i(l+1)}}^2 = S_{x_i}^2 \left(1 + e_{x_{i(l+1)}}\right)$,

such that $E\left(e_{y(l)}^{\otimes 2}\right)=\varphi_{(l)}S_y^4A_{yy}^*,\ E\left(e_{y(l+1)}^{\otimes 2}\right)=\varphi_{(l+1)}S_y^4A_{yy}^*,\ E\left(e_{y(l)}^2\right)=\varphi_{(l)}S_y^4\lambda_{yy}^*,$
$E\left(e_{x_{i(l)}}^2\right)=\varphi_{(l)}S_{x_i}^4\lambda_{x_ix_i}^*,\ E\left(e_{x_{i(l+1)}}^2\right)=E\left(e_{x_{i(l)}}e_{x_{i(l+1)}}\right)=\varphi_{(l+1)}S_{x_i}^4\lambda_{x_ix_i}^*,$
$E\left(e_{y(l)}^{\otimes}e_{x_{i(l)}}\right)=\varphi_{(l)}S_y^2S_{x_i}^2\lambda_{yx_i}^*,\ E\left(e_{y(l)}^{\otimes}e_{x_{i(l+1)}}\right)=\varphi_{(l+1)}S_y^2S_{x_i}^2\lambda_{yx_i}^*,$

where $A_{yy}^*=\gamma_{2y}+\frac{2+\gamma_{2u}\left(1-\theta_y\right)^2}{\theta_y^2}$, $\theta_y=\frac{S_y^2}{S_y^2+S_u^2}$, $\gamma_{2y}=\beta_{2(y)}-3$ and $\gamma_{2u}=\beta_{2(u)}-3$, here $\beta_{2(y)}$ and $\beta_{2(u)}$ are the population co-efficients of kurtosis for the variable y and u.

Let $\lambda_{x_ix_i}^*=\lambda_{x_ix_i}-1$, $\lambda_{yx_i}^*=\lambda_{yx_i}-1$, $\mu_{yx_i}^*=\mu_{yx_i}-\mu_y\mu_{x_i}$, $\mu_y=S_y^2$, $\mu_{x_i}=S_{x_i}^2$ and $\varphi_{(l)}=\frac{1}{n_{(l)}}$.

Also $\lambda_{ts}=\frac{\mu_{22(t,s)}}{\mu_{20(t,s)}\mu_{02(t,s)}}=\frac{\mu_{ts}}{\mu_t\mu_s}$ or $t=y,x_i$ and $s=y,x_i$ $(i=1,2,....,r,r+1,....,q)$,

where $\mu_{ab(t,s)}=\frac{\sum_{i=1}^N\left(t_i-\bar{T}\right)^a\left(s_i-\bar{S}\right)^b}{N-1}$.

For $a=2$ and $b=2\Rightarrow\mu_{22(t,s)}=\frac{\sum_{i=1}^N\left(t_i-\bar{T}\right)^2\left(s_i-\bar{S}\right)^2}{N-1}$.

For $a=0$ and $b=2\Rightarrow\mu_{02(t,s)}=\frac{\sum_{i=1}^N\left(s_i-\bar{S}\right)^2}{N-1}$.

For $a=2$ and $b=0\Rightarrow\mu_{20(t,s)}=\frac{\sum_{i=1}^N\left(t_i-\bar{T}\right)^2}{N-1}$.

3. Generalized Regression-Type Estimators Using Multi-Auxiliary Variables

In this section, the estimators are formulated under the proposed setup.

3.1. Generalized regression-type estimators using multi-auxiliary variables under multi-phase sampling in the presence of ME under CIC. Let $s_{y(l)}^{\otimes 2}$ and $s_{x_{i(l)}}^2$ be the sample variances of the study variable y under measurement error and the i-th auxiliary variable $(i=1,2,...,r,r+1,...,q)$ respectively. The population variance $S_{x_i}^2$ $(i=1,2,...,r,r+1,...,q)$ of all multi-auxiliary variables is known. We consider the following generalized multi-phase regression-type estimator for population variance S_y^2 using α_i $(i=1,2,...,r,r+1,...,q)$ as unknown constants.

$$\hat{S}_{y1}^{\otimes 2}=s_{y(l)}^{\otimes 2}+\sum_{i=1}^{q}\alpha_i\left(S_{x_i}^2-s_{x_{i(l)}}^2\right). \tag{3.1}$$

In terms of e's, we have

$$\hat{S}_{y1}^{\otimes 2}-S_y^2=S_y^2e_{y(l)}^{\otimes}-\sum_{i=1}^{q}\alpha_iS_{x_i}^2e_{x_{i(l)}}. \tag{3.2}$$

Squaring (3.2) and then taking expectation, we get $MSE\left(\hat{S}_{y1}^{\otimes 2}\right)$ as

$$MSE\left(\hat{S}_{y1}^{\otimes 2}\right)=E\left(S_y^2e_{y(l)}^{\otimes}-\sum_{i=1}^{q}\alpha_iS_{x_i}^2e_{x_{i(l)}}\right)^2. \tag{3.3}$$

For optimum value of $\alpha_i=(-1)^{i+1}\frac{\left|\Lambda_{yx_i}\right|_{(y\bar{x}_q)}}{\left|\Lambda_{xx(q\times q)}\right|}$ $(i=1,2,...,q)$, the resulting minimum $MSE\left(\hat{S}_{y1}^{\otimes 2}\right)$, to first order of approximation, is given by

$$(3.4)\qquad MSE\left(\hat{S}_{y1}^{\otimes 2}\right)_{\min} = \varphi_{(l)} S_y^4 \left[A_{yy}^* - \sum_{i=1}^{q} (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|} \frac{\mu_{x_i}\lambda_{yx_i}^*}{\mu_y}\right].$$

Let $\Re^2_{s_y^2.s_{\tilde{x}_q}^2} = \sum_{i=1}^{q} (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|} \frac{\mu_{yx_i}^*}{\mu_y^2}$, then (3.4) can be written as

$$(3.5)\qquad MSE\left(\hat{S}_{y1}^{\otimes 2}\right)_{\min} = \varphi_{(l)} S_y^4 \left[A_{yy}^* - \Re^2_{s_y^2.s_{\tilde{x}_q}^2}\right].$$

Remark 3.1.1: Single-phase sampling using q auxiliary variables

For full information case using q multi-auxiliary variables, we replace l by 1, which is the case of simple random sampling. The estimator given in (3.1) becomes

$$(3.6)\qquad \hat{S}_{y1}^{\otimes 2\dagger} = s_{y_{(1)}}^{\otimes 2} + \sum_{i=1}^{q} \alpha_i \left(S_{x_i}^2 - s_{x_{i(1)}}^2\right).$$

For the optimum values of $\alpha_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|}$ $(i = 1, 2, ..., q)$, the resulting minimum $MSE\left(\hat{S}_{y1}^{\otimes 2\dagger}\right)$, to first order of approximation, is given by

$$(3.7)\qquad MSE\left(\hat{S}_{y1}^{\otimes 2\dagger}\right)_{\min} = \varphi_{(1)} S_y^4 \left[A_{yy}^* - \Re^2_{s_y^2.s_{\tilde{x}_q}^2}\right].$$

Here $|\Lambda_{yx_i}|_{(y\tilde{x}_q)}$ is the determinant of matrix of population variances of the variables $y, x_1, ..., x_q$ and $|\Lambda_{xx(q\times q)}|$ is the determinant of matrix of population variances of the variables $x_1, ..., x_q$.

Remark 3.1.2: Single-phase sampling using q auxiliary variables in the absence of measurement error

Let the observations of variable of interest y be recorded without an error. Substituting $S_u^2 = 0$ in (3.5), we get $A_{yy}^* = \lambda_{yy}^*$,

$$(3.8)\qquad MSE\left(\hat{S}_{y1}^{2}\right)_{\min} = \varphi_{(1)} S_y^4 \left[\lambda_{yy}^* - \Re^2_{s_y^2.s_{\tilde{x}_q}^2}\right].$$

3.2. Generalized regression-type estimators using multi-auxiliary information under multi-phase sampling in the presence of ME under IIC. Let $s_{x_{i(l)}}^2$ and $s_{x_{i(l+1)}}^2$ be sample variances of the auxiliary variables x_i $(i = 1, 2, \dots, r, r+1, \dots, q)$ at l-th and $(l+1)$-th phases respectively with the sample size $n_{(l)}$ and $n_{(l+1)}$ having the population variance $S_{x_i}^2$. Also $s_{y_{(l+1)}}^{\otimes 2}$ be the sample variances of the study variable y of size $n_{(l+1)}$ selected at $(l+1)$-th phase. The population variance $S_{x_i}^2$ $(i = 1, 2, ..., r, r+1, ..., q)$ on some auxiliary variables is known. We formulate the generalized regression-type estimator for the estimation of unknown finite population variance S_y^2 using α_i, δ_i $(i = 1, 2, ..., r)$ and γ_i $(i = r+1, r+2, ..., q)$ as unknown constants.

$$\begin{aligned}\hat{S}_{y2}^{\otimes 2} =& s_{y_{(l+1)}}^{\otimes 2} + \sum_{i=1}^{r} \alpha_i \left(S_{x_i}^2 - s_{x_{i(l)}}^2\right) + \sum_{i=1}^{r} \delta_i \left(S_{x_i}^2 - s_{x_{i(l+1)}}^2\right)\\ (3.9)\qquad & + \sum_{i=r+1}^{q} \gamma_i \left(s_{x_{i(l)}}^2 - s_{x_{i(l+1)}}^2\right).\end{aligned}$$

In terms of e's, we have

(3.10)
$$\hat{S}_{y2}^{\otimes 2} - S_y^2 = \left[S_y^2 e_{y(l+1)}^{\otimes} - \sum_{i=1}^{r} \alpha_i S_{x_i}^2 e_{x_{i(l)}} - \sum_{i=1}^{r} \delta_i S_{x_i}^2 e_{x_{i(l+1)}} + \sum_{i=r+1}^{q} \gamma_i S_{x_i}^2 \left(e_{x_{i(l)}} - e_{x_{i(l+1)}} \right) \right].$$

Squaring (3.10) and then taking expectation, we get

$$MSE\left(\hat{S}_{y2}^{\otimes 2}\right) = E\left[S_y^2 e_{y(l+1)}^{\otimes} - \sum_{i=1}^{r} \alpha_i S_{x_i}^2 e_{x_{i(l)}} - \sum_{i=1}^{r} \delta_i S_{x_i}^2 e_{x_{i(l+1)}} \right.$$
$$\left. + \sum_{i=r+1}^{q} \gamma_i S_{x_i}^2 \left(e_{x_{i(l)}} - e_{x_{i(l+1)}} \right) \right]^2. \tag{3.11}$$

For the optimum values of $\alpha_i = (-1)^{i+1} \left(\frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|} - \frac{|\Lambda_{yx_i}|_{(y\bar{x}_r)}}{|\Lambda_{xx(r\times r)}|} \right)$, $\delta_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_r)}}{|\Lambda_{xx(r\times r)}|}$ $(i = 1, 2, ..., r)$ and $\gamma_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|}$ $(i = r+1, r+2, ..., q)$, the resulting minimum $MSE\left(\hat{S}_{y2}^{\otimes 2}\right)$, to first order of approximation, is given by

$$MSE\left(\hat{S}_{y2}^{\otimes 2}\right)_{\min} = S_y^4 \left[\varphi_{(l+1)} A_{yy}^* - \varphi_{(l)} \sum_{i=1}^{r} (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_r)}}{|\Lambda_{xx(q\times q)}|} \frac{\mu_{yx_i}^*}{\mu_y^2} \right.$$
$$\left. + \left(\varphi_{(l)} - \varphi_{(l+1)}\right) \sum_{i=1}^{q} (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|} \frac{\mu_{yx_i}^*}{\mu_y^2} \right]. \tag{3.12}$$

Let $\Re^2_{s_y^2 \cdot s_{\bar{x}_r}^2} = \sum_{i=1}^{r} (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|} \frac{\mu_{yx_i}^*}{\mu_y^2}$. then (3.12) can be written as

$$MSE\left(\hat{S}_{y2}^{\otimes 2}\right)_{\min} = S_y^4 \left[\varphi_{(l+1)} A_{yy}^* - \varphi_{(l)} \Re^2_{s_y^2 \cdot s_{\bar{x}_r}^2} + \left(\varphi_{(l)} - \varphi_{(l+1)}\right) \Re^2_{s_y^2 \cdot s_{\bar{x}_q}^2} \right]. \tag{3.13}$$

Remark 3.2.1: Two-phase sampling using q auxiliary variables

For the case of two-phase sampling using q multi-auxiliary variables, we replace l by 1. The estimator given in (3.9) becomes

(3.14)
$$\hat{S}_{y2}^{\otimes 2\dagger} = s_{y(2)}^{\otimes 2} + \sum_{i=1}^{r} \alpha_i \left(S_{x_i}^2 - s_{x_{i(1)}}^2 \right) + \sum_{i=1}^{r} \delta_i \left(S_{x_i}^2 - s_{x_{i(2)}}^2 \right) + \sum_{i=r+1}^{q} \gamma_i \left(s_{x_{i(1)}}^2 - s_{x_{i(2)}}^2 \right).$$

For optimum values of $\alpha_i = (-1)^{i+1} \left(\frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|} - \frac{|\Lambda_{yx_i}|_{(y\bar{x}_r)}}{|\Lambda_{xx(r\times r)}|} \right)$, $\delta_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|}$ $(i = 1, 2, ..., r)$ and $\gamma_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\bar{x}_q)}}{|\Lambda_{xx(q\times q)}|}$ $(i = r+1, r+2, ..., q)$, the minimum $MSE\left(\hat{S}_{y2}^{\otimes 2\dagger}\right)$, to first order of approximation, is given by

(3.15)
$$MSE\left(\hat{S}_{y2}^{\otimes 2\dagger}\right)_{\min} = S_y^4 \left[\varphi_{(2)} \left(A_{yy}^* - \Re^{\otimes 2}_{s_y^2 \cdot s_{\bar{x}_r}^2} \right) + \left(\varphi_{(1)} - \varphi_{(2)}\right) \left(\Re^{\otimes 2}_{s_y^2 \cdot s_{\bar{x}_q}^2} - \Re^{\otimes 2}_{s_y^2 \cdot s_{\bar{x}_r}^2} \right) \right].$$

Remark 3.2.2: Two-phase sampling using q auxiliary variables in the absence of measurement error

Let the observations of variable of interest y be recorded without an error. Substituting $S_u^2 = 0$ in (3.13), we get $A_{yy}^* = \lambda_{yy}^*$, so

$$MSE\left(\hat{S}_{y2}^2\right)_{\min} = S_y^4\left[\varphi_{(2)}\lambda_{yy}^* - \varphi_{(1)}\Re^2_{s_y^2.s_{\tilde{x}_r}^2} + \left(\varphi_{(1)} - \varphi_{(2)}\right)\Re^2_{s_y^2.s_{\tilde{x}_q}^2}\right]. \tag{3.16}$$

3.3. Generalized regression-type estimators using multi-auxiliary information under multi-phase sampling in the presence of ME under NIC. Let $s_{y(l+1)}^{\otimes 2}$ and $s_{x_{i(l+1)}}^2$ be the sample variances of the study variable y under measurement error and the i-th auxiliary variable $(i = 1, 2, ..., r, r+1, ..., q)$ respectively at $(l+1)$-th phase, whereas $s_{x_{i(l)}}^2$be the sample variance of i-th auxiliary variable at l-th phase. The population variance $S_{x_i}^2$ $(i = 1, 2, ..., r, r+1, ..., q)$ of all multi-auxiliary variables is unknown. We consider the following generalized regression-type estimator for population variance S_y^2 under no information case using α_i $(i = 1, 2, ..., r, r+1, ..., q)$ as unknown constant.

$$\hat{S}_{y3}^{\otimes 2} = s_{y(l+1)}^{\otimes 2} + \sum_{i=1}^{q}\alpha_i\left(s_{x_{i(l)}}^2 - s_{x_{i(l+1)}}^2\right). \tag{3.17}$$

To the first order of approximation, we write (3.17) as

$$\hat{S}_{y3}^{\otimes 2} - S_y^2 = S_y^2 e_{y(l+1)}^{\otimes} + \sum_{i=1}^{q} S_{x_i}^2\alpha_i\left(e_{x_{i(l)}} - e_{x_{i(l+1)}}\right). \tag{3.18}$$

Squaring (3.18) and then taking expectation, we get MSE as

$$MSE\left(\hat{S}_{y3}^{\otimes 2}\right) = S_y^4 E\left[e_{y(l+1)}^{\otimes} + \sum_{i=1}^{q}\alpha_i\left(e_{x_{i(l)}} - e_{x_{i(l+1)}}\right)\right]^2. \tag{3.19}$$

For optimum value of $\alpha_i = (-1)^{i+1}\frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|}(i = 1, 2, ..., q)$,the resulting minimum $MSE\left(\hat{S}_{y3}^{\otimes 2}\right)$, to first order of approximation, is given by

$$\begin{aligned} MSE\left(\hat{S}_{y3}^{\otimes 2}\right)_{\min} =& S_y^4\left[\varphi_{(l+1)}\left(A_{yy}^* - \sum_{i=1}^{q}(-1)^{i+1}\frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|}\frac{\mu_{yx_i}^*}{\mu_y^2}\right)\right. \\ &\left. +\varphi_{(l)}\sum_{i=1}^{q}(-1)^{i+1}\frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|}\frac{\mu_{yx_i}^*}{\mu_y^2}\right]. \end{aligned} \tag{3.20}$$

We can write (3.20) in compact form as

$$MSE\left(\hat{S}_{y3}^{\otimes 2}\right)_{\min} = S_y^4\left[\varphi_{(l+1)}A_{yy}^* + \left(\varphi_{(l)} - \varphi_{(l+1)}\right)\Re^2_{s_y^2.s_{\tilde{x}_q}^2}\right]. \tag{3.21}$$

Remark 3.3.1: Two-phase sampling using q auxiliary variables

For no information case using q auxiliary variables, we replace l by 1, which is the case of two-phase sampling. The estimator given in (3.17) becomes

$$\hat{S}_{y3}^{\otimes 2\dagger} = s_{y(2)}^{\otimes 2} + \sum_{i=1}^{q}\alpha_i\left(s_{x_{i(1)}}^2 - s_{x_{i(2)}}^2\right). \tag{3.22}$$

For optimum value of $\alpha_i = (-1)^{i+1} \frac{|\Lambda_{yx_i}|_{(y\tilde{x}_q)}}{|\Lambda_{xx(q\times q)}|}$ $(i = 1, 2, ..., q)$, the resulting minimum $MSE\left(\hat{S}_{y3}^{\otimes 2\dagger}\right)$, to first order of approximation, is given by

$$(3.23) \quad MSE\left(\hat{S}_{y3}^{\otimes 2\dagger}\right)_{\min} = S_y^4\left[\varphi_{(2)}A_{yy}^* + \left(\varphi_{(1)} - \varphi_{(2)}\right)\Re_{s_y^2.s_{\tilde{x}_q}^2}^{\otimes 2}\right].$$

Remark 3.3.2: Two-phase sampling using q auxiliary variables in the absence of measurement error

Let the observations of variable of interest y be recorded without an error. Substituting $S_u^2 = 0$ in (3.21), we get $A_{yy}^* = \lambda_{yy}^*$,

$$(3.24) \quad MSE\left(\hat{S}_{y3}^{2}\right)_{\min} = S_y^4\left[\varphi_{(2)}\lambda_{yy}^* + \left(\varphi_{(1)} - \varphi_{(2)}\right)\Re_{s_y^2.s_{\tilde{x}_q}^2}^{2}\right].$$

4. Efficiency Comparison

To obtain the efficiency of proposed estimators, we compare the mean square errors of proposed multi-phase regression-type variance estimators under measurement error with the estimators assumed to be free of error.
By (3.7) and (3.8), (3.15) and (3.16), (3.23) and (3.24), it is evident that

$$(4.1) \quad \lambda_{yy}^* < A_{yy}^*.$$

Note: The Condition (4.1) is always true.

Remark: The numerical comparison is made under the efficiency conditions given above.

5. Data Description

Population 1: (Source: Mukherjee et al. [8])
The fertility data is based on 64 countries. Let y =Total fertility rate, 1980–1985, the average number of children born to a woman, using age specific fertility rates for a given year, x_1 = Child mortality, the number of deaths of children under age 5 in a year per 1000 live births, x_2 =Female literacy rate, (percent) and x_3 =Per capita GNP (in billions) in 1980.

$$N = 64, S_y^2 = 2.277, S_{x_1}^2 = 5772.670, S_{x_2}^2 = 676.409, S_{x_3}^2 = 7429417.00,$$
$$\bar{Y} = 5.549, \bar{X}_1 = 141.500, \bar{X}_2 = 51.188, \bar{X}_3 = 1401.250, S_u^2 = 1.255,$$
$$\lambda_{yy} = 2.773, \lambda_{x_1x_1} = 2.341, \lambda_{x_2x_2} = 1.631, \lambda_{x_3x_3} = 34.046,$$
$$\lambda_{yx_1} = 1.458, \lambda_{yx_2} = 1.069, \lambda_{yx_3} = 0.540, \lambda_{x_1x_2} = 1.415,$$
$$\lambda_{x_1x_3} = 1.921, \lambda_{x_2x_3} = 0.372, A_{yy} = 1.234.$$

Population 2:(Source: Gujarati [6])
The data is baesd on the demand for chicken in USA, 1960-1982. Let y = Per capita consumption of chickens in pounds, x_1 = Real disposable income per capita in dollars, x_2 = Real retail price of chicken per pound (in cents) and x_3 = Real retail price of pork per pound (in cents).

$$N = 23, S_y^2 = 54.360, S_{x_1}^2 = 381735.00, S_{x_2}^2 = 123.592, S_{x_3}^2 = 1240.710,$$
$$\bar{Y} = 39.669, \bar{X}_1 = 1035.065, \bar{X}_2 = 47.995, \bar{X}_3 = 90.400, S_u^2 = 3.987,$$
$$\lambda_{yy} = 2.03, \lambda_{x_1x_1} = 2.696, \lambda_{x_2x_2} = 1.756, \lambda_{x_3x_3} = 1.951, \lambda_{yx_1} = 2.094,$$
$$\lambda_{yx_2} = 1.541, \lambda_{yx_3} = 1.758, \lambda_{x_1x_2} = 1.997, \lambda_{x_1x_3} = 2.145, \lambda_{x_2x_3} = 1.755,$$

$A_{yy} = 1.033$.

Population 3:(Source: Vandaele [20])
The data is based on the crime rate data of USA in 1960. Let y =Number of offenses reported to police per million population, x_1 =Number of males of age 14-24 per 1000 population, x_2 =Indicator variable for southern states and x_3 =Mean number of years of schooling times 10 for persons age 25 or older.

$N = 47, S_y^2 = 1495.853, S_{x_1}^2 = 151.516, S_{x_2}^2 = 0.229, S_{x_3}^2 = 124.076,$
$\bar{Y} = 90.508, \bar{X}_1 = 137.511, \bar{X}_2 = 0.340, \bar{X}_3 = 105.406, S_u^2 = 1428.881,$
$\lambda_{yy} = 3.859, \lambda_{x_1x_1} = 3.684, \lambda_{x_2x_2} = 1.423, \lambda_{x_3x_3} = 1.896, \lambda_{yx_1} = 0.456,$
$\lambda_{yx_2} = 0.743, \lambda_{yx_3} = 1.041, \lambda_{x_1x_2} = 1.354, \lambda_{x_1x_3} = 1.356, \lambda_{x_2x_3} = 1.220,$
$A_{yy} = 2.703$.

Table 1. MSE of proposed ratio-type estimators $\hat{S}_{y1}^{\otimes 2\dagger}$, $\hat{S}_{y2}^{\otimes 2\dagger}$ and $\hat{S}_{y3}^{\otimes 2\dagger}$

Estimators	Pop.1	Pop.2	Pop.3
$\hat{S}_{y1}^{\otimes 2\dagger}$	0.408	119.840	17.957
	0.112	**2.05**	**42.106**
$\hat{S}_{y2}^{\otimes 2\dagger}$	0.709	453.212	62.581
	0.134	**14.781**	**1.408**
$\hat{S}_{y3}^{\otimes 2\dagger}$	0.554	280.270	19.282
	0.379	**2.050**	**42.106**

*The results written in Table 1 in bold format are the absolute values of measurement error.

6. Conclusion

In general, the presence of measurement error in the survey data invalidates the results. The goal of this study was to show how measurement error is to be seperated in case of multi-phase sampling using multi-auxiliary variables for estimation of population variance S_y^2. The values of absolute measurement error are shown in Table 1. It is also evident that the condition (4.1) holds for all the populations. Hence, the use of proposed estimators are highly preferred in the cases of multi-phase sampling under CIC, IIC and NIC.

Acknowledgments
The authors wish to thank the editor and the anonymous referees for their suggestions which led to improvement in the earlier version of the manuscript.

References

[1] Allen, J., Singh, H.P. and Smarandache, F. *A family of estimators of population mean using multi-auxiliary information in presence of measurement error,* International Journal of Social Economics, 30 (7): 837-849, 2003.

[2] Bhushan, S., Singh, R. K. and Pandey, A. *Some generalized classes of double sampling regression type estimators using auxiliary information,* Science Vision, 11 (1): 2-6, 2011.

[3] Cochran W.G. Errors of measurement in statistics. *Technometrics*, 10: 637-666, 1968.

[4] Diana, T. and Giordan, M. *Finite population variance estimation in presence of measurement error,* Communications in Statistics-Theory and Methods, 41: 4302-431, 2012.

[5] Dorfman, A. H. *A note on variance estimation for regression estimator for double sampling*, Journal of American Statistical Association, 89: 137- 140, 1994.

[6] Gujarati, D. M. *Basic Econometrics. The McGraw-Hill Companies*, 2004.

[7] Isaki, C.T. *Variance estimation using auxiliary information*,Journal of American Statistical Association, 381: 117–123, 1983.

[8] Mukherjee, C., White, H. and Whyte, M. *Econometrics and Data Analysis for Developing Countries*, *Routledge, London*, 1998.

[9] Mukerjee R., Rao T.J. and Vijayan K. *Regression-type estimators using multiple auxiliary information*, Australian Journal of Statistics, 29 (3): 244–254, 1998.

[10] Raj, D. *On a method of using multi-auxiliary information in sample surveys*, Journal of American Statistical Association, 60: 270-277, 1965.

[11] Shukla, D., Pathak, S. and Thakur, N. S. *An estimator for mean estimation in presence of measurement error*, Research and Reviews: A Journal of Statistics, 2 (1), 1-8, 2012 a.

[12] Shukla, D., Pathak, S. and Thakur, N. S. *Class(es) of factor-type estimator(s) in presence of measurement error*, Journal of Applied Statistical Methods, 11 (2), 336- 347, 2012 b.

[13] Singh, H.P. and Karpe, N. *Ratio-product estimator for population mean in the presence of measurement errors*, Journal of Applied Statistical Science, 16: 49-64, 2008 a.

[14] Singh, H.P. and Karpe, N. *Estimation of population variance using auxiliary information in the presence of measurement errors*, Statistics in Transition, 9 (3): 443-470, 2008b.

[15] Singh, H.P. and Karpe, N. *A class of estimators using auxiliary information for estimating finite population variance in presence of measurement errors*, Communications in Statistics-Theory and Methods, 38: 734-741, 2009.

[16] Singh H. P. and Solanki, R.S. *Improved estimation of finite population variance using auxiliary information,* Communications in Statistics-Theory and Methods, 42 (15), 2718-2730, 2013 a.

[17] Singh H. P. and Solanki, R.S. *A new procedure for variance estimation in simple random sampling using auxiliary information,* Statistical Papers, 54 (2), 479-497, 2013 b.

[18] Singh, S. *Estimation of finite population variance using double sampling*, Aligarh Journal of Statistics, 11: 53-56, 1991.

[19] Solanki R. S. and Singh H. P. *An improved class of estimators for the population variance,* Model Assisted Statistics and Applications, 8 (3), 229-238, 2013.

[20] Vandaele, W. *Participation in Illegitimate Activities: Erlich Revisted. Deterrence and Incapacitation*, National Academy of Sciences. 270–335, 1978.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (4) (2016), 1307 – 1320

Equivariant estimation of common location parameter of two exponential populations using censored samples

Manas Ranjan Tripathy*†

Abstract

In this paper, we consider the problem of estimating common location parameter of two exponential populations using type-II censored samples when the scale parameters are unknown. The loss function is taken as the quadratic loss. First, we derive a class of affine equivariant estimators, which includes the maximum likelihood estimator (MLE) and the uniformly minimum variance unbiased estimator (UMVUE). A sufficient condition for improving estimators in the class is derived. Consequently, estimators dominating the MLE and the UMVUE in terms of the risk values are obtained. An example is given to compute the estimates using our result. Finally a simulation study has been carried out to numerically compare the risk functions of all the estimators.

Keywords: Brewster - Zidek technique, Equivariant estimators, Inadmissibility, Quadratic loss function, Relative risk performances, Type-II censoring.

2000 AMS Classification: 62F10, 62C15

Received : 14.09.2014 *Accepted* : 30.06.2015 *Doi* : 10.15672/HJMS.20157411600

1. Introduction

Suppose $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(r)}$ $(2 \leq r \leq m)$ and $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(s)}$ $(2 \leq s \leq n)$ be the ordered observations taken from two exponential populations $Ex(\mu, \sigma_1)$ and $Ex(\mu, \sigma_2)$ respectively. Here $Ex(\mu, \sigma_i)$ denotes the exponential distribution with density function

$$f(t, \mu, \sigma_i) = \frac{1}{\sigma_i} \exp\{-(t-\mu)/\sigma_i\}, \quad t > \mu, \sigma_i > 0, \\ -\infty < \mu < \infty; \; i = 1, 2. \tag{1.1}$$

*Department of Mathematics, National Institute of Technology, Rourkela-769008, India, Email: manasmath@yahoo.co.in
†Corresponding Author.

The problem is to estimate the common location parameter μ (minimum guarantee time) when the scale parameters σ_1, σ_2 (residual life times) are unknown, with respect to the loss function,

$$L(d, \underline{\alpha}) = \left(\frac{d-\mu}{\sigma_1}\right)^2, \tag{1.2}$$

where d is an estimate for μ and $\underline{\alpha} = (\mu, \sigma_1, \sigma_2)$.

The model (1.1) under consideration arises naturally in the study of reliability, life testing and survival analysis and has applications in industry, engineering, business and social science. For example, two brands of electronic devices having $m(\geq 2)$ and $n(\geq 2)$ number of units respectively put for a life testing experiment. Due to some constraints (may be time or cost) the experimenter could able to observe only the $r(\leq m)$ and $s(\leq n)$ failure times respectively. It is assumed that, the life times of each units are random and follow exponential distributions having same minimum guarantee time. The problem we consider, comes under the umbrella of estimation problems "estimation of parameters of a distribution function using censored samples". For some more examples on related model one may refer to Suresh [13]. Basically, the censoring schemes available are type-I (number of failures are random), type-II (censoring time is random) and random censoring (both may be random) or some modifications of these. We consider the conventional type-II right censoring sampling scheme which is a particular case of progressive type-II censoring scheme. For some results on estimation of parameters of exponential distributions using various such conventional censoring schemes one may refer to Lawless [9] and Johnson et al. [8]. For some reference on estimation of parameters using progressive type-II censored samples one may refer to Chandrasekar et al. [4], Madi [12] and Wang et al. [14] and the references cited there in. Some applications of these types of models have been discussed in Balakrishnan and Aggarwala [1] and Balakrishnan and Cramer [2]. It is very surprising to see in the literature that, a very little attention has been paid for estimation of a common mean/location (or in general common parameter) when incomplete data (censored samples) are available from the population. In that regard, Chiou and Cohen [5] considered the model in (1.1) under type-II censoring scheme and estimate the common location parameter μ, when the scale parameters are unknown. They obtained the maximum likelihood estimate (MLE) and the uniformly minimum variance unbiased estimate (UMVUE) for μ. They have also generalized the results to $k = 3$ exponential populations. Elfessi and Pal [6] considered the problem of estimation of common scale parameter of several exponential populations under type-II censoring scheme. They provided stein type testimators for the common scale parameter and used this to construct estimators for the location parameters.

In the case of full sample (that is $r = m$ and $s = n$) probably, Ghosh and Razmpour [7] was the first to consider the problem of estimation of μ. They obtained the MLE, a modification to the MLE (MMLE) and the UMVUE for μ. Asymptotic and numerical comparisons of these estimators were done in terms of bias and mean squared error. Their simulation study shows that, the MLE is dominated by both the MMLE and the UMVUE. Jin and Pal [11] considered the problem of estimation of common location parameter of several exponential populations and suggested a class of estimators which dominates the MLE under a class of convex loss functions. For some early results on the estimation of common location of exponential populations we refer to Jin and Crouse [10] and the references there in.

In this article, we consider the model in (1.1) under the conventional type-II censoring, which was considered earlier by Chiou and Cohen [5] and estimated the common location parameter μ with respect to a quadratic loss function. The aim of the present work is twofold, one is to propose a wide class of estimators which include the MLE, the MMLE (we propose in next section) and the UMVUE for μ. Secondly, we derive a sufficient condition that helps in obtaining estimators which dominate estimators belonging to this class. The rest of the work is organized as follows. In Section 2, we present the model and discuss some basic results. In Section 3, a general class of estimators has been proposed and sufficient conditions for improving estimators in the class has been derived. This class contains the MLE, MMLE and the UMVUE for μ. Using the results of section 3, estimators dominating the MLE and the UMVUE have been obtained. In Section 4, a massive simulation study has been carried out to numerically compare the risk performances of all these estimators.

2. Some Basic Results

In this section, we discus the model and derive some basic estimators such as the MLE, a modification to the MLE (MMLE) and the uniformly minimum variance unbiased estimator (UMVUE) for the common location parameter μ, when the scale parameters are unknown.

Specifically, let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(r)}$, $(2 \leq r \leq m)$ be the r ordered observations taken from a random sample of size m which follows $Ex(\mu, \sigma_1)$ as in (1.1). Similarly, let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(s)}$, $(2 \leq s \leq n)$ be the s ordered observations from a random sample of size n which follows $Ex(\mu, \sigma_2)$ as in (1.1). We assume that the two random samples drawn are stochastically independent. The joint probability density function of $\underline{X}_r = (X_{(1)}, X_{(2)}, \cdots, X_{(r)})$ and $\underline{Y}_s = (Y_{(1)}, Y_{(2)}, \cdots, Y_{(s)})$ is given by

$$f(\underline{x}_r, \underline{y}_s) = M \exp\Bigg\{ -\frac{\sum_{i=1}^{r}(x_{(i)} - \mu) + (m-r)(x_{(r)} - \mu)}{\sigma_1} \tag{2.1}$$
$$-\frac{\sum_{j=1}^{s}(y_{(j)} - \mu) + (n-s)(y_{(s)} - \mu)}{\sigma_2}\Bigg\},$$

where, $\mu \leq x_{(1)} \leq x_{(2)} \cdots \leq x_{(r)}$; $\mu \leq y_{(1)} \leq y_{(2)} \cdots \leq y_{(s)}$; $-\infty < \mu < \infty$, $\sigma_1 > 0$, $\sigma_2 > 0$ and

$$M = \frac{m(m-1)\cdots(m-r+1)n(n-1)\cdots(n-s+1)}{\sigma_1^r \sigma_2^s}.$$

It can be observed that, the maximum likelihood estimator (MLE) of μ is $Z = \min(X_{(1)}, Y_{(1)}) = d_{ML}$ (say). The MLEs of both σ_1 and σ_2 can be obtained by differentiating the log-likelihood function with respect to σ_i $(i = 1, 2)$ and equating to zero. These are obtained as,

$$\hat{\sigma}_1 = \frac{\sum_{i=1}^{r}(X_{(i)} - Z) + (m-r)(X_{(r)} - Z)}{r},$$
$$\hat{\sigma}_2 = \frac{\sum_{j=1}^{s}(Y_{(j)} - Z) + (n-s)(Y_{(s)} - Z)}{s}.$$

Let us introduce the new random variables

$$U_1 = \frac{\sum_{i=1}^{r} X_{(i)} + (m-r)X_{(r)}}{m}, \quad \text{and} \quad U_2 = \frac{\sum_{j=1}^{s} Y_{(j)} + (n-s)Y_{(s)}}{n}.$$

For our model, a sufficient statistic is (U_1, U_2, Z) (see Chiou and Cohen [5]). Further, the joint probability density function of (U_1, U_2, Z) is given by,

$$f_{U_1,U_2,Z}(u_1, u_2, z) = K\Big[\frac{(u_1 - z)^{r-2}(u_2 - z)^{s-1}}{\Gamma s \Gamma(r-1)} + \frac{(u_1 - z)^{r-1}(u_2 - z)^{s-2}}{\Gamma r \Gamma(s-1)}\Big]$$
$$\exp\Bigg\{ -\frac{m(u_1 - \mu)}{\sigma_1} - \frac{n(u_2 - \mu)}{\sigma_2}\Bigg\}, \quad u_1 > x_{(1)}, u_2 > y_{(1)}, z > \mu \tag{2.2}$$

where

$$K = \frac{m^r n^s}{\sigma_1^r \sigma_2^s}.$$

It should be noted that the details of derivation of the joint probability density function of (U_1, U_2, Z) has been omitted here for brevity, however for equal sample sizes one may refer to Chiou and Cohen [5].

The probability density function of Z is given by

$$f_Z(z) = p \exp\{-p(z-\mu)\}, \quad z > \mu, \tag{2.3}$$

where $p = \frac{m}{\sigma_1} + \frac{n}{\sigma_2}$. It can be noted that, $E(Z) = \mu + \frac{1}{p}$. Motivated by Ghosh and Razmpour [7], we propose a modification to the MLE d_{ML} as,

$$d_{MM} = Z - \frac{1}{\hat{p}}, \tag{2.4}$$

where, we have the MLE for p as $\hat{p} = \frac{m}{\hat{\sigma}_1} + \frac{n}{\hat{\sigma}_2}$. It can be further noticed that the statistics $(U_1 - Z, U_2 - Z)$ and Z are independent. Using the complete and sufficient statistic $(U_1 - Z, U_2 - Z, Z)$, it is easy to observe that the UMVUE of μ is given by,

$$d_{MV} = Z - \frac{(U_1 - Z)(U_2 - Z)}{(r-1)(U_2 - Z) + (s-1)(U_1 - Z)}, \tag{2.5}$$

(see Chiou and Cohen [5] for $m = n$ and $r = s$).

In the next section, we prove a general inadmissibility result for affine equivariant class of estimators and as a consequence, estimators dominating the MLE d_{ML} and the UMVUE d_{MV} in terms of risk values have been obtained.

3. A Sufficient Condition for Improving Equivariant Estimators

In this section, we introduce the concept of invariance to our problem and obtain some inadmissibility conditions for estimators belonging to the affine equivariant class.

Let $G = \{g_{a,b} : g_{a,b}(x) = ax + b, a > 0, -\infty < b < \infty\}$ be an affine group of transformations. Let us use the notation $V_1 = U_1 - Z$, $V_2 = U_2 - Z$. Under the transformation $g_{a,b}$, the sufficient statistics, $V_1 \to aV_1$, $V_2 \to aV_2$ and $Z \to aZ + b$. The set of parameters being transformed as $\mu \to a\mu + b$, $\sigma_i \to a\sigma_i, i = 1, 2$. In order that, the loss function (2.1) to be invariant, the decision rule d must satisfy the equation,

$$d(aZ + b, aV_1, aV_2) = ad(Z, V_1, V_2) + b.$$

Taking choice for $b = -aZ$, where $a = \frac{1}{V_1}$, and rearranging the terms, we obtain the form of an affine equivariant estimator based on (Z, V_1, V_2) for estimating μ as,

$$\begin{aligned} d(Z, V_1, V_2) &= Z + V_1 \Psi(V), \\ &= d_\Psi, \quad \text{(say)}, \end{aligned} \tag{3.1}$$

where $\Psi(V)$ is any function of $V = \frac{V_2}{V_1}$.

Further, define a function Ψ_0, for the affine equivariant estimator d_Ψ (as given in (3.1)) as,

$$\Psi_0(v) = \begin{cases} -\frac{1}{r+s}\max(v,1), & \text{if } \Psi(v) < -\frac{1}{r+s}\max(v,1) \\ \Psi(v), & \text{if } -\frac{1}{r+s}\max(v,1) \le \Psi(v) \le -\frac{1}{r+s}\min(v,1), \\ -\frac{1}{r+s}\min(v,1), & \text{if } \Psi(v) > -\frac{1}{r+s}\min(v,1) \end{cases} \tag{3.2}$$

Next, we present the main result of this section which helps in obtaining the improved estimators for μ.

3.1. Theorem. *For the affine equivariant estimator d_Ψ given in (3.1), define the function Ψ_0 as given in (3.2) and the loss function be the affine invariant loss (1.2). The estimator d_Ψ is inadmissible and is improved by d_{Ψ_0}, if there exist some values of parameters $(\mu, \sigma_1, \sigma_2)$ such that, $P(d_\Psi \neq d_{\Psi_0}) > 0$.*

Proof. The proof of this theorem can be done by using a result of Brewster and Zidek [3]. So, consider the conditional risk function of d_Ψ given $V = v$.

$$\begin{aligned} R((d_\Psi, \underline{\alpha})|V = v) &= \frac{1}{\sigma_1^2} E[(d_\Psi - \mu)^2 | V = v], \\ &= \frac{1}{\sigma_1^2} E[(Z + V_1 \Psi(V) - \mu)^2 | V = v]. \end{aligned} \tag{3.3}$$

The above risk function (3.3) is a convex function in Ψ. Hence, the minimizing value of $\Psi(V)$ for fixed values of V is obtained as,

$$\hat{\Psi}(v, \sigma_1, \sigma_2) = -\frac{1}{p} \frac{E(V_1 | V = v)}{E(V_1^2 | V = v)}. \tag{3.4}$$

To evaluate the above expression in (3.4), we have the joint probability density function of (U_1, U_2, Z) as given in (2.2). Let us use the transformation $V_1 = U_1 - Z$, $V_2 = U_2 - Z$ and $Z = Z$. The inverse transformation is given by $U_1 = V_1 + Z$, $U_2 = V_2 + Z$, and $Z = Z$. The jacobian is obtained as $J = 1$. Hence, the joint probability density function of (Z, V_1, V_2) is obtained as,

$$\begin{aligned} f_{V_1,V_2,Z}(v_1, v_2, z) = \frac{m^r n^s}{\sigma_1^r \sigma_2^s} \Big[\frac{v_1^{r-1} v_2^{s-2}}{\Gamma r \Gamma(s-1)} + \frac{v_1^{r-2} v_2^{s-1}}{\Gamma s \Gamma(r-1)} \Big] \exp\{ -\frac{m}{\sigma_1}(v_1 + z - \mu) - \\ \frac{n}{\sigma_2}(v_2 + z - \mu)\}, \\ v_1 > 0, v_2 > 0, z > \mu. \end{aligned}$$

Using the independence of (V_1, V_2) and Z one can easily write the joint probability density function of (V_1, V_2) and is given by,

$$\begin{aligned} f_{V_1,V_2}(v_1, v_2) = \frac{m^r n^s p^{-1}}{\sigma_1^r \sigma_2^s} \Big[\frac{v_1^{r-1} v_2^{s-2}}{\Gamma r \Gamma(s-1)} + \frac{v_1^{r-2} v_2^{s-1}}{\Gamma s \Gamma(r-1)} \Big] \exp\Big\{ -\frac{m}{\sigma_1} v_1 - \frac{n}{\sigma_2} v_2 \Big\}, \\ v_1 > 0, v_2 > 0. \end{aligned}$$

We need to calculate the conditional density of V_1 given V. Let us use the transformation, $V = \frac{V_2}{V_1}$, $V_1 = V_1$. The inverse transformation is given by $V_2 = V V_1$, $V_1 = V_1$. The jacobian of this transformation is obtained as V_1. Hence the joint probability density function of (V_1, V) is obtained as,

$$\begin{aligned} f_{V_1,V}(v_1, v) = \frac{m^r n^s p^{-1}}{\sigma_1^r \sigma_2^s} \Big[\frac{v_1^{r+s-2} v^{s-2}}{\Gamma r \Gamma(s-1)} + \frac{v_1^{r+s-2} v^{s-1}}{\Gamma s \Gamma(r-1)} \Big] \exp\Big\{ -\frac{m}{\sigma_1} v_1 - \frac{n}{\sigma_2} v v_1 \Big\}, \\ v_1 > 0, v > 0. \end{aligned}$$

The marginal density function of V is given by

$$\begin{aligned} f_V(v) = \frac{m^r n^s p^{-1} \Gamma(r+s-1)}{\sigma_1^r \sigma_2^s} \Big(\frac{m}{\sigma_1} + \frac{n}{\sigma_2} v \Big)^{1-r-s} \Big[\frac{v^{s-2}}{\Gamma r \Gamma(s-1)} + \frac{v^{s-1}}{\Gamma s \Gamma(r-1)} \Big], \\ v > 0. \end{aligned}$$

It is easy to observe that, the conditional probability density function of V_1 given $V = v$, is a gamma distribution with shape parameter $r + s - 1$ and scale parameter $\frac{\sigma_1 \sigma_2}{m\sigma_2 + n\sigma_1 v}$.

Here the gamma probability density function with a shape parameter α and a scale parameter β is defined as,

$$g(x,\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad x > 0, \alpha > 0, \beta > 0.$$

So, the conditional expectations are calculated and obtained as

$$E(V_1|V=v) = \frac{(r+s-1)\sigma_1\sigma_2}{m\sigma_2 + n\sigma_1 v}, \tag{3.5}$$

and

$$E(V_1^2|V=v) = (r+s-1)(r+s)\Big(\frac{\sigma_1\sigma_2}{m\sigma_2 + n\sigma_1 v}\Big)^2. \tag{3.6}$$

Substituting these conditional expectations from (3.5) and (3.6) in (3.4), and simplifying, we have the minimizing choice of $\hat{\Psi}(\tau, v)$ for fixed v as,

$$\hat{\Psi}(\tau, v) = -\frac{m + n\tau v}{(r+s)(m+n\tau)}, \tag{3.7}$$

where $\tau = \frac{\sigma_1}{\sigma_2} > 0$, and $v > 0$.

In order to apply the Brewster Zidek orbit-by-orbit improvement technique (see Brewster and Zidek [3]), we need to find the supremum and infimum of $\hat{\Psi}(\tau, v)$ with respect to τ for fixed v. Let $h(\tau) = -\frac{m+n\tau v}{m+n\tau}$. It can be easily seen that, $h(\tau)$ is an increasing function in τ if and only if $v < 1$ and decreasing if and only if $v \geq 1$. We consider two separate cases for obtaining the supremum and infimum of $\hat{\Psi}(v)$, depending upon $v < 1$ or $v \geq 1$.

Case-I: $v < 1$. In this case, the supremum and infimum of the function $\Psi(\tau, v)$ for fixed values of v, are obtained as,

$$\sup_{\tau>0} \hat{\Psi}(v,\tau) = -\frac{v}{r+s}, \quad \text{and} \quad \inf_{\tau>0} \hat{\Psi}(v,\tau) = -\frac{1}{r+s}.$$

Case-II: $v \geq 1$. For this case we have the supremum and infimum of $\Psi(\tau, v)$ as,

$$\sup_{\tau>0} \hat{\Psi}(v,\tau) = -\frac{1}{r+s}, \quad \text{and} \quad \inf_{\tau>0} \hat{\Psi}(v,\tau) = -\frac{v}{r+s}.$$

Utilizing the results from Case-I and II, we can easily define the function $\Psi_0(v)$ as in (3.2). Now applying the orbit by orbit improvement technique of [3] (see Theorem 3.1.1 in Brewster and Zidek [3]), the proof follows. $\square$

Next, our target is to apply the results of Theorem 3.1, and provide improved estimators for μ which will perform better than the MLE d_{ML} and the UMVUE d_{MV} in terms of risk values. The class considered above contains the MLE d_{ML}, the modified MLE d_{MM} and the UMVUE d_{MV}. Hence, expressing d_{ML} and d_{MV} in the form (3.1), we have

$$\begin{aligned} d_{ML} &= Z + V_1\Psi_{ML}(V), \text{where } \ \Psi_{ML}(V) = 0, \\ d_{MV} &= Z + V_1\Psi_{MV}(V), \text{where } \ \Psi_{MV}(V) = -\frac{V}{(r-1)V + (s-1)}. \end{aligned}$$

Let us define the new estimators for μ as,

$$d_{MLI} = \begin{cases} Z - \frac{V_2}{r+s}, & \text{if } V_1 > V_2, \\ Z - \frac{V_1}{r+s}, & \text{if } V_1 \leq V_2 \end{cases} \tag{3.8}$$

and

$$d_{MVI} = \begin{cases} Z - \frac{V_1}{r+s}\max(V,1), & \text{if } \Psi_{MV}(V) < -\frac{1}{r+s}\max(V,1), \\ d_{MV}, & \text{otherwise.} \end{cases} \tag{3.9}$$

Next, we present the result in the form of a theorem, regarding improvement over the MLE d_{ML} and the UMVUE d_{MV}, for estimating μ.

3.2. Theorem. *Let the loss function be the quadratic loss as in (1.2).*

- *The estimator d_{ML} (MLE) is inadmissible and is improved by d_{MLI}.*
- *The estimator d_{MVI} improves upon d_{MV} (UMVUE), if there exists some values of parameters $(\mu, \sigma_1, \sigma_2)$ such that, $P(d_{MV} \neq d_{MVI}) > 0$.*

Proof. The proof follows by an application of Theorem 3.1. The choice of $\Psi_{ML} = 0 > -\frac{v}{r+s}$ (when $v < 1$) and also $\Psi_{ML} = 0 > -\frac{1}{r+s}$ (when $v \geq 1$). Hence, replacing these choices by their respective supremum values, we get the estimator defined in (3.8), which has smaller risk values than d_{ML} by an application of Theorem 3.1. Also we note that, for estimator d_{MV}, the choice $P(\Psi_{MV}(V) < -\frac{1}{r+s}) > 0$ (when $v < 1$) and $P(\Psi_{MV}(V) < -\frac{v}{r+s}) > 0$ (when $v \geq 1$). Replacing $\Psi_{MV}(V)$ by these extreme values in d_{MV} we get the required estimator d_{MVI} as given in (3.9), which has smaller risk values than d_{MV}. □

Let us define $\Psi_1 = -\frac{1}{r+s}\max(v,1)$ and $\Psi_2 = -\frac{1}{r+s}\min(v,1)$.

3.1. Remark. Though the estimator d_{MM} is a member of the class considered in (3.1) (we can write $d_{MM} = Z + V_1\Psi_{MM}(V)$, where $\Psi_{MM}(V) = -\frac{V}{rV+s}$), it can not be improved by using our result in Theorem 3.1, as it can be seen that, $P(\Psi_{MM}(V) \in [\Psi_1, \Psi_2]) = 1$.

3.2. Remark. The class of estimators $D_\Psi = \{d_\Psi : \Psi_1 \leq \Psi \leq \Psi_2\}$ form a complete class for estimating common location parameter μ when the loss is (1.2).

Next, we present an example where our model fits well and compute the estimates for the minimum guarantee time.

3.1. Example. (Simulated Data) Suppose two brands of electronic devices each having 30 units are placed for a life testing experiment. It is known that, the lifetimes (in hours) of each unit follows an exponential distribution with same minimum guarantee time. The experimenter could able to observe only 15 failures (in hours) from each brands of devices because of some constraints. The data for both the brands are obtained as,

Brand 1: 1417.70, 1458.49, 2963.76, 3969.39, 5995.44, 6939.76, 7048.85, 7768.59, 7844.87, 8824.96, 9190.34, 9321.34, 9434.04, 10793.03, 12881.22.

Brand 2: 462.71, 659.86, 1187.35, 1295.99, 1370.69, 2050.36, 2305.46, 2633.27, 3176.41, 3297.63, 3413.95, 3806.01, 4571.04, 4639.71, 6059.09.

On the basis of above data, we have computed the statistic values as $Z = 462.7199$, $T_1 = 9506.285$, and $T_2 = 3931.15$. The various estimates for μ have been computed as $d_{ML} = 462.7199$, $d_{MLI} = 331.6816$, $d_{MM} = 277.3143$, $d_{MV} = 264.0711$ and $d_{MVI} = 264.0711$. It can be noted that the condition for improvement over d_{MV} (that is $\Psi_{MV} < -\frac{1}{r+s}\max(w,1)$) is not satisfied. So, in this case we will not get improved estimator for d_{MV}. In this situation, we recommend to use d_{MM}.

4. Numerical Comparisons

In this section, we compare numerically the simulated risk values of all the estimators proposed in previous sections for estimating μ. For this purpose, we have generated 20,000 type-II censored random samples each from two exponential populations (1.1) with a common location parameter μ and different scale parameters σ_1, σ_2. The loss function is taken as (1.2). We use Monte-Carlo simulation method to compute the risk values of each estimator. The accuracy of simulation has been checked and the error is of the order of 10^{-3}. It can be easily seen that with respect to the loss (1.2), the risk functions of all the estimators are function of τ (>0) for fixed sample sizes. Though the values of τ can lie in the interval $(0, \infty)$ theoretically, to avoid simulation error we present the risk values for τ up to 4. Let us define the percentage of relative risk improvements (RRI) of all estimators with respect to the MLE as,

$$\begin{aligned} R(MLI) &= \frac{d_{ML} - d_{MLI}}{d_{ML}} \times 100, \ R(MM) = \frac{d_{ML} - d_{MM}}{d_{ML}} \times 100 \\ R(MV) &= \frac{d_{ML} - d_{MV}}{d_{ML}} \times 100, \ R(MVI) = \frac{d_{ML} - d_{MVI}}{d_{ML}} \times 100. \end{aligned}$$

Further we define the censoring factors ($CF1$, $CF2$) for both the populations as the ratio of number of observed samples to total number of samples. That is for first population $CF1 = r/m$ and for second population $CF2 = s/n$. It can be noticed that the censoring factors $CF1$ and $CF2$ always lie between 0 and 1. A massive simulation study has been carried out by considering various combinations of sample sizes. However, for illustration purpose, we present (in Table 4.1-4.3) the percentage of relative risk performances of d_{MLI}, d_{MM}, d_{MV} and d_{MVI} over d_{ML} for equal and unequal sample sizes. Specifically in Table 4.1 we present the percentage of relative risk performances for sample sizes $(16, 16)$ and $(24, 24)$. The first column gives the values of τ. Corresponding to one value of τ, there corresponds three values of relative risk performances for an estimators. These three values are obtained for $CF1 = CF2 = 0.25, 0.50, 0.75$ respectively. Similarly in Tables 4.2-4.3 the relative risk performances have been presented for unequal sample sizes. We have also plotted the graph of the RRI values of the improved estimators with respect to MLE in Figures 1 and 2 for $CF1 = CF2 = 0.25$ and $CF1 = CF2 = 0.5$ respectively. It can be seen that, as the values of $CF1$ and $CF2$ become close to 1, the amount of improvements for d_{MVI} over d_{MV} is marginal.

The following conclusions can be made from our simulation study and Table 4.1-4.3.

(i) The risk values of all the estimators are decreasing as τ increases, with respect to the loss function (1.2). Further, as τ becomes large the risk values of all the estimators converge to some constant value. The percentage of relative risk performances of each estimator with respect to MLE increases as censoring factors ($CF1$ and $CF2$) increase for fixed sample sizes.

(ii) When the sample sizes are small, and for small values of τ, the percentage of relative risk improvement for d_{MM} is maximum (near about 46%). For moderate values of τ the estimator d_{MVI} has the best percentage of relative risk performance (near about 46.5%). For large values of τ the estimator d_{MM} performs the best (near about 45%).

(iii) For moderate sample sizes, and for small values of τ the estimator d_{MM} performs the best(about 47%). When τ values are moderate the estimators d_{MM} and d_{MVI} are good competitors of each other. For large values of τ the estimator d_{MM} performs the best (about 48.5%).

(iv) For large sample sizes, and for small values of τ the estimator d_{MVI} has the best performance (48%). For moderate values of τ the estimators d_{MM} and d_{MVI} are competing each other. For large values of τ the estimator d_{MM} has the best percentage of relative risk performance (48%).

(vi) As the sample sizes increase for fixed censoring factors ($CF1$ and $CF2$) the amount of percentage of improvements of d_{MLI} over d_{ML} increases. Also the amount of improvement of d_{MVI} over d_{MV} increases as sample sizes increase. The percentage of risk improvement of d_{MVI} over d_{MV} is near about 2.5% and this value decreases as $CF1$ and $CF2$ become close to 1. The percentage of improvement for d_{MLI} over d_{ML} has been seen near about 45.5%. This validates the findings of theoretical results in Section 3.

(vii) It has been also noticed that for small and large values of τ (that is when the standard deviations vary significantly) the percentage of relative risk improvements for d_{MVI} is very marginal. A similar type of observations were made for other combinations of m, n and r, s and we omit the tables here.

(viii) Combining the facts (ii)-(iv), we recommend using d_{MM} for all sample sizes. Though the estimator performs better theoretically (around 2.5% improvement from simulation study) than d_{MM}, we do not recommend using it as it is not a smooth estimator.

5. Conclusions

In this article, we have considered the model that was earlier considered by Chiou and Cohen [5] for exponential populations. Specifically, we have considered the estimation of common location parameter μ of two exponential populations when the samples are type-II right censored in a decision theoretic approach. First we propose a broad class of estimators (which are equivariant under an affine group of transformations) for the common location parameter μ. Interestingly, this class contains the MLE and the UMVUE for μ. Then we provide a sufficient condition which may be useful for improving certain estimators in this class. Using our results of Theorem 3.1, we have obtained an estimator which dominates the MLE significantly (the percentage of relative risk improvement is between 28% to 46%). However, the improved estimator obtained for the UMVUE has marginal percentage of risk improvements. The theoretical results are well supported by a simulation study. It should be noted that, a very little attention has been given by the researchers in the recent past for the problem considered in this article. Our work revisits the problem and will definitely help the researchers to search new estimators which may work better than the usual one.

Table 4.1: Relative risk performances of different estimators for μ with CF1=CF2=0.25,0.50,0.75

$\tau\downarrow$	m=n=16 and r=s=4,8,12				m=n=24 and r=s=6,12,18			
	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$
0.25	33.53 38.07 39.74	40.45 45.33 47.04	37.89 44.88 46.92	38.28 44.90 46.92	36.95 39.62 40.54	44.43 47.01 47.75	43.73 46.90 47.70	43.78 46.90 47.70
0.50	36.14 42.32 42.81	41.51 46.77 47.06	40.22 46.60 47.03	41.34 46.70 47.07	40.33 43.48 44.71	44.93 47.42 48.30	44.38 47.29 48.21	44.63 47.30 48.22
0.75	36.95 43.37 45.40	41.19 46.40 47.66	39.84 46.38 47.56	41.28 46.58 47.64	41.06 44.76 46.91	44.75 47.12 48.74	44.71 47.21 48.73	45.08 47.25 48.76
1.00	36.91 43.05 46.24	41.13 45.57 48.07	40.54 45.32 48.03	41.84 45.61 48.11	41.36 45.68 47.17	44.61 47.54 48.40	44.24 47.62 48.37	44.80 47.67 48.41
1.25	37.04 43.31 45.66	41.11 46.07 47.77	39.91 46.08 47.74	41.40 46.22 47.81	41.69 45.66 47.00	45.15 47.69 48.60	44.95 47.67 48.62	45.35 47.71 48.63
1.50	36.94 43.30 44.08	41.47 46.55 46.73	40.52 46.41 46.71	41.78 46.56 46.76	40.81 45.85 45.93	44.64 48.58 48.43	44.51 48.57 48.48	44.84 48.61 48.49
1.75	36.55 42.78 44.06	41.29 46.62 47.46	40.00 46.38 47.43	41.09 46.54 47.46	40.19 44.19 45.06	44.35 47.39 48.22	43.84 47.29 48.24	44.23 47.31 48.25
2.00	36.31 41.34 43.12	41.67 45.79 47.04	40.28 45.66 46.94	41.48 45.80 46.96	39.85 43.58 43.79	44.55 47.44 47.76	44.17 47.35 47.84	44.45 47.36 47.84
2.25	35.56 41.15 42.19	41.56 46.22 47.11	40.84 46.08 47.14	41.65 46.13 47.14	39.05 42.34 44.53	44.24 46.86 48.81	43.88 46.76 48.74	44.10 46.77 48.74
2.50	35.49 40.18 42.22	41.43 45.91 47.26	40.01 45.92 47.12	40.72 45.95 47.12	38.77 42.26 43.32	44.60 47.50 48.38	44.42 47.42 48.35	44.56 47.42 48.35
2.75	34.69 40.42 41.84	40.89 46.30 47.48	39.52 46.05 47.33	40.17 46.08 47.34	38.29 41.04 42.76	44.12 46.71 48.28	43.57 46.64 48.24	43.66 46.64 48.24
3.00	34.69 39.39 41.43	41.25 45.78 47.70	39.37 45.62 47.66	40.01 45.65 47.66	37.65 41.31 42.30	44.02 47.33 47.90	43.49 47.19 47.78	43.61 47.19 47.78
3.25	33.76 39.01 40.58	40.31 45.49 46.69	38.73 45.08 46.46	39.18 45.12 46.46	37.41 41.10 41.26	43.81 47.42 47.63	43.07 47.30 47.62	43.16 47.30 47.62
3.50	33.60 38.88 39.85	40.59 45.76 46.48	39.24 45.45 46.35	39.73 45.46 46.35	36.49 40.77 41.16	43.66 47.60 47.74	43.39 47.44 47.70	43.46 47.45 47.70
3.75	33.37 38.39 40.37	40.29 45.22 47.54	38.57 44.71 47.41	38.95 44.72 47.41	36.26 40.13 41.13	43.28 47.33 47.95	42.58 47.26 47.88	42.65 47.26 47.88
4.00	33.25 38.84 40.19	40.55 45.89 47.51	38.77 45.35 47.35	39.17 45.36 47.35	36.25 39.91 40.77	43.72 47.02 47.89	43.12 46.82 47.82	43.17 46.82 47.82

Table 4.2: Relative risk performances of different estimators for μ with CF1=CF2=0.25,0.50,0.75

$\tau \downarrow$	m=12, n=16 and r=3,6,9;s=4,8,12				m=16,n=24 and r=4,8,12;s=6,12,18			
	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$
0.25	29.25 35.04 36.97	37.89 44.60 46.18	34.50 44.27 45.94	35.38 44.29 45.94	31.45 36.23 37.14	41.49 46.53 47.65	40.45 46.27 47.62	40.63 46.28 47.62
0.50	32.74 39.42 41.94	39.09 44.97 47.14	37.52 44.80 46.96	39.11 44.89 47.01	36.07 40.79 42.25	42.83 46.90 47.81	42.52 46.94 47.79	43.10 47.00 47.80
0.75	34.76 42.23 44.05	39.84 45.88 47.15	38.18 45.60 47.23	40.33 45.91 47.29	39.25 43.45 45.70	43.96 46.94 48.37	43.31 47.04 48.32	44.16 47.11 48.36
1.00	35.28 42.72 45.23	39.89 45.92 47.48	38.37 45.90 47.53	40.58 46.24 47.61	39.79 44.68 46.89	43.63 46.95 48.56	43.20 46.86 48.58	44.07 47.01 48.63
1.25	36.23 42.67 45.66	40.66 45.58 47.63	39.31 45.63 47.45	41.34 45.82 47.57	39.76 45.22 46.41	43.25 47.16 47.99	42.80 46.95 47.99	43.49 47.06 48.04
1.50	35.91 42.88 45.20	40.14 45.88 47.53	38.42 45.75 47.47	40.25 45.98 47.52	40.43 44.96 46.86	44.05 47.30 48.62	43.65 47.34 48.57	44.32 47.42 48.58
1.75	35.92 42.04 44.36	40.24 45.31 47.13	38.61 45.13 47.06	40.13 45.32 47.10	40.30 43.86 46.41	43.92 46.20 47.99	43.42 46.00 47.99	43.95 46.11 48.04
2.00	35.56 41.86 43.84	40.27 45.49 47.06	38.44 45.20 47.01	40.20 45.32 47.04	39.64 44.45 44.86	43.22 47.31 47.42	42.43 47.22 47.37	42.85 47.24 47.37
2.25	35.96 41.39 43.64	40.75 45.43 47.33	38.80 45.25 47.26	40.09 45.35 47.28	40.18 43.16 45.15	44.16 46.47 48.31	43.59 46.39 48.35	43.89 46.42 48.35
2.50	35.56 41.00 43.34	40.52 45.26 47.24	38.81 44.94 47.08	39.93 45.03 47.10	39.32 42.95 44.75	43.34 46.28 47.94	42.48 46.01 47.84	42.83 46.02 47.84
2.75	34.77 41.15 42.67	39.67 45.70 46.74	37.73 45.34 46.49	38.64 45.41 46.49	38.48 42.86 44.14	42.70 46.61 47.98	41.93 46.39 47.98	42.24 46.40 47.98
3.00	35.79 40.39 42.31	41.03 45.18 46.84	39.09 44.86 46.68	39.89 44.90 46.69	38.51 42.14 44.36	43.07 46.35 48.21	42.41 46.29 48.15	42.64 46.31 48.15
3.25	34.58 40.11 42.24	39.71 45.23 47.07	37.22 44.88 46.86	38.10 44.95 46.87	38.88 42.37 43.46	43.40 46.53 47.78	42.37 46.30 47.77	42.59 46.31 47.77
3.50	34.35 39.99 42.06	39.71 44.93 47.33	37.56 44.35 47.23	38.19 44.37 47.24	38.68 42.19 43.44	43.24 46.74 47.60	42.33 46.56 47.47	42.46 46.56 47.47
3.75	34.03 39.45 40.85	39.71 44.74 46.03	37.80 44.28 45.86	38.46 44.29 45.86	38.35 41.49 42.85	43.18 46.00 47.12	42.32 45.80 46.96	42.45 45.80 46.96
4.00	33.97 39.61 40.81	39.39 45.10 46.32	36.59 44.54 46.20	37.23 44.57 46.20	38.19 41.79 43.13	43.11 46.23 47.75	42.12 45.85 47.63	42.27 45.85 47.63

Table 4.3: Relative risk performances of different estimators for μ with CF1=CF2=0.25,0.50,0.75

$\tau \downarrow$	m=16, n=12 and r=4,8,12;s=3,6,9				m=24,n=16 and r=6,12,18;s=4,8,12			
	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$	$R(MLI)$	$R(MM)$	$R(MV)$	$R(MVI)$
0.25	34.11 39.67 41.39	39.83 45.15 46.83	37.83 44.67 46.62	38.35 44.67 46.63	38.28 42.23 43.01	43.11 46.88 47.51	42.09 46.57 47.36	42.21 46.58 47.36
0.50	35.57 41.87 43.98	40.03 45.42 47.17	38.08 45.15 47.14	39.47 45.32 47.14	40.31 44.64 45.64	43.99 47.61 48.10	43.14 47.54 47.99	43.66 47.59 48.00
0.75	36.05 43.15 45.30	40.50 46.10 47.41	39.30 46.11 47.40	41.18 46.31 47.47	40.32 45.22 46.67	43.70 47.34 48.21	43.15 47.20 48.16	43.80 47.33 48.19
1.00	35.94 43.05 44.87	40.60 46.17 47.09	38.90 46.18 47.08	41.29 46.51 47.21	40.17 44.81 46.24	44.11 47.11 47.89	43.82 47.00 47.90	44.59 47.16 47.94
1.25	35.13 41.50 45.11	40.18 45.10 47.91	38.90 45.03 47.81	41.13 45.38 47.98	38.67 43.57 46.13	43.17 46.53 48.47	42.73 46.40 48.41	43.59 46.55 48.46
1.50	34.86 41.50 43.66	40.35 45.10 47.15	38.80 45.03 47.14	40.80 45.38 47.21	38.29 42.46 44.54	43.47 46.55 47.79	43.09 46.57 47.79	43.73 46.66 47.80
1.75	32.93 40.78 42.76	38.83 45.89 47.15	37.01 45.80 47.06	39.12 46.06 47.13	36.77 40.84 43.34	42.72 45.61 47.87	42.33 45.53 47.82	42.98 45.60 47.85
2.00	32.43 39.94 41.20	38.96 45.70 46.73	37.52 45.42 46.72	39.24 45.57 46.77	36.54 40.41 42.46	43.45 46.29 48.01	43.26 46.25 48.02	43.82 46.30 48.04
2.25	32.32 38.61 40.91	39.25 45.06 46.86	37.39 44.98 46.69	39.06 45.06 46.73	35.80 40.29 41.55	43.27 47.02 48.12	42.83 46.98 48.15	43.30 47.01 48.16
2.50	31.72 38.31 40.09	38.94 45.50 46.68	37.08 45.25 46.59	38.51 45.36 46.60	35.01 39.40 40.80	42.83 46.87 48.27	41.95 46.76 48.31	42.47 46.79 48.32
2.75	31.51 37.16 39.62	38.91 44.36 47.24	36.52 44.06 47.22	37.72 44.10 47.23	34.06 38.19 40.20	42.22 46.03 48.16	41.42 45.84 48.11	41.78 45.86 48.11
3.00	30.73 35.56 38.97	38.31 43.53 46.56	35.25 43.24 46.28	36.63 43.33 46.29	33.13 37.67 39.28	41.87 46.34 47.99	41.10 46.24 48.01	41.40 46.25 48.01
3.25	29.81 36.33 37.90	37.87 44.46 46.33	35.43 43.97 46.25	36.51 44.00 46.25	32.66 37.27 38.66	41.73 46.28 47.66	40.70 46.04 47.58	41.01 46.04 47.58
3.50	30.06 35.70 37.02	38.35 44.30 45.54	35.35 43.80 45.39	36.40 43.83 45.39	32.39 36.51 37.52	41.80 46.21 47.06	40.94 46.17 47.03	41.11 46.18 47.03
3.75	29.39 35.29 37.56	37.91 44.30 46.46	34.99 43.79 46.20	35.90 43.80 46.20	31.94 36.09 37.51	41.65 45.99 47.30	40.53 45.77 47.19	40.73 45.78 47.19
4.00	29.17 35.16 36.46	38.14 44.41 45.71	35.76 43.83 45.53	36.55 43.85 45.53	31.38 35.89 37.51	41.08 46.07 47.38	39.80 45.74 47.18	39.91 45.75 47.18

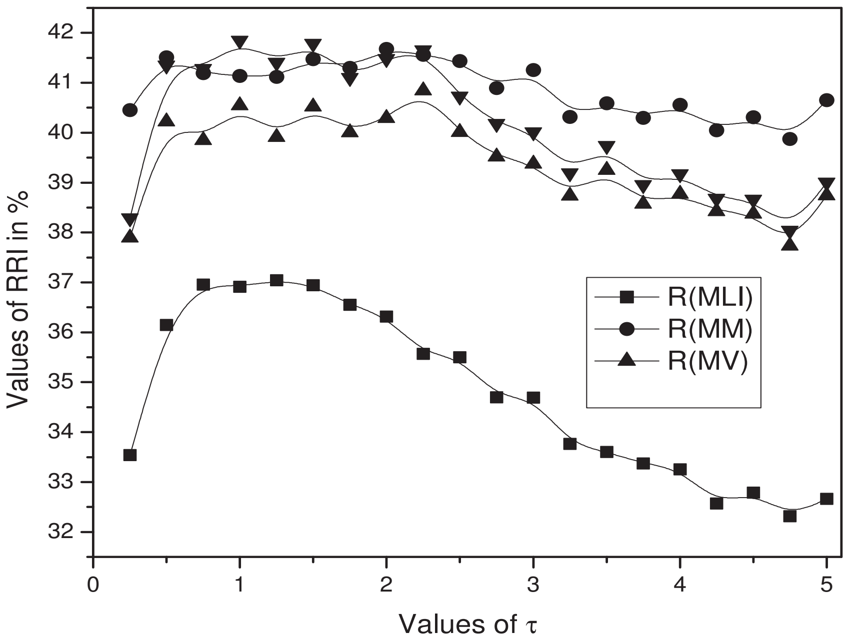


Figure 1. Comparison of RRI in % of improved estimators for μ when $m = n = 16$ and $r = s = 4$.

Acknowledgements: The author thanks two reviewers and the chief editor for their valuable comments which helped in improving the presentation of the paper. The author would also like to thank DST (SERB), [SR/FTP/MS-037/2012 dated 4/10/2013] New Delhi, India and the Director, NIT Rourkela for providing financial support.

References

[1] Balakrishnan, N. and Aggarwala, R. *Progrssive Censoring: Theory, Methods and Applications*, Birkhauser, Boston, 2000.

[2] Balakrishnan, N. and Cramer, E. *The Art of Progressive Censoring: Applications to Reliability and Quality*, Springer, New York., 2014.

[3] Brewster, J. F. and Zidek, J. V. *Improving on equivariant estimators*, The Annals of Statistics **2** (1), 21-38, 1974.

[4] Chandrasekar, B., Alexander, L. T. and Balakrishnan, N. *Equivariant estimation for parameters of exponential distributions based on type-II progressively censored samples*, Communications in Statistics- Theory ans Methods **31** (10), 1675-1686, 2002.

[5] Chiou, W. and Cohen, A. *Estimating the common location parameter of exponential distributions with censored samples*, Naval Research Logistics Quarterly **31**, 475-482, 1984.

[6] Elfessi, A. and Pal, N. *On location and scale parameters of exponential distributions with censored observations*, Communications in Statistics- Theory and Methods **20** (5-6), 1579-1592, 1991.

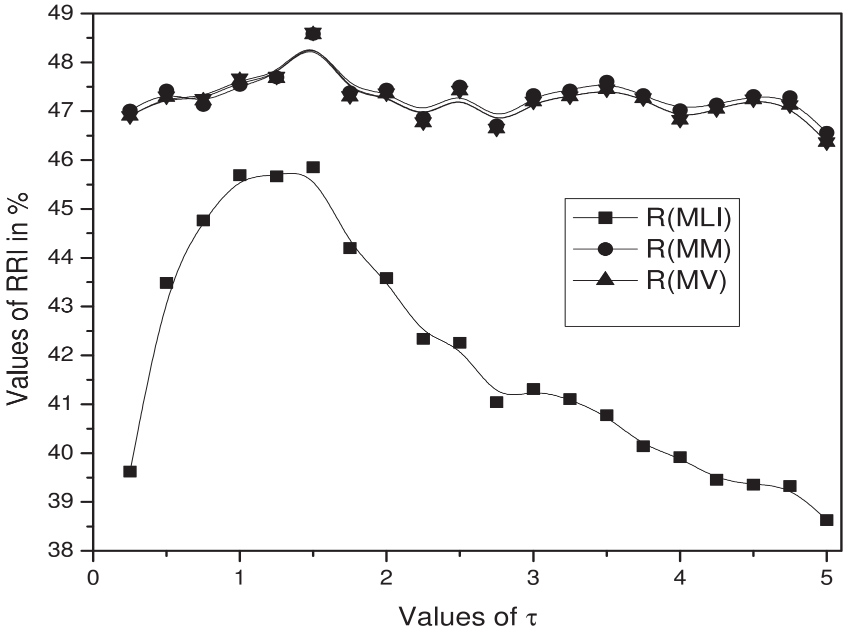


Figure 2. Comparison of RRI in % of improved estimators for μ when $m = n = 24$ and $r = s = 12$.

[7] Ghosh, M. and Razmpour, A. *Estimation of the common location parameter of several exponentials*, Sankhya: The Indian Journal of Statistics **46** Series A, Pt. 3, 383-394, 1984.

[8] Johnson, N. L., Kotz, S. and Balakrishnan, N. *Continuous Univariate Distributions, Vol-1,* John Wiley & Sons, INC, New York, 2004.

[9] Lawless, J. F. *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 1982.

[10] Jin, C. and Crouse, H. *A note on the common location parameter of several exponential populations*, Communications in Statistics-Theory and Methods **27** (11), 2777-2789, 1998.

[11] Jin, C. and Pal, N. *On common location of several exponentials under a class of convex loss functions*, Calcutta Statistical Association Bulletin **42** (167-168), 191-200, 1992.

[12] Madi, M. T. *A note on the equivariant estimation of an exponential scale using progressively censored data*, Journal of Statistical Planning and Inference **140** 1437-1440, 2010.

[13] Suresh, R. P. *Estimation of location and scale parameters in a two-parameter exponential distribution from a censored sample*, Statistical Methods **6** (1), 82-89, 2004.

[14] Wang, B. X., Yu, K. and Jones, M. C. *Inference under progressively type-II right-censored sampling for certain lifetime distributions*, Technometrics **52** (4), 453-460, 2010.

ADVICE TO AUTHORS

Hacettepe Journal of Mathematics and Statistics publishes short to medium length research papers and occasional survey articles written in English. All papers are refereed to international standards.

Address for Correspondence

Editorial Office,
Hacettepe Journal of Mathematics and Statistics,
Hacettepe University,
Faculty of Science,
Department of Mathematics,
06532 Beytepe,
Ankara,
Turkey.

E-mail: hjms@hacettepe.edu.tr
Tel: Editors: + 90 312 297 7898
+ 90 312 297 7850 / 122
Associate Editor: + 90 312 297 7880
Fax : + 90 312 299 2017
Web Page : http//www.hjms.hacettepe.edu.tr

Advise to Authors : The style of articles should be lucid but concise. The text should be preceded by a short descriptive title and an informative abstract of not more than 100 words. Keywords or phrases and the 2010 AMS Classification should also be included.The main body of the text should be divided into numbered sections with appropriate headings. Items should be numbered in the form 2.4. Lemma, 2.5. Definition. These items should be referred to in the text using the form Lemma 2.4., Definition 2.5. Figures and tables should be incorporated in the text. A numbered caption should be placed above them. References should be punctuated according to the following examples, be listed in alphabetical order according to in the (first) author' s surname, be numbered consecutively and referred to in the text by the same number enclosed in square brackets. Only recognized abbreviations of the names of journals should be used.

[1] Banaschewski, B. *Extensions of topological spaces*, Canad. Math. Bull. **7** (1), 1–22, 1964.

[2] Ehrig, H. and Herrlich, H. *The construct PRO of projection spaces: its internal structure*, in: Categorical methods in Computer Science, Lecture Notes in Computer Science **393** (Springer-Verlag, Berlin, 1989), 286–293.

[3] Hurvich, C. M. and Tsai, C. L. *Regression and time series model selection in small samples*, Biometrika **76** (2), 297–307, 1989.

[4] Papoulis, A. *Probability random variables and stochastic process* (McGraw-Hill, 1965).

INSTRUCTIONS FOR AUTHORS

Preparation of Manuscripts : Manuscripts will be typeset using the LATEX typesetting system. Authors should prepare the article using the HJMS style before submission by e-mail. The style bundle contains the styles for the article and the bibliography as well as example of formatted document as pdf file. The authors should send their paper directly to hjms@hacettepe.edu.tr with the suggestion of area editor and potential reviewers for their submission. The editorial office may not use these suggestions, but this may help to speed up the selection of appropriate reviewers.

Copyright : No manuscript should be submitted which has previously been published, or which has been simultaneously submitted for publication elsewhere. The copyright in a published article rests solely with the Faculty of Science of Hacettepe University, and the paper may not be reproduced in whole in part by any means whatsoever without prior written permission.

Notice : The content of a paper published in this Journal is the sole responsibility of the author or authors, and its publication does not imply the concurrence of the Editors or the Publisher.