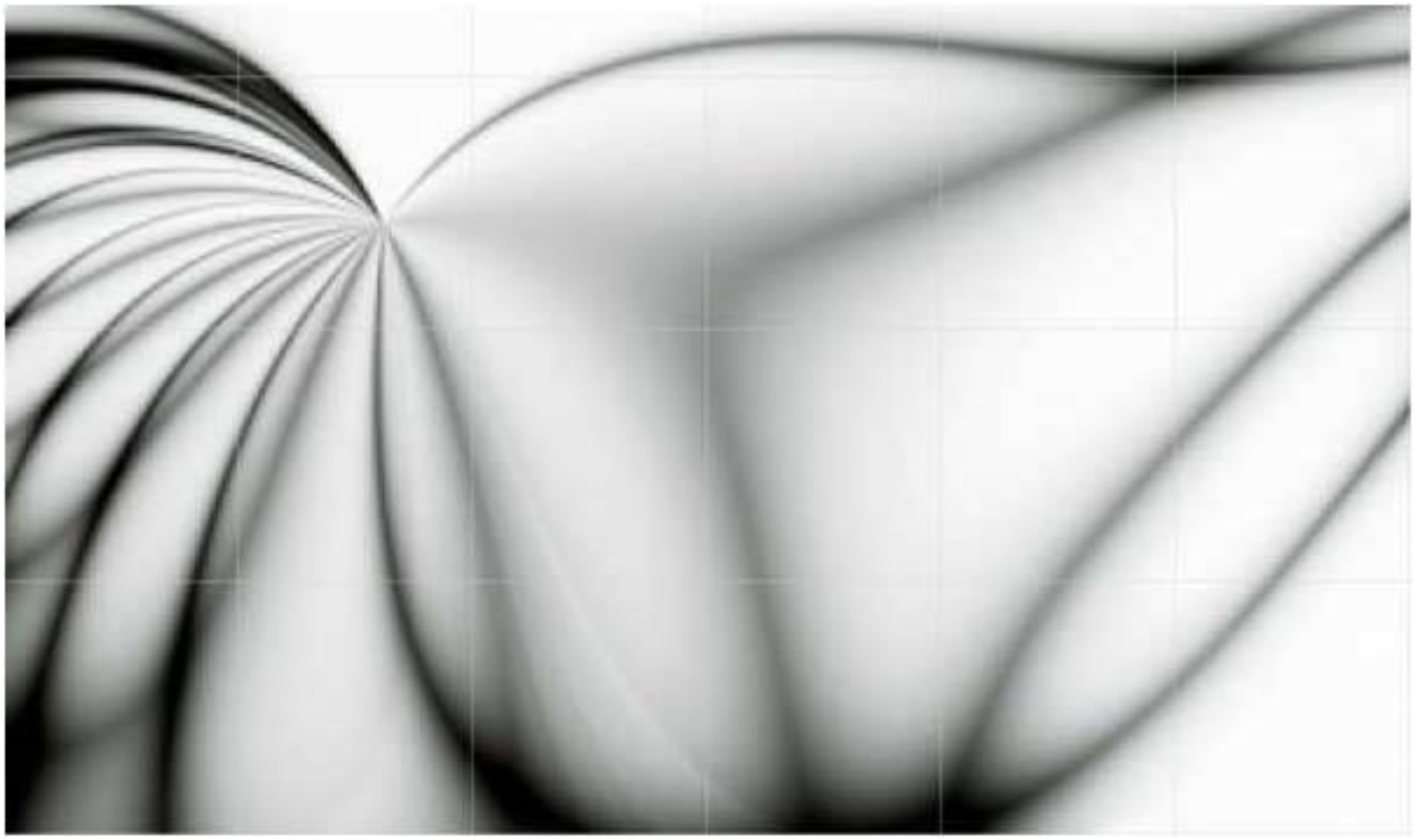
IJOCTA
An International Journal of Optimization and Control: Theories & Applications
2010

ISSN:2146-0957
eISSN:2146-5703

Volume:8 Number:2
July 2018

An International Journal of Optimization and Control:

Theories & Applications



www.ijocta.com

IJOCTA
An International Journal of Optimization and Control: Theories & Applications
2010

www.ijocta.com
info@ijocta.com

An International Journal of Optimization and Control: Theories & Applications

Volume: 8, Number: 2

July 2018

Publisher & Owner *(Yayımcı & Sahibi)*:
Prof. Dr. Ramazan YAMAN
Balikesir University, Faculty of Engineering, Department of Industrial Engineering, Cagis Campus, 10145, Balikesir, Turkey
Balıkesir Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Çağış Kampüsü, 10145, Balıkesir, Türkiye

ISSN: 2146-0957
eISSN: 2146-5703

Press *(Basımevi)*:
Bizim Dijital Matbaa (SAGE Publishing), Kazım Karabekir Street, Kültür Market, No:7 / 101-102, İskitler, Ankara, Turkey
Bizim Dijital Matbaa (SAGE Yayıncılık), Kazım Karabekir Caddesi, Kültür Çarşısı, No:7 / 101-102, İskitler, Ankara, Türkiye

Date Printed *(Basım Tarihi)*:
July 2018
Temmuz 2018

Responsible Director (Sorumlu Müdür):
Prof. Dr. Ramazan YAMAN

IJOCTA is an international, bi-annual, and peer-reviewed journal indexed/ abstracted by *(IJOCTA, yılda iki kez yayımlanan ve aşağıdaki indekslerce taranan/dizinlenen uluslararası hakemli bir dergidir)*:
Cabell's Directories, DOAJ, EBSCO Databases, JournalSeek, Google Scholar, Index Copernicus, International Abstracts in Operations Research, JournalTOCs, Mathematical Reviews (MathSciNet), ProQuest, Scopus, Ulakbim Engineering and Basic Sciences Database (Tubitak), Ulrich's Periodical Directory, and Zentralblatt Math.

crossref

We Use Plagiarism Detection iThenticate Learn More »

iThenticate and ijocta.balikesir.edu.tr are granted by Balikesir University.

An International Journal of Optimization and Control:
Theories & Applications
Volume: 8 Number: 2
July 2018

CONTENTS

An International Journal of Optimization and Control:
Theories & Applications
Volume: 8 Number: 2
July 2018

CONTENTS

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957eISSN: 2146-5703
Vol.8, No.2, pp.127-136 (2018)
https://doi.org/10.11121/ijocta.01.2018.00452

RESEARCH ARTICLE

Finite element-based hybrid techniques for advection-diffusion-reaction processes

Murat Sari, Huseyin Tunc

Department of Mathematics,Yildiz Technical University, Turkey
sarim@yildiz.edu.tr, huseyin.tunc526@std.yildiz.edu.tr

ARTICLE INFO

Article history:
Received: 31 January 2017
Accepted: 24 January 2018
Available Online: 4 February 2018

Keywords:
Advection-diffusion-reaction process
Galerkin method
Taylor-Galerkin method
B-spline
Finite element method.

AMS Classification 2010:
65M60, 65M22, 65N30, 75R50

ABSTRACT

In this paper, numerical solutions of the advection-diffusion-reaction (ADR) equation are investigated using the Galerkin, collocation and Taylor-Galerkin cubic B-spline finite element methods in strong form of spatial elements using an α-family optimization approach for time variation. The main objective of this article is to capture effective results of the finite element techniques with B-spline basis functions under the consideration of the ADR processes. All produced results are compared with the exact solution and the literature for various versions of problems including pure advection, pure diffusion, advection-diffusion, and advection-diffusion-reaction equations. It is proved that the present methods have good agreement with the exact solution and the literature.

1. Introduction

Consider the following advection-diffusion-reaction equation with given initial and boundary conditions:

$$C_t = DC_{xx} - VC_x - \theta C, t \geq 0, a \leq x \leq b \quad (1)$$

$$C(x,0) = f_0(x) \quad (2)$$

$$C(a,t) = g_0(t) \text{ or } \frac{\partial C}{\partial x} = h_0(t) \text{ at } x = a \quad (3)$$

$$C(b,t) = g_1(t) \text{ or } \frac{\partial C}{\partial x} = h_1(t) \text{ at } x = b. \quad (4)$$

Many quantities are encountered in various field of science such as mass, heat, energy, velocity, and concentration represented in the advection-diffusion-reaction (ADR) equation as the dependent variable C. The ADR equation has great importance in different areas, especially those involving fluid flow [1,2]. The ADR equation models various physical and chemical processes, as stated in the literature [3], such as heat transfer in draining film, dispersion of tracers in porous media, the spread of pollutants in rivers and stream, the dispersion of dissolved material in estuaries and coastal sea, etc. When the advection is dominant to the diffusion in the equation, the exact solutions mostly fail and thus diverge. In these cases, the effective numerical methods need to be constructed to obtain accurate and stable results of the model equation.

Nowadays, B-spline basis functions are main interest of many researchers to find out effective numerical solutions of partial differential equations [4,5]. Various versions of finite element methods have profoundly been analyzed in the literature. For instance; Least-squares B-spline finite element method was used by Dag et al. [6], a cubic B-spline collocation method was introduced by Goh [7], an upwind finite element method was organized by Ramakrishman [8], the quartic and quintic B-spline methods were used by Korkmaz and Dag [9] for their own problems. In the study of Irk et al. [3], an extended cubic B-spline collocation method was also considered. In addition to finite element-based methods, some other numerical methods were also taken into consideration in dealing with the ADR processes [1,10].

This study discovers some finite element based hybrid techniques to analyze the model problems encountered in broad range of science. To integrate the resulted system of ordinary differential equations α-family of time approximation is performed and fully discrete algebraic equations are obtained in terms of the parameters. Note that the strong form of the ADR equation (1) is accepted, as opposed to the weak form commonly used in the literature, since the strong form leads to computationally more economic and more accurate results.

All produced results are compared with the literature and exact solutions. Various test problems involving pure advection, pure diffusion, advection-diffusion and advection-diffusion-reaction are demonstrated

*Corresponding author

with quantitatively and qualitatively produced results.

2. Numerical methods

2.1 Galerkin method

To solve equation (1) with given boundary conditions (3)-(4) and initial condition (2), the Galerkin cubic B-spline finite element method is used for spatial approximation. The selection of these types of basis functions is very suitable and advantageous. The well-known advantages of using cubic B-splines are the continuity of the approximate solution and the first and second order-derivatives at all region.

The interval $[a,b]$ is partitioned into N finite elements. Each element has equal length h and element nodes are discretized as $a = x_0 < x_1 < \cdots < x_N = b$ where $x_{l+1} = x_l + h$ $(l = 0,1. \dots, N-1)$. Let φ_l be the cubic B-spline basis functions [11] as follows

$$\varphi_l(x) = \frac{1}{h^3}\begin{cases} (x-x_{l-2})^3 \\ h^3 + 3h^2(x-x_{l-1}) + 3h(x-x_{l-1})^2 - 3(x-x_{l-1})^3 \\ h^3 + 3h^2(x_{l+1}-x) + 3h(x_{l+1}-x)^2 - 3(x_{l+1}-x)^3 \\ (x_{l+2}-x)^3 \\ 0 \end{cases}, x \in \begin{cases} [x_{l-2}, x_{l-1}] \\ [x_{l-1}, x_l] \\ [x_l, x_{l+1}] \\ [x_{l+1}, x_{l+2}] \\ \text{otherwise.} \end{cases} \quad (5)$$

The corresponding cubic B-spline basis functions include the set of splines $\{\varphi_{-1}, \varphi_0, \dots, \varphi_{N+1}\}$ and the global approximation function $\tilde{C}_N(x,t)$ can be expressed as

$$\tilde{C}_N(x,t) = \sum_{l=-1}^{N+1} \beta_l(t)\varphi_l(x) \quad (6)$$

where $\beta_i(t)$ is the time part of approximation function $\tilde{c}_N(x,t)$ and is to be determined from the time approximation.

To compute element matrices, it is required to use local coordinate system considering (5) and $\sigma = x - x_l$ where $0 \le \sigma \le h$, the basis functions are expressed in the following form

$$\begin{cases} \varphi_{l-1} \\ \varphi_l \\ \varphi_{l+1} \\ \varphi_{l+2} \end{cases} = \frac{1}{h^3}\begin{cases} (h-\sigma)^3 \\ h^3 + 3h^2(h-\sigma) + 3h(h-\sigma)^2 - 3(h-\sigma)^3 \\ h^3 + 3h^2\sigma + 3h\sigma^2 - 3\sigma^3 \\ \sigma^3 \end{cases}, x \in \begin{cases} [x_{l-2}, x_{l-1}] \\ [x_{l-1}, x_l] \\ [x_l, x_{l+1}] \\ [x_{l+1}, x_{l+2}] \\ otherwise. \end{cases} \quad (7)$$

Each finite element $[x_l, x_{l+1}]$ is covered by the set of four cubic B-splines $\{\varphi_{l-1}, \varphi_l, \varphi_{l+1}, \varphi_{l+2}\}$. Table 1 shows the values of φ_l , φ_l' and φ_l'' at the end points of element $[x_l, x_{l+1}]$. Local approximation function on the element $[x_l, x_{l+1}]$ is defined as follows

$$\tilde{C}_N(x,t) = \sum_{i=l-1}^{l+2} \beta_i(t)\varphi_i(x) \quad (8)$$

Table 1. Values of approximate function and its derivatives at the end points of the element

x	x_{l-2}	x_{l-1}	x_l	x_{l+1}	x_{l+2}
φ_l	0	1	4	1	0
φ_l'	0	-3/h	0	3/h	0
φ_l''	0	$6/h^2$	$-12/h^2$	$6/h^2$	0

Values of the local approximation function $\tilde{c}_N(x,t)$ and its first two derivatives at the end points of the interval $[x_l, x_{l+1}]$ is defined in terms of time dependent quantities $\beta_l(t)$ using (8) and Table 1. The corresponding values are thus:

$$\begin{aligned} \tilde{C}_N(x_l,t) &= \beta_{l-1} + 4\beta_l + \beta_{l+1} \\ \tilde{C}_N(x_{l+1},t) &= \beta_l + 4\beta_{l+1} + \beta_{l+2} \\ \tilde{C}_N'(x_l,t) &= \frac{3}{h}(\beta_{l+1} - \beta_{l-1}) \\ \tilde{C}_N'(x_{l+1},t) &= \frac{3}{h}(\beta_{l+2} - \beta_l) \\ \tilde{C}_N''(x_l,t) &= \frac{6}{h^2}(\beta_{l+1} - 2\beta_l + \beta_{l-1}) \\ \tilde{C}_N''(x_{l+1},t) &= \frac{6}{h^2}(\beta_{l+2} - 2\beta_{l+1} + \beta_l) \end{aligned} \quad (9)$$

Now it is time to apply the Galerkin approach. By considering element $[x_l, x_{l+1}]$, let us consider the strong form of equation (1) over the interval $[x_l, x_{l+1}]$, one can then write

$$\int_{x_l}^{x_{l+1}} v\left(\frac{\partial C}{\partial t} + \mathrm{V}\frac{\partial C}{\partial x} + \theta C - D\frac{\partial^2 C}{\partial x^2}\right) dx = 0. \quad (10)$$

The test function v is selected to be equal to the cubic B-spline basis functions. This approach is known as the Galerkin approach in the finite element society. Use of (8) and local coordinate system (7) transforms equation (10) to the following relation

$$\sum_{j=l-1}^{l+2}\left[\int_0^h \varphi_i\varphi_j d\sigma\right]\frac{d\beta_j^e}{dt} + V\sum_{j=l-1}^{l+2}\left[\int_0^h \varphi_i\varphi_j' d\sigma\right]\beta_j^e + \theta\sum_{j=l-1}^{l+2}\left[\int_0^h \varphi_i\varphi_j d\sigma\right]\beta_j^e -$$

$$D\sum_{j=l-1}^{l+2}\left[\int_0^h \varphi_i\varphi_j''d\sigma\right]\beta_j^e = 0 \quad (11)$$

or in a matrix notation

$$M^e\frac{d\beta^e}{dt} + VL^e\beta^e + \theta M^e\beta^e - DK^e\beta^e = 0 \quad (12)$$

where

$$M_{ij}^e = \int_0^h \varphi_i\varphi_j d\sigma,$$
$$L_{ij}^e = \int_0^h \varphi_i\varphi_j' d\sigma,$$
$$K_{ij}^e = \int_0^h \varphi_i\varphi_j'' d\sigma, \quad (13)$$
$$\beta^e = (\beta_{l-1}, \beta_l, \beta_{l+1}, \beta_{l+2})^T$$

where $i,j = l-1, l, l+1, l+2$ for the element $[x_l, x_{l+1}]$. In equation (11); M^e, K^e and L^e are (4×4) time independent matrices. After the assembling process of each element and imposing the boundary conditions the matrix form will finally be

$$M^*\frac{d\beta}{dt} + VL^*\beta + \theta M^*\beta - DK^*\beta + R = 0 \quad (14)$$

where M^*, L^* and K^* are $(N+1)\times(N+1)$ matrices, $\beta = (\beta_0, \dots, \beta_N)^T$ is the unknown time approximation vector and R is an $((N+1)\times 1)$ time dependent vector. By considering Dirichlet boundary conditions $(3)-(4)$, R is defined as follows

$$R = (g_0P_{21}, g_0P_{31}, g_0P_{41}, 0, \dots, 0, g_1P_{(N)(N+3)}, g_1P_{(N+1)(N+3)}, g_1P_{(N+2)(N+3)})^T \quad (15)$$

$$P_{ij} = (M^{**} + VL^{**} + \theta M^{**} - DK^{**})_{ij}. \quad (16)$$

Note that M^{**}, L^{**} and K^{**} are the assembled matrices before imposing the boundary conditions. Thus, equation (14) is a system of ordinary differential equations, which is integrated using α-family of time approximation.

2.2 Collocation method

Let us reconsider expressions (5)-(9) and model equation (1) with the following collocation points

$$x_m = a + m*h,\ m = 0,1,\dots,N \quad (17)$$

where $\omega^m = [x_m, x_{m+1}]$ is the $m-th$ element. Here h and N indicate the element size and the number of total element, respectively. Then the model equation can be written as follows

$$\frac{\partial\tilde{C}_N(x,t)}{\partial t} = D\frac{\partial^2\tilde{c}_N(x,t)}{\partial x^2} - V\frac{\partial\tilde{C}_N(x,t)}{\partial x} - \theta\tilde{C}_N(x,t). \quad (18)$$

Use of expressions (9) yields the following equation

$$\dot{\beta}_{m-1} + 4\dot{\beta}_m + \dot{\beta}_{m+1} =$$
$$D\frac{6}{h^2}(\beta_{m-1} - 2\beta_m + \beta_{m+1}) - V\frac{3}{h}(-\beta_{m-1} + \beta_{m+1})$$
$$-\theta(\beta_{m-1} + 4\beta_m + \beta_{m+1}) \quad (19)$$

where $\dot{\beta}$ stands for the time differentiation. For all values of $m, N+1$ equations are obtained. In matrix notation, the corresponding equations can be rewritten as

$$M\frac{d\beta}{dt} + VL\beta + \theta M\beta - DK\beta = 0 \quad (20)$$

where M, L and K are $(N+1)\times(N+3)$ time independent matrices. After imposing Dirichlet boundary conditions $(3)-(4)$, equation (20) can be written as

$$M^*\frac{d\beta}{dt} + VL^*\beta + \theta M^*\beta - DK^*\beta + R = 0 \quad (21)$$

where M^*, L^* and $K^*$are $(N+1)\times(N+1)$ time independent matrices and $\beta = (\beta_0, \dots, \beta_N)^T$ is the unknown time approximation vector. R is an $((N+1)\times 1)$ time dependent vector and defined by

$$R = (g_0(t)P_{11}, 0, \dots, 0, g_1(t)P_{(N+1)(N+3)})^T \quad (22)$$

$$P_{ij} = (M^* + VL^* + \theta M^* - DK^*)_{ij.} \quad (23)$$

Thus, consideration of a suitable time integration method for equation (21) gives us the solution of equation (1) with conditions (2)-(4).

2.3 Taylor-Galerkin method

The third approximation method in solving equation (1) is the Taylor-Galerkin method being effective for many problems represented by differential equations. The main idea of the method is that the time approximation based on Taylor series expansion is performed before the spatial discretization. After performing the time discretization, the Galerkin method is used for the spatial approximation by utilizing B-splines basis functions (5). The order of the TGFEM schemes can be determined by the truncation error of the Taylor expansion. In this study, we prefer to use the second order TGFEM schemes for the numerical solution of equation (1). Use of the Taylor expansion of the function C with respect to t gives rise to,

$$C_t^n = \frac{C^{n+1}-C^n}{dt} - \frac{dt}{2}C_{tt}^n - O((dt)^2). \quad (24)$$

Taking derivative of equation (1) with respect to t leads to,

$$C_{tt} = (-VC_x - \theta C + DC_{xx})_t^n = -V(C_t^n)_x - \theta C_t^n + D(C_t^n)_{xx}. \quad (25)$$

At the very moment, there are several ways to approximate the second time derivative of function C. First one is that all C_t^n terms in equation (25) can be replaced by Euler time stepping, i.e. $C_t^n = \frac{C^{n+1}-C^n}{dt}$. This selection has the same as the Galerkin method. For the sake of brevity, it is preferred to use a different way to approximate the second time derivative of function C. It is noticeable that, the first order time derivative of the function, i.e. C_t, can be rewritten by using ADR equation (1) itself. Euler time stepping is used for the diffusion term of equation (25) while the

rest of the terms are dealt with the considered technique. Thus, equation (25) can be re-expressed as follows,

$$C_{tt}^n - V(-VC_x^n - \theta C^n + DC_{xx}^n)_x - \theta(-VC_x^n - \theta C^n + DC_{xx}^n +) + D\left(\frac{C^{n+1}-C^n}{dt}\right)_{xx}. \quad (26)$$

Substitution of equations (1) and (26) into equation (24) and doing some mathematical manipulations lead to the following semi-discrete system,

$$C^{n+1} - \frac{dt}{2}DC_{xx}^{n+1} = \left(1 - \theta dt + \frac{\theta^2(dt)^2}{2}\right)C^n - (Vdt - \theta V(dt)^2)C_x^n + \left(\frac{dt}{2}D + \frac{V^2dt^2}{2} - \frac{\theta V(dt)^2}{2}\right)C_{xx}^n - \frac{VD(dt)^2}{2}C_{xxx}^n. \quad (27)$$

Using the Galerkin approach for equation (27) under the consideration of equations (5)-(8) and doing some algebraic operations, one obtains the following iteration system,

$$\left[M^e - \gamma_1 \mathrm{K}^e\right]\beta^{n+1} = \left[\gamma_2 M^e - \gamma_3 L^e + \gamma_4 K^e - \gamma_5 S^e\right]\beta^n \quad (28)$$

and

$$\mathrm{K}_{ij}^e = \int_0^h \varphi_i\varphi_j''\, d\sigma,\ \mathrm{S}_{ij}^e = \int_0^h \varphi_i\varphi_j'''\, d\sigma$$

$$\gamma_1 = \frac{dt}{2}D,\ \gamma_2 = 1 - \theta dt + \frac{\theta^2(dt)^2}{2}, \quad (29)$$

$$\gamma_3 = Vdt - \theta V(dt)^2,$$

$$\gamma_4 = \frac{dt}{2}D + \frac{V^2(dt)^2}{2} - \frac{\theta V(dt)^2}{2},$$

$$\gamma_5 = \frac{VD(dt)^2}{2}$$

where $i, j = l-1, l, l+1, l+2$ for the element $[x_l, x_{l+1}]$. After assembling the procedure and imposing the Dirichlet boundary conditions, the matrix equation will then be

$$[M^* - \gamma_1 \mathrm{K}^*]\beta^{n+1} = [\gamma_2 M^* - \gamma_3 \mathrm{L}^* + \gamma_4 \mathrm{K}^* - \gamma_5 S^*]\beta^n + R^* \quad (30)$$

where M^*, L^*, K^* and S^* are $(N+1)\times(N+1)$ matrices and $n = 0,1,\dots,T$ for $T = \frac{T_{final}}{dt}$. $R^*$is an $((N+1)\times 1)$ time dependent residual vector resulting from boundary conditions. Assuming the case of Dirichlet boundary conditions $(3)-(4)$, $R^* = R_1 - R_2$ and the required matrices are defined as follows

$$\mathrm{M}_{ij}^e = \int_0^h \varphi_i\varphi_j d\sigma,\ \mathrm{L}_{ij}^e = \int_0^h \varphi_i\varphi_j'\, d\sigma$$

$$R_1 = (g_0^n P_{21}, g_0^n P_{31}, g_0^n P_{41}, 0, \dots, 0, g_1^n P_{(N)(N+3)}, g_1^n P_{(N+1)(N+3)}, g_1^n P_{(N+2)(N+3)})^T$$

$$R_2 = (g_0^{n+1} P_{21}^1, g_0^{n+1} P_{31}^1, g_0^{n+1} P_{41}^1, 0, \dots, 0, g_1^{n+1} P_{(N)(N+3)}^1, g_1^{n+1} P_{(N+1)(N+3)}^1, g_1^{n+1} P_{(N+2)(N+3)}^1)^T$$

and

$$P_{ij} = [M^{**} - \gamma_1 \mathrm{K}^{**}]_{ij.},\ P_{ij}^1 = [\gamma_2 M^* + \gamma_4 \mathrm{K}^* - \gamma_5 S^*]_{ij.}.$$

Note that M^{**} and K^{**} are the assembled matrices before prescribing the boundary conditions. Equation (30) is a recursive relation between β^n and β^{n+1}. By obtaining β^0 we can calculate the solution vector for each time step.

2.4 α-family of time approximation

To solve the ODE systems (14) and (21), α-family of time approximation is preferred since the method is easy to implement, satisfies the unconditional stability by the dependence of the selection of the parameter α and has the required accuracy. As stated in [12], the α-family of approximation can be defined as

$$\{\beta\}_{s+1} = \{\beta\}_s + dt\{\beta\}_{s+\alpha} \quad (31)$$

$$\{\beta\}_{s+\alpha} = (1-\alpha)\{\dot{\beta}\}_s + \alpha\{\dot{\beta}\}_{s+1} \quad (32)$$

or

$$dt[(1-\alpha)\{\dot{\beta}\}_s + \alpha\{\dot{\beta}\}_{s+1}] = \{\beta\}_{s+1} - \{\beta\}_s \quad (33)$$

where $0 \le \alpha \le 1$, $t_{s+1} - t_s = dt$ and $\dot{\beta}$ stands for the time differentiation. Using the same steps of the procedure given in [12], both equation (14) and equation (21) give

$$[\mathrm{M}^* + \alpha \mathrm{dt}(VL^* + \theta M^* - D\mathrm{K}^*)]\{\beta\}_{s+1} = [\mathrm{M}^* - (1-\alpha)\mathrm{dt}(VL^* + \theta M^* - D\,\mathrm{K}^*)]\{\beta\}_s + R_{s+1} - R_s \quad (34)$$

where s represents the time index. $\{\beta\}_0$ can be obtained under the consideration of initial condition (2). Then, by using recursive relation (34), the other solutions are computed.

3. Numerical illustrations

This section is devoted to numerical illustration of the various test problems for the advection-diffusion-reaction processes by considering quantitative and qualitative results. Accuracy and stability of the obtained results are figured out by demonstrating error norms and pointwise solutions. Produced results are compared with the literature and exact solutions. To evaluate error norms of the present results we prefer to use the following norm definitions,

$$E_i = \left|C_i^{exact} - C_i^{numerical}\right|,$$

$$L_\infty = max_i|C_i^{exact} - C_i^{numerical}|,$$

$$L_2 = \sqrt{h\sum_{i=1}^N |C_i^{exact} - C_i^{numerical}|^2}.$$

In the ADR equation, there are two important problem parameters need to be considered, the Peclet and the Courant numbers. Non-dimensional parameter, Courant (Cr) number, gives the fractional distance relative to the grid spacing travelled due to advection in a single time step $Cr = (Vdt)/h$. This parameter especially plays an important role when we need to determine stability conditions of the considered numerical approaches. The Peclet number is another crucial non-dimensional parameter which compares the characteristic time for dispersion and diffusion given a length scale with the characteristic time for advection, i.e. $Pe = (Vh)/D$ where the parameters are as in equation (1).

Problem 1 [6] Pure advection in an infinitely long channel:

Initially, we consider pure advection problem, i.e. $D = 0$ and $\theta = 0$. The analytic solution of the problem of interest is as follows [6]

$$C(x,t) = 10\exp(-\frac{1}{\rho^2}(x - x_0 - Vt)^2) \quad (35)$$

where ρ is the real problem parameter. This solution construct a transportation of an initial concentration of 10 height units whose peak value is at x_0 initially along an infinitely long channel as well as it maintains its own shape during the propagation. The parameters of the problem are taken to be $V = 0.5$, $x_0 = 2000\,m$ and $\rho = 264$. To compare with the literature [6], all parameters are taken to be equal. The final propagation time is $t_f = 9600\,s$ while the initial and boundary conditions are as follows

$$C(x,0) = 10\exp\left(-\frac{1}{264^2}(x - 2000)^2\right), \quad (36)$$

$$C(0,t) = 0,\ \frac{\partial C(9000,t)}{\partial x} = 0. \quad (37)$$

In Figure 1, we demonstrate the propagation of the initial pulses up to 9600 s by considering the Galerkin method for the parameters $h = 50$ and $dt = 10$. The comparison of the absolute errors produced by the Galerkin, collocation and Taylor-Galerkin methods are given in Figure 2 for the values of $V = 0.5$, $h = 100\,m$, $dt = 50\,s$ and $t = 9600\,s$. As seen in Figure 2, the Galerkin method has been seen to be more accurate than the rest of the considered methods. The error norms and peak location of the concentration are compared with the literature [6] and exact peak location in Table 2 for various values of the Courant numbers, i.e. $Cr = (Vdt)/h$. Because of the stability condition in Table 2, the Taylor-Galerkin method is not preferred to use. The method is not stable for higher Courant numbers.

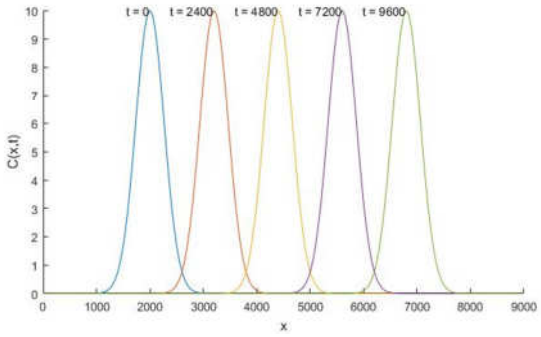


Figure 1. Propagation of initial pulse with constant wave speed $V = 0.5$ for various time values up to 9600 s.

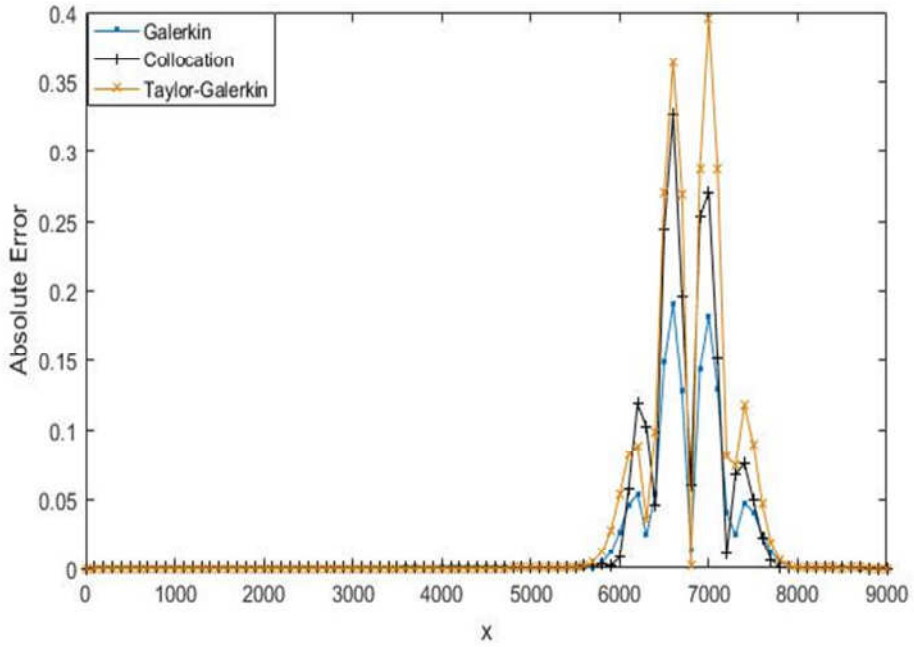


Figure 2. Comparison of the absolute error norms of various schemes with $V = 0.5$, $h = 100\,m$, $dt = 50\,s$ and $t = 9600\,s$.

Table 2. Comparison of the error norms produced with various values of the Courant number and $Pe = \infty$ in Problem 1.

		FEMQSF [6]			GFEM			CFEM			
Cr	h	Peak	L_2	L_∞	Peak	L_2	L_∞	Peak	L_2	L_∞	Exact Peak
0.125	200	8.7529	32.874	1.350	9.9216	5.5198	0.2783	9.1319	37.0409	1.2897	10
0.25	100	9.647	10.596	0.494	9.9864	4.0050	0.1899	9.9402	6.4028	0.3258	10
0.5	50	9.864	7.984	0.380	9.9865	4.0032	0.1896	9.9847	4.1390	0.1976	10
1.0	25	9.943	7.881	0.377	9.9865	4.0032	0.1896	9.9864	4.0115	0.1901	10
2.0	12.5	9.956	7.899	0.378	9.9944	4.0032	0.1896	9.9944	4.0037	0.1897	10

Problem 2 [13] Pure diffusion on a finite line

Now consider pure diffusion problem for selection of $V = 0, D = 1/\pi^2$ and $\theta = 0$ in the ADR equation (1) with the following exact solution [13],

$$C(x,t) = \exp(-t)\sin(\pi x). \quad (38)$$

The problem has homogenous Dirichlet boundary conditions and initial condition can be taken from the exact solution (38). In Table 3, we compare maximum error norms of the present schemes with the literature [13] and among each other. As seen in Section 2, the same discretized equations have been obtained for the Galerkin and the Taylor-Galerkin methods in case of pure diffusion. As realized in Table 3, the present Galerkin scheme produces better accuracy comparison to the literature [13] and the current collocation scheme. Thus the computed results have been seen to represent the related physical problem. Yet, comparison of absolute errors has been seen both qualitatively and quantitatively in Figure 3 for $h = 0.01\ m$, $dt = 0.0001\ s$ and $t = 1\ s$. Figure 4 illustrates the diffusion process of the initial concentration with the diffusion constant $D = 1/\pi^2$ by using the Galerkin approach for $h = 0.01\ m$ and $dt = 0.0001\ s$.

Table 3. Comparison of the error norms produced with various values of the Peclet number and $Cr = 0$ in Problem 2.

		CN [13]	GFEM		CFEM		TGFEM	
$h = dt$	Pe	L_∞	L_∞	L_2	L_∞	L_2	L_∞	L_2
0.2	1.973	1.1E-2	1.1E-3	8.7E-4	1.2E-2	9.4E-3	1.1E-3	8.7E-4
0.1	0.986	2.7E-3	3.0E-6	2.1E-4	3.3E-3	2.3E-3	3.0E-6	2.1E-4
0.05	0.493	6.8E-4	7.6E-5	5.4E-5	8.3E-4	5.8E-4	7.6E-5	5.4E-5
0.025	0.246	1.7E-4	1.9E-5	1.3E-5	2.0E-4	1.4E-4	1.9E-5	1,3E-5
0.0125	0.123	4.2E-5	4.7E-6	3.3E-6	5.2E-5	3.6E-5	4.7E-6	3,3E-6
0.00625	0.061	1.1E-5	1.1E-6	8.4E-7	1.3E-5	9.2E-6	1.1E-6	8,4E-7

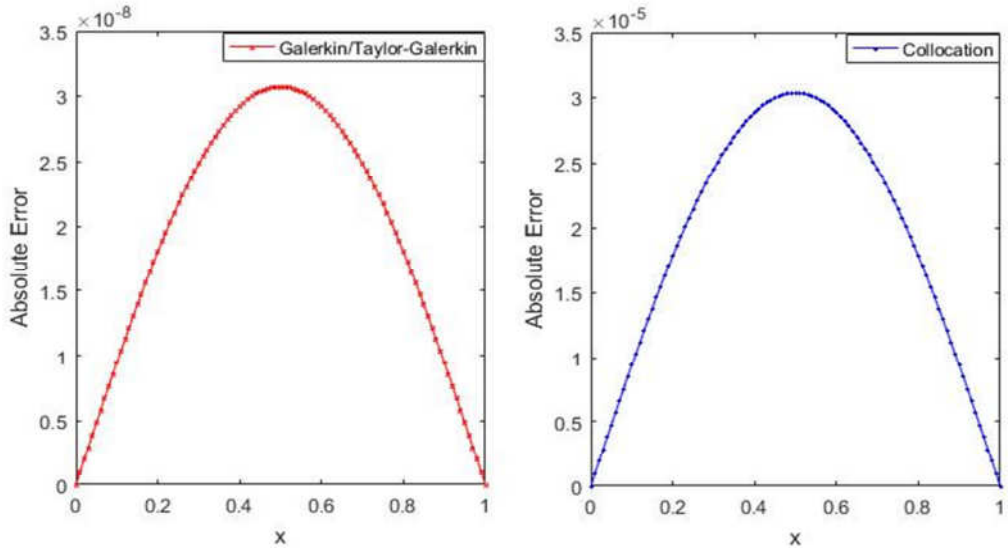


Figure 3. Comparison of the absolute errors of the Galerkin and collocation schemes for$h = 0.01\ m$, $dt = 0.0001\ s$ and $t = 1\ s$ in Problem 2.

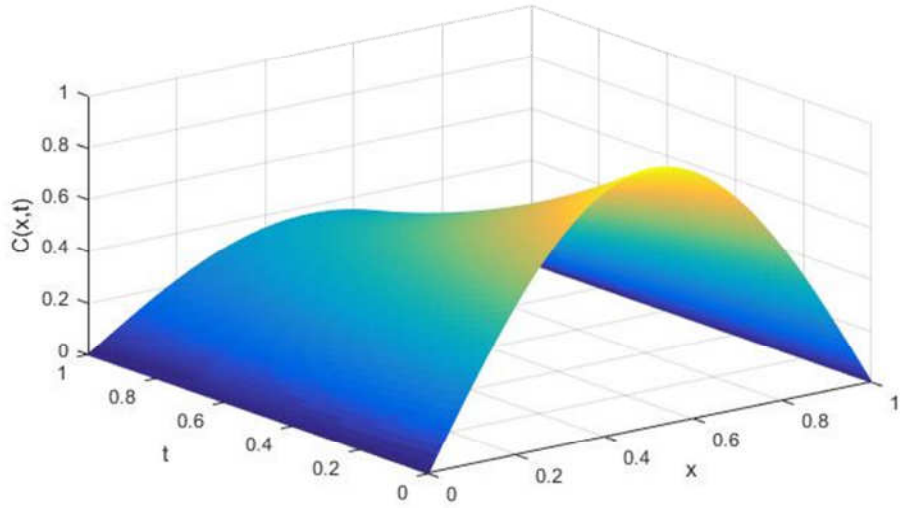


Figure 4. Diffusion process of the initial concentration for $D = 1/\pi^2$, $h = 0.01\ m$ and $dt = 0.0001\ s$ in Problem 2.

Problem 3 [14,10] Advection-diffusion process on a finite line

Consider the advection-diffusion process with the choice of $V = 1, v = 0.01$ and $\theta = 0$ in the ADR equation for which the exact solution, in a region bounded by $0 \le x \le 1$, is [10]:

$$C(x,t) = \frac{0.025}{\sqrt{0.000625 + 0.02t}} \exp\left(-\frac{(x+0.5-t)^2}{0.00125+0.04t}\right). \quad (39)$$

The boundary conditions of the problem can be written from exact solution (40) as follows:

$$f(x) = \exp\left(-\frac{(x+0.5)^2}{0.00125}\right),$$

$$g_0(t) = \frac{0.025}{\sqrt{0.000625+0.02t}} \exp\left(-\frac{(0.5-t)^2}{0.00125+0.04t}\right) \quad (40)$$

$$g_1(t) = \frac{0.025}{\sqrt{0.000625+0.02t}} \exp\left(-\frac{(1.5-t)^2}{0.00125+0.04t}\right). \quad (41)$$

The produced results and exact solution are compared in Table 4 for the various values of the Courant number and various values of α at the peak location of the concentration, $x = 0.5$ and $t = 1$. The present solutions are compared with the work of Kadalbajoo and Arora [14] in Table 5 under the consideration of the values $= 0.1, h = 0.01$, $Pe = 1$ and $t = 1$. As seen in Tables 4-5, the Galerkin method seems to be more accurate than the other suggested methods. It is also seen that the Galerkin and the Taylor-Galerkin methods are preferable comparison to the literature [14]. For various choices of the parameters, qualitative behavior of the model problem is sketched in Figures 5-6. There has been seen to be good agreement among the suggested methods (see Figure 6). Thus the produced results showed that the current problem has represented the physical behaviour very well. It is important to note that the collocation method takes less computational time than both the Galerkin and the Taylor-Galerkin methods while the collocation method is of a bit cruder results than the others.

Table 4. Comparison of the peak value for $h = 0.01$, $Pe = 1$, $x = 0.5$and $t = 1$ in Problem 3.

dt	Cr	$Exact$	$GFEM$	$CFEM$	$TGFEM$
0.0005	0.025	0.17407766	0.17407779	0.17403093	0.17407778
0.0010	0.050	0.17407766	0.17407819	0.17403132	0.17407819
0.0020	0.100	0.17407766	0.17407975	0.17403288	0.17408007
0.0033	0.167	0.17407766	0.17408344	0.17403655	-
0.0040	0.200	0.17407766	0.17408597	0.17403907	-
0.0080	0.400	0.17407766	0.17411025	0.17406324	-
0.0100	0.500	0.17407766	0.17412783	0.17408073	-

Table 5. A comparison of present solutions with the literature [14] and exact solution at various nodal points and at $t = 1$ in Problem 3.

X	Exact	TGBS2L [14]	TGBS2HQ [14]	TGBS2HL [14]	GFEM	CFEM	TGFEM
0.1	0.0035992	0.0036360	0.0036055	0.0036070	0.0035991	0.0035981	0.0035995
0.2	0.0196422	0.0196436	0.0196048	0.0195969	0.0196419	0.0196473	0.0196428
0.3	0.0660098	0.0656522	0.0658887	0.0658624	0.0660105	0.0660280	0.0660080
0.4	0.1366027	0.1362141	0.1366971	0.1367171	0.1366054	0.1365920	0.1365975
0.5	0.1740776	0.1746041	0.1744042	0.1744743	0.1740782	0.1740313	0.1740782
0.6	0.1366027	0.1373258	0.1365433	0.1365294	0.1365989	0.1366069	0.1366108
0.7	0.0660098	0.0659361	0.0657310	0.0656721	0.0660075	0.0660477	0.0660133
0.8	0.0196422	0.0193197	0.0196024	0.0195950	0.0196433	0.0196494	0.0196397
0.9	0.0035992	0.0034754	0.0036699	0.0036847	0.0035953	0.0035902	0.0035948

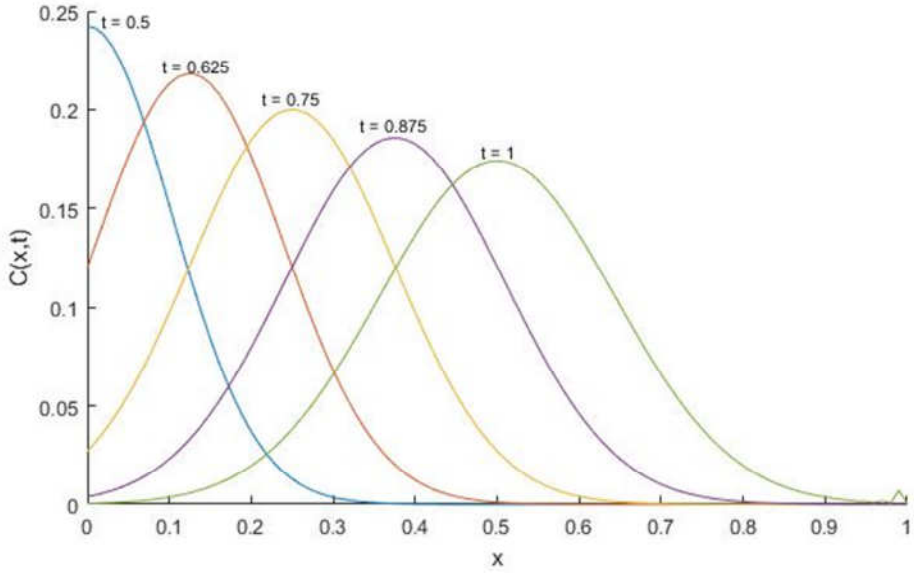


Figure 5. GFEM solution at various times in Problem 3 for $h = 0.01, dt = 0.001,\ Cr = 0.1$ and $Pe = 1$.

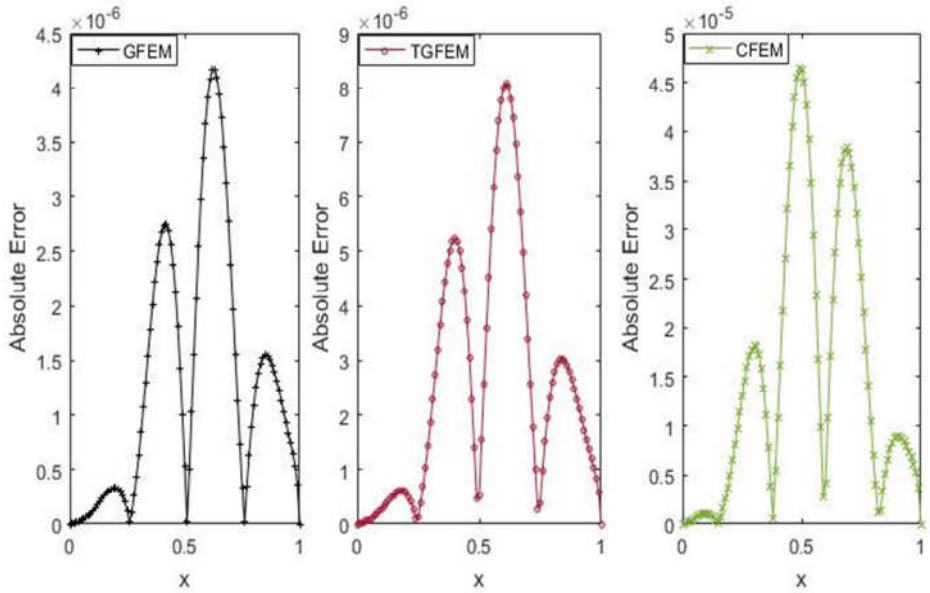


Figure 6. Absolute errors of the suggested methods at with the parameters $t = 1, h = 0.01, dt = 0.001, Cr = 0.1$ and $Pe = 1$ in Problem 3.

Problem 4. Advection-diffusion-reaction process on a finite line

Let us now consider the ADR equation with arbitrary values of V, θ and with homogeneous Dirichlet boundary conditions. Taking following initial condition

$$C(x,0) = \sin(\pi x) \tag{42}$$

leads us to derive the following Fourier series exact solution,

$$C(x,t) = \sum_{n=0}^{\infty} A_n \exp\left(-\left(\frac{Dn^2}{\pi^2 L^2} + \frac{V^2}{4D} + \theta\right)\right) exp\left(\frac{V^2}{2D}x\right) sin\left(\frac{n\pi x\pi}{L}\right) \tag{43}$$

where $x \in [0, L]$, $t > 0$ and

$$A_n = \int_0^L sin\left(\frac{n\pi x}{L}\right) \sin(\pi x)\, exp\left(-\frac{V^2}{2D}x\right) dx \tag{44}$$

For the sake of simplicity, the upper bound of spatial domain is taken to be $L = 1$. For various values of the parameters, absolute errors of the considered numerical methods are compared in Figures 7-8.

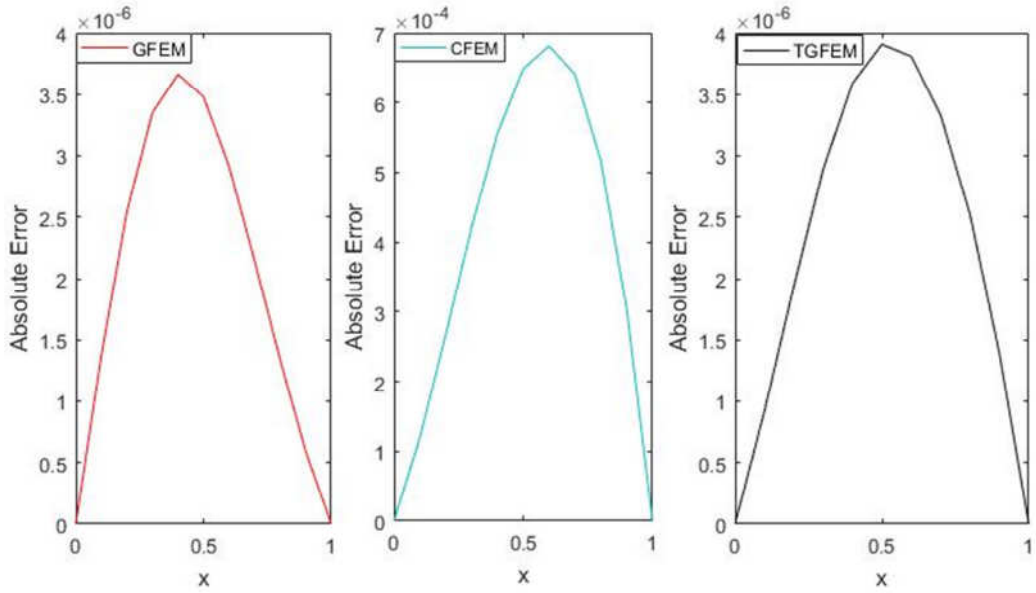


Figure 7. Comparison of the currently suggested methods in terms of the absolute errors for $t = 0.1$, $h = 0.05, dt = 0.001$, $D = 1, V = 1$ and $\theta = 0.5$ in Problem 4.

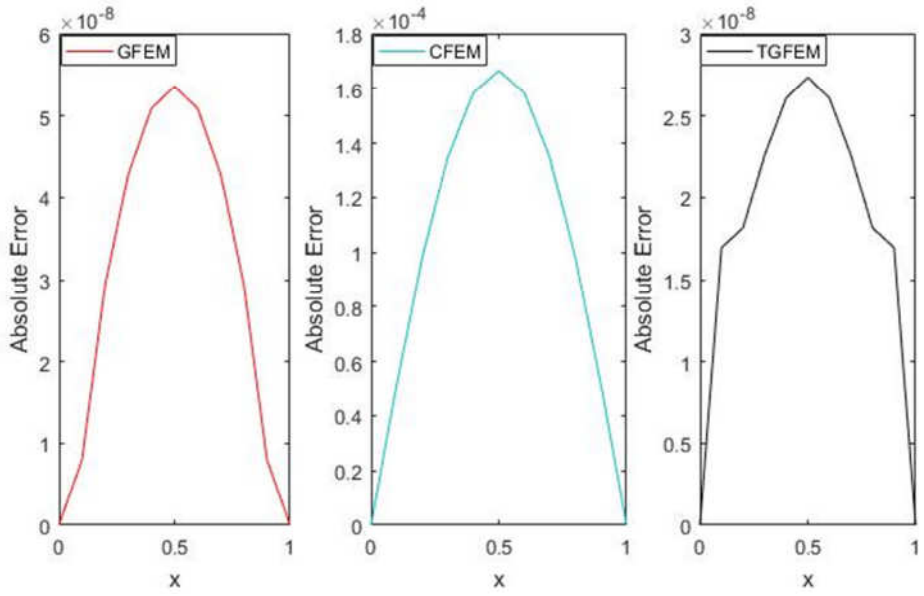


Figure 8. Comparison of the currently suggested methods in terms of the absolute errors for $t = 0.1$, $h = 0.05, dt = 0.001$, $D = 0.1, V = 0$ and $\theta = 1$ in Problem 4.

4. Conclusions and recommendation

This paper has concentrated on numerically analyzing the advection-diffusion-reaction equation by considering various finite element techniques including Galerkin, collocation and Taylor-Galerkin methods. To profoundly understand the physical processes represented by the model equation, various choices of the problem have been discussed both qualitatively and quantitatively. Comparison of the computed results showed that the Galerkin finite element method is more accurate and more versatile than the Taylor-Galerkin and the collocation methods. However, note that the collocation method takes less computational time than both the Galerkin and the Taylor-Galerkin methods while the collocation method is a bit cruder than the others. Further studies can concentrate on capturing the response of the same physical model with various forms of forcing terms.

References

[1] Dehghan, M. (2004). Weighted finite difference techniques for the one-dimensional advection-diffusion equation. *Applied Mathematics and Computation*, 147, 307-319.

[2] Spalding, D.B. (1972). A novel finite difference formulation for differential expressions ivolving both first and second derivatives. *International Journal for Numerical Methods in Engineering*, 4, 551-559.

[3] Irk, D., Dag, I.& Tombul, M.(2015). Extended cubic B-spline solution pf the advection-diffusion equation.*KSCE Journal of Civil Engineering*, 19, 929-934.

[4] Aksan, E.N. (2006). Quadratic B-spline finite element method for numerical solution of the Burgers equation. *Applied Mathematics and Computation*, 174, 884-896.

[5] Dag, I., Saka, B. & Boz, A. (2005). B-spline Galerkin methods for numerical solutions of Burgers' equation.*Applied Mathematics and Computation*, 166, 506-522.

[6] Dag, I., Irk, D. & Tombul, M. (2006). Least-squares finite element method for the advection diffusion equation.*Applied Mathematics and Computation*, 173, 554-565

[7] Goh, J., Majid, A.A., & Ismail, A.I. Md. (2012). Cubic B-spline collocation method for one-dimensional heat and advection-diffusion equations.*Journal of Applied Mathematics*, 37, 1-8.

[8] Ramakrishnan, C.V. (1979). An upwind finite element scheme for the unsteady convective diffusive transport equation.*Applied Mathematical Modelling*, 3, 280-284.

[9] Korkmaz, A., & Dag, I. (2016). Quartic and quintic B-spline methods for advection-diffusion equation. *Applied Mathematics and Computation*, 274, 208-219.

[10] Sari, M., Gürarslan, G. & Zeytinoglu, A. (2010). High-order finite difference schemes for solving the advection diffusion equation. *Mathematical and Computational Applications*, 15(3), 449-460.

[11] Prenter, P.M. (1975). Splines and Variational Methods. John Wiley & Sons, New York.

[12] Reddy, J. D. (2006). An Introduction to the Finite Element Method. McGraw-Hill. Singapore.

[13] Sun, H.& Zhang, J. (2003). A high order compact boundary value method for solving one dimensional heat equations. *Numerical Methods for Partial Differential Equations*, 19, 846—857.

[14] Kadalbajoo, M. K. & Arora, P. (2010). Taylor-Galerkin B-spline finite element method for the one-dimensional advection-diffusion equation. *Numerical Methods for Partial Differential Equations*, 26(5), 1206–1223.

Murat Sari *is an Associate Professor in the Department of Mathematics at Yildiz Technical University, Istanbul, Turkey. His current research interests include numerical solutions of differential equations, computational methods, computational fluid dynamics, mathematical modelling. He has about 50 international publications, over 30 proceedings, book/translation chapters. He supervised 12 graduate students and has been supervising 8 graduate students.*

Huseyin Tunc *is a Ph.D. student in the department of Mathematics at Yildiz Technical University. His research interests include approximate solutions of partial differential equations, fluid dynamics, finite element method and high order time integration methods.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.137-144 (2018)
http://doi.org/10.11121/ijocta.01.2018.00541

RESEARCH ARTICLE

New extensions of Chebyshev type inequalities using generalized Katugampola integrals via Pólya-Szegö inequality

Erhan Set[a*], Zoubir Dahmani[b] and İlker Mumcu[a]

[a] *Department of Mathematics, Faculty of Science and Arts, Ordu University, Ordu, Turkey*
[b] *Laboratory of Pure and Applied Mathematics (LPAM), Faculty of Exact Sciences, University of Mostaganem Abdelhamid Ben Badis (UMAB), Mostaganem, Algeria*
erhanset@yahoo.com, zzdahmani@yahoo.fr, mumcuilker@msn.com

ARTICLE INFO

Article History:
Received 15 September 2017
Accepted 18 February 2018
Available 09 March 2018

Keywords:
Pólya-Szegö inequality
Chebyshev inequality
Beta function
Katugampola fractional integral operators

AMS Classification 2010:
26A33, 26D10, 33B20

ABSTRACT

A number of Chebyshev type inequalities involving various fractional integral operators have, recently, been presented. In this work, motivated essentially by the earlier works and their applications in diverse research subjects, we establish some new Pólya-Szegö inequalities involving generalized Katugampola fractional integral operator and use them to prove some new fractional Chebyshev type inequalities which are extensions of the results in the paper: [On Pólya-Szegö and Chebyshev type inequalities involving the Riemann-Liouville fractional integral operators, J. Math. Inequal, 10(2) (2016)].

1. Introduction and preliminaries

This article is based on the well known Chebyshev functional [1]:

$$\begin{aligned} & T(f,g) \\ = & \frac{1}{b-a}\int_a^b f(x)\,g(x)\,dx \\ & -\left(\frac{1}{b-a}\int_a^b f(x)\,dx\right)\left(\frac{1}{b-a}\int_a^b g(x)\,dx\right), \end{aligned}$$

where f and g are two integrable functions which are synchronous on $[a,b]$, i.e.

$$(f(x)-f(y))(g(x)-g(y))\geq 0$$

for any $x,y\in[a,b]$, then the Chebyshev inequality states that $T(f,g)\geq 0$.
For some recent counterparts, generalizations of Chebyshev inequality, the reader is refer to [2–6].

We also need to introduce the Pólya and Szegö inequality [7]:

$$\frac{\int_a^b f^2(x)dx\int_a^b g^2(x)dx}{(\int_a^b f(x)g(x)dx)^2} \leq \frac{1}{4}\left(\sqrt{\frac{MN}{mn}}+\sqrt{\frac{mn}{MN}}\right)^2.$$

Using the above Pólya-Szegö inequality, Dragomir and Diamond [8] established the following Grüss type inequality:

Theorem 1. *Let $f,g:[a,b]\to\mathbb{R}_+$ be two integrable functions so that*

$$0<m\leq f(x)\leq M<\infty$$

and

$$0<n\leq g(x)\leq N<\infty$$

*Corresponding Author

for a.e. $x \in [a,b]$.
Then, we have

$$\begin{aligned}&|T(f,g;a,b)|\\&\leq \frac{1}{4}\frac{(M-m)(N-n)}{\sqrt{mnMN}}\\&\times\frac{1}{b-a}\int_a^b f(x)dx\frac{1}{b-a}\int_a^b g(x)dx.\end{aligned} \tag{1}$$

The constant $\frac{1}{4}$ *is best possible in (1) in the sense it can not be replaced by a smaller constant.*

For our purpose, we recall some other preliminaries: We note that the beta function $B(\alpha,\beta)$ is defined by (see, e.g. [9, Section 1.1])

$$B(\alpha,\,\beta) = \begin{cases}\displaystyle\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt & (\Re(\alpha)>0;\ \Re(\beta)>0)\\ \dfrac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} & \left(\alpha,\,\beta\in\mathbb{C}\setminus\mathbb{Z}_0^-\right),\end{cases} \tag{2}$$

where Γ is the familiar Gamma function. Here and in the following, let $\mathbb{C}$, $\mathbb{R}$, $\mathbb{R}^+$ and $\mathbb{Z}_0^-$ be the sets of complex numbers, real numbers, positive real numbers and non-positive integers, respectively, and let $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$.

Definition 1. *(see, e.g., [10], [11]) Let* $[a,b]$ $(-\infty < a < b < \infty)$ *be a finite interval on the real axis* $\mathbb{R}$. *The Riemann-Liouville fractional integrals(left-sided) of order* $\alpha \in \mathbb{C}, Re(\alpha) > 0$ *of a real function* $f \in L(a,b)$, *is defined:*

$$\left(J_{a+}^{\alpha}f\right)(x) := \frac{1}{\Gamma(\alpha)}\int_a^x \frac{f(t)}{(x-t)^{1-\alpha}}\,dt \quad (x>a). \tag{3}$$

Definition 2. *(see, e.g., [10], [11]) Let* (a,b) $(0 \leq a < b \leq \infty)$ *be a finite or infinite interval on the half-axis* $\mathbb{R}^+$. *The Hadamard fractional integrals(left-sided) of order* $\alpha \in \mathbb{C}, Re(\alpha) > 0$ *of a real function* $f \in L(a,b)$ *are defined by*

$$\left(H_{a+}^{\alpha}f\right)(x) := \frac{1}{\Gamma(\alpha)}\int_a^x \left(\log\frac{x}{t}\right)^{\alpha-1}\frac{f(t)}{t}\,dt \quad (a<x<b) \tag{4}$$

Definition 3. *(see, e.g., [10], [11]) Let* (a,b) $(-\infty \leq a < b \leq \infty)$ *be a finite or infinite interval on the half-axis* $\mathbb{R}^+$. *Also let* $Re(\alpha) > 0$, $\sigma > 0$ *and* $\eta \in \mathbb{C}$. *The Erdelyi-Kober fractional integrals (left-sided) of order* $\alpha \in \mathbb{C}$ *of a real function* $f \in L(a,b)$ *are defined by*

$$\begin{aligned}&\left(I_{a+,\sigma,\eta}^{\alpha}f\right)(x)\\&= \frac{\sigma x^{-\sigma(\alpha+\eta)}}{\Gamma(\alpha)}\int_a^x \frac{t^{\sigma(\eta+1)-1}}{(x^{\alpha}-t^{\alpha})^{1-\alpha}}f(t)\,dt\\&\quad (0\leq a<x<b\leq\infty).\end{aligned} \tag{5}$$

Definition 4. *[12] Let* $[a,b] \subset \mathbb{R}$ *be a finite interval. The Katugampola fractional integrals (left-sided) of order* $\alpha \in \mathbb{C}$, $\rho > 0$ $Re(\alpha > 0)$ *of a real function* $f \in X_c^p(a,b)$ *are defined by*

$$\begin{aligned}&\left({}^{\rho}I_{a+}^{\alpha}f\right)(x)\\&:= \frac{\rho^{1-\alpha}}{\Gamma(\alpha)}\int_a^x \frac{t^{\rho-1}}{(x^{\rho}-t^{\rho})^{1-\alpha}}f(t)\,dt. \qquad (x>a)\end{aligned} \tag{6}$$

Definition 5. *(see, e.g., [10], [11]) Let a continuous function by parts in* $\mathbb{R} = (-\infty,\infty)$. *The Liouville fractional integrals (left-sided) of order* $\alpha \in \mathbb{C}$, $\Re(\alpha) > 0$ *of a real function* f, *are defined by*

$$\begin{aligned}&\left(I_{+}^{\alpha}f\right)(x)\\&:= \frac{1}{\Gamma(\alpha)}\int_{-\infty}^x \frac{f(t)}{(x-t)^{1-\alpha}}\,dt \qquad (x\in\mathbb{R}).\end{aligned} \tag{7}$$

Here, the space $X_c^p(a,b)$ $(c \in \mathbb{R},\ 1 \leq p \leq \infty)$ consists of those complex-valued Lebesgue measurable functions φ on (a,b) for which $\|\varphi\|_{X_c^p} < \infty$, with

$$\|\varphi\|_{X_c^p} = \left(\int_a^b |x^c\varphi(x)|^p\frac{dx}{x}\right)^{1/p} \qquad (1\leq p<\infty)$$

and

$$\|\varphi\|_{X_c^p} = \mathtt{esssup}_{x\in(a,b)}[x^c|\varphi(x)|].$$

In particular, when $c = 1/p$ $(1 \leq p < \infty)$, the space $X_c^p(a,b)$ coincides with the classical $L^p(a,b)$ space.
Let $0 \leq a < x < b \leq \infty$. Also, let $\varphi \in X_c^p(a,b)$, $\alpha,\rho \in \mathbb{R}^+$, and β, η, $\kappa \in \mathbb{R}$. Then, the fractional integrals (left-sided and right-sided) of a function φ are defined, respectively, by (see [13])

$$\begin{aligned}&\left({}^{\rho}I_{a+,\eta,\kappa}^{\alpha,\beta}\varphi\right)(x)\\&:= \frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\alpha)}\int_a^x \frac{\tau^{\rho(\eta+1)-1}}{(x^{\rho}-\tau^{\rho})^{1-\alpha}}\varphi(\tau)\,d\tau\end{aligned} \tag{8}$$

and

$$\left({}^{\rho}I^{\alpha,\beta}_{b-,\eta,\kappa}\varphi\right)(x) \tag{9}$$
$$:= \frac{\rho^{1-\beta}x^{\rho\eta}}{\Gamma(\alpha)}\int_x^b \frac{\tau^{\kappa+\rho-1}}{(\tau^\rho - x^\rho)^{1-\alpha}}\varphi(\tau)\,d\tau.$$

Remark 1. The fractional integral (8) contains five well-known fractional integrals as its particular cases (see also [13–15]):

(i) Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in (8), the integral operator (8) reduces to the Riemann-Liouville fractional integral (3) (see also [10, p. 69]).
(ii) Setting $\kappa = 0$, $\eta = 0$, $a = -\infty$ and $\rho = 1$ in (8), the integral operator (8) reduces to the Liouville fractional integral (7) (see also [10, p.79]).
(iii) Setting $\beta = \alpha$, $\kappa = 0$, $\eta = 0$, and taking the limit $\rho \to 0^+$ with L'Hôspital's rule in (8), the integral operator (8) reduces to the Hadamard fractional integral (4) (see also [10, p. 110]).
(iv) Setting $\beta = 0$ and $\kappa = -\rho(\alpha+\eta)$ in (8), the integral operator (8) reduces to the Erdélyi-Kober fractional integral (5) (see also [10, p. 105]).
(v) Setting $\beta = \alpha$, $\kappa = 0$ and $\eta = 0$ in (8), the integral operator (8) reduces to the Katugampola fractional integral (6) (see also [12]).

The principle aim of the present paper is to establish new Pólya-Szegö inequalities and other of Chebyshev type by using generalized Katugampola fractional integration theory.

2. Main Results

In this section, we establish some new Chebyshev type inequalities involving the Katugampola fractional integration approach. Thanks to (2), we obtain (see [15, Eq. (3.1)])

$$\begin{aligned}&\frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\alpha)}\int_0^x \frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}d\tau\\ &= \frac{x^{\kappa+\rho(\eta+\alpha)}\,\Gamma(\eta+1)}{\rho^\beta\,\Gamma(\alpha+\eta+1)}\\ &= \Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\end{aligned} \tag{10}$$

$$\left(\alpha,\, x\in\mathbb{R}^+;\ \beta,\,\rho,\,\eta,\,\kappa\in\mathbb{R}\right).$$

We also let

$$\left({}^{\rho}I^{\alpha,\beta}_{0+,\eta,\kappa}\varphi\right)(x) := \left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\varphi\right)(x).$$

Lemma 1. *Let $\beta,\kappa\in\mathbb{R}$, x, α, $\rho\in\mathbb{R}^+$, and $\eta\in\mathbb{R}^+_0$. Let f and g be two positive integrable functions on $[0,\infty)$. Assume that there exist four positive integrable functions υ_1, υ_2, ω_1 and ω_2, such that:*

$$\begin{aligned}&0<\upsilon_1(\tau)\le f(\tau)\le \upsilon_2(\tau)\\ &0<\omega_1(\tau)\le g(\tau)\le \omega_2(\tau)\\ &(\tau\in[0,x], x>0)\end{aligned} \tag{11}$$

Then the following inequality holds:

$$\frac{{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2 f^2\right\}(x)\,{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2 g^2\right\}(x)}{\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{(\upsilon_1\omega_1+\upsilon_2\omega_2)fg\right\}(x)\right)^2}\le\frac{1}{4}. \tag{12}$$

Proof. From (11), for $\tau\in[0,x]$, $x>0$, we can write

$$\left(\frac{\upsilon_2(\tau)}{\omega_1(\tau)}-\frac{f(\tau)}{g(\tau)}\right)\ge 0 \tag{13}$$

and

$$\left(\frac{f(\tau)}{g(\tau)}-\frac{\upsilon_1(\tau)}{\omega_2(\tau)}\right)\ge 0. \tag{14}$$

Multiplying (13) and (14), we get

$$\left(\frac{\upsilon_2(\tau)}{\omega_1(\tau)}-\frac{f(\tau)}{g(\tau)}\right)\left(\frac{f(\tau)}{g(\tau)}-\frac{\upsilon_1(\tau)}{\omega_2(\tau)}\right)\ge 0.$$

From the above inequality, we can write

$$\begin{aligned}&(\upsilon_1(\tau)\omega_1(\tau)+\upsilon_2(\tau)\omega_2(\tau))\,f(\tau)g(\tau)\\ &\ge \omega_1(\tau)\omega_2(\tau)f^2(\tau)+\upsilon_1(\tau)\upsilon_2(\tau)g^2(\tau).\end{aligned} \tag{15}$$

Multiplying both sides of (15) by

$$\frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\alpha)}\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}$$

and integrating the resulting inequality with respect to τ over $(0,x)$, we get

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{(\upsilon_1\omega_1+\upsilon_2\omega_2)fg\right\}(x)\\ &\ge {}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2 f^2\right\}(x)+{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2 g^2\right\}(x)\end{aligned}$$

Applying the AM-GM inequality, i.e. $(a+b \geq 2\sqrt{ab}, \quad a,b \in \mathbb{R}^+)$, we have

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{(\upsilon_1\omega_1+\upsilon_2\omega_2)fg\right\}(x)\\&\geq 2\sqrt{{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2 f^2\right\}(x){}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2 g^2\right\}(x)}\end{aligned}$$

which implies that

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2 f^2\right\}(x){}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2 g^2\right\}(x)\\&\leq \frac{1}{4}\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{(\upsilon_1\omega_1+\upsilon_2\omega_2)fg\right\}(x)\right)^2.\end{aligned}$$

So, we get the desired result. □

Corollary 1. *If $\upsilon_1 = m$, $\upsilon_2 = M$, $\omega_1 = n$ and $\omega_2 = N$, then we have*

$$\begin{aligned}&\frac{\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}f^2\right)(x)\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}g^2\right)(x)}{\left(\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}fg\right)(x)\right)^2}\\ \leq\ &\frac{1}{4}\left(\sqrt{\frac{mn}{MN}}+\sqrt{\frac{MN}{mn}}\right)^2\end{aligned}$$

Remark 2. *Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in Lemma 1, yields the inequality in [16, Lemma 3.1].*

Lemma 2. *Let $\beta, \kappa \in \mathbb{R}$, x, α, θ, $\rho \in \mathbb{R}^+$, and $\eta \in \mathbb{R}_0^+$. Let f and g be two positive integrable functions on $[0,\infty)$. Assume that there exist four positive integrable functions υ_1, υ_2, ω_1 and ω_2 satisfying condition (11). Then the following inequality holds:*

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2\right\}(x)\\&\times{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{f^2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{g^2\right\}(x)\\&\leq\frac{1}{4}\Big({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1 g\right\}(x)\\&+{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_2 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\upsilon_2 g\right\}(x)\Big)^2.\end{aligned}\tag{16}$$

Proof. From (11), we get

$$\left(\frac{\upsilon_2(\tau)}{\omega_1(\xi)}-\frac{f(\tau)}{g(\xi)}\right)\geq 0$$

and

$$\left(\frac{f(\tau)}{g(\xi)}-\frac{\upsilon_1(\tau)}{\omega_2(\xi)}\right)\geq 0$$

which lead to

$$\begin{aligned}&\left(\frac{\upsilon_1(\tau)}{\omega_2(\xi)}+\frac{\upsilon_2(\tau)}{\omega_1(\xi)}\right)\frac{f(\tau)}{g(\xi)}\\&\geq\frac{f^2(\tau)}{g^2(\xi)}+\frac{\upsilon_1(\tau)\upsilon_2(\tau)}{\omega_1(\xi)\omega_2(\xi)}.\end{aligned}\tag{17}$$

Multiplying both sides of (17) by $\omega_1(\xi)\omega_2(\xi)g^2(\xi)$, we get

$$\begin{aligned}&\upsilon_1(\tau)f(\tau)\omega_1(\xi)g(\xi)+\upsilon_2(\tau)f(\tau)\omega_2(\xi)g(\xi)\\&\geq\omega_1(\xi)\omega_2(\xi)f^2(\tau)+\upsilon_1(\tau)\upsilon_2(\tau)g^2(\xi).\end{aligned}\tag{18}$$

Multiplying both sides of (18) by

$$\frac{\rho^{2(1-\beta)}x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}$$

and integrating the resulting inequality with respect to τ and ξ over $(0,x)^2$, we get

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1 g\right\}(x)\\&+{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_2 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\upsilon_2 g\right\}(x)\\&\geq{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{f^2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2\right\}(x)\\&+{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{g^2\right\}(x).\end{aligned}$$

Applying the AM-GM inequality, we have

$$\begin{aligned}&{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1 g\right\}(x)\\&+{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_2 f\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\upsilon_2 g\right\}(x)\\&\geq 2\sqrt{{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{f^2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{\omega_1\omega_2\right\}(x)}\\&\times\sqrt{{}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}\left\{\upsilon_1\upsilon_2\right\}(x){}^{\rho}I^{\theta,\beta}_{\eta,\kappa}\left\{g^2\right\}(x)}.\end{aligned}$$

So, we get the desired inequality of (16). □

Corollary 2. *If $\upsilon_1 = m$, $\upsilon_2 = M$, $\omega_1 = n$ and $\omega_2 = N$, then we have*

$$\begin{aligned}&\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\\&\times\frac{\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}f^2\right)(x)\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}g^2\right)(x)}{\left(\left({}^{\rho}I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^{\rho}I^{\theta,\beta}_{\eta,\kappa}g\right)(x)\right)^2}\\&\leq\frac{1}{4}\left(\sqrt{\frac{mn}{MN}}+\sqrt{\frac{MN}{mn}}\right)^2\end{aligned}$$

Remark 3. *Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in Lemma 2 yields the inequality in [16, Lemma 3.3].*

Lemma 3. *Suppose that all assumptions of Lemma 2 are satisfied. Then, we have:*

$$\begin{aligned} &{}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{f^2\right\}(x)\,{}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{g^2\right\}(x) \\ &\leq {}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{\frac{\upsilon_2 fg}{\omega_1}\right\}(x)\,{}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{\frac{\omega_2 fg}{\upsilon_1}\right\}(x). \end{aligned} \tag{19}$$

Proof. Using the condition (11), we get

$$\begin{aligned} &\frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\alpha)}\int_0^x \frac{\tau^{\rho(\eta+1)-1}}{(x^{\rho}-\tau^{\rho})^{1-\alpha}}f^2(\tau)d\tau \\ &\leq \frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\alpha)}\int_0^x \frac{\tau^{\rho(\eta+1)-1}}{(x^{\rho}-\tau^{\rho})^{1-\alpha}}\frac{\upsilon_2(\tau)}{\omega_1(\tau)}f(\tau)g(\tau)d\tau \end{aligned}$$

which leads to

$${}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{f^2\right\}(x) \leq {}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{\frac{\upsilon_2 fg}{\omega_1}\right\}(x). \tag{20}$$

Similarly, we have

$$\begin{aligned} &\frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\theta)}\int_0^x \frac{\xi^{\rho(\eta+1)-1}}{(x^{\rho}-\xi^{\rho})^{1-\theta}}g^2(\xi)d\xi \\ &\leq \frac{\rho^{1-\beta}x^{\kappa}}{\Gamma(\theta)}\int_0^x \frac{\xi^{\rho(\eta+1)-1}}{(x^{\rho}-\xi^{\rho})^{1-\theta}}\frac{\omega_2(\xi)}{\upsilon_1(\xi)}f(\xi)g(\xi)d\xi, \end{aligned}$$

which implies

$${}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{g^2\right\}(x) \leq {}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{\frac{\omega_2 fg}{\upsilon_1}\right\}(x). \tag{21}$$

Multiplying (20) and (21), we get the inequality of (19). □

Corollary 3. *If $\upsilon_1 = m$, $\upsilon_2 = M$, $\omega_1 = n$ and $\omega_2 = N$, then we have*

$$\frac{\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f^2\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g^2\right)(x)}{\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}fg\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}fg\right)(x)} \leq \frac{MN}{mn}.$$

Remark 4. *Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in Lemma 3 yields the inequality in [16, Lemma 3.4].*

Theorem 2. *Let $\beta, \kappa \in \mathbb{R}$, x, α, θ, $\rho \in \mathbb{R}^+$, and $\eta \in \mathbb{R}_0^+$. Let f and g be two positive integrable functions on $[0,\infty)$. Assume also that there exist four positive integrable functions υ_1, υ_2, ω_1 and ω_2 satisfying the condition (11). Then the following inequality holds:*

$$\begin{aligned} &\Big|\Lambda_{x,\kappa}^{\rho,\beta}(\alpha,\eta)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}fg\right)(x) \\ &+\Lambda_{x,\kappa}^{\rho,\beta}(\theta,\eta)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}fg\right)(x) \\ &-\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}g\right)(x) \\ &-\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}f\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g\right)(x)\Big| \\ \leq\ &|G_1(f,\upsilon_1,\upsilon_2)(x)+G_2(f,\upsilon_1,\upsilon_2)(x)|^{1/2} \\ &\times|G_2(g,\omega_1,\omega_2)(x)+G_2(g,\omega_1,\omega_2)(x)|^{1/2} \end{aligned} \tag{22}$$

where

$$\begin{aligned} &G_1(f,\upsilon_1,\upsilon_2)(x) \\ &=\frac{\Lambda_{x,\kappa}^{\rho,\beta}(\theta,\eta)}{4} \\ &\times\frac{\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{(\upsilon_1+\upsilon_2)f\right\}(x)\right)^2}{{}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\left\{\upsilon_1\upsilon_2\right\}(x)} \\ &-\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}f\right)(x) \end{aligned}$$

and

$$\begin{aligned} &G_2(f,\omega_1,\omega_2)(x) \\ &=\frac{\Lambda_{x,\kappa}^{\rho,\beta}(\alpha,\eta)}{4} \\ &\times\frac{\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{(\omega_1+\omega_2)f\right\}(x)\right)^2}{{}^{\rho}I_{\eta,\kappa}^{\theta,\beta}\left\{\omega_1\omega_2\right\}(x)} \\ &-\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}f\right)(x). \end{aligned}$$

Proof. Let f and g be two positive integrable functions on $[0,\infty)$. For $\tau,\xi \in (0,x)$ with $x>0$, we define $H(\tau,\xi)$ as

$$H(\tau,\xi) = (f(\tau)-f(\xi))\,(g(\tau)-g(\xi)),$$

equivalently,

$$\begin{aligned} &H(\tau,\xi) \\ &= f(\tau)g(\tau)+f(\xi)g(\xi)-f(\tau)g(\xi)-f(\xi)g(\tau). \end{aligned} \tag{23}$$

Multiplying both sides of (23) by

$$\frac{\rho^{2(1-\beta)}x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\frac{\tau^{\rho(\eta+1)-1}}{(x^{\rho}-\tau^{\rho})^{1-\alpha}}\frac{\xi^{\rho(\eta+1)-1}}{(x^{\rho}-\xi^{\rho})^{1-\theta}}$$

and double integrating the resulting inequality with respect to τ and ξ over $(0,x)^2$, we get

$$\begin{aligned}
&\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\\
&\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}H(\tau,\xi)d\tau d\xi\\
&=\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}fg\right)(x)\\
&+\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}fg\right)(x)\\
&-\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}g\right)(x)\\
&-\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}g\right)(x). \qquad (24)
\end{aligned}$$

Applying the Cauchy-Schwarz inequality for double integrals, we can write

$$\begin{aligned}
&\left|\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\right.\\
&\left.\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}H(\tau,\xi)d\tau d\xi\right|\\
&\le\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\\
&\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}f^2(\tau)d\tau d\xi+\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\\
&\times\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}f^2(\xi)d\tau d\xi\\
&-2\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\\
&\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}f(\tau)f(\xi)d\tau d\xi\Bigg]^{1/2}\\
&\times\Bigg[\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\\
&\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}g^2(\tau)d\tau d\xi+\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\\
&\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{\theta}}g^2(\xi)d\tau d\xi\\
&-2\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\\
&\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}g(\tau)g(\xi)d\tau d\xi\Bigg]^{1/2}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\left|\frac{\rho^2(1-\beta)x^{2\kappa}}{\Gamma(\alpha)\Gamma(\theta)}\int_0^x\int_0^x\frac{\tau^{\rho(\eta+1)-1}}{(x^\rho-\tau^\rho)^{1-\alpha}}\right.\\
&\left.\times\frac{\xi^{\rho(\eta+1)-1}}{(x^\rho-\xi^\rho)^{1-\theta}}H(\tau,\xi)d\tau d\xi\right|\\
&\le\left[\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f^2\right)(x)\right.\\
&+\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f^2\right)(x)\\
&\left.-2\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\right]^{1/2}\\
&\times\left[\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}g^2\right)(x)\right.\\
&+\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}g^2\right)(x)\\
&\left.-2\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}g\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}g\right)(x)\right]^{1/2}. \qquad (25)
\end{aligned}$$

Applying Lemma 1 with $\omega_1(\tau)=\omega_2(\tau)=g(\tau)=1$, we get

$$\begin{aligned}
&\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f^2\right)(x)\\
&\le\frac{\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)}{4}\frac{\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}\{(\upsilon_1+\upsilon_2)f\}(x)\right)^2}{{}^\rho I^{\alpha,\beta}_{\eta,\kappa}\{\upsilon_1\upsilon_2\}(x)}.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f^2\right)(x)\\
&-\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\\
&\le\frac{\Lambda^{\rho,\beta}_{x,\kappa}(\theta,\eta)}{4}\frac{\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}\{(\upsilon_1+\upsilon_2)f\}(x)\right)^2}{{}^\rho I^{\alpha,\beta}_{\eta,\kappa}\{\upsilon_1\upsilon_2\}(x)}\\
&-\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\\
&=G_1(f,\upsilon_1,\upsilon_2)(x). \qquad (26)
\end{aligned}$$

and

$$\begin{aligned}
&\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f^2\right)(x)\\
&-\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\\
&\le\frac{\Lambda^{\rho,\beta}_{x,\kappa}(\alpha,\eta)}{4}\frac{\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}\{(\omega_1+\omega_2)f\}(x)\right)^2}{{}^\rho I^{\theta,\beta}_{\eta,\kappa}\{\omega_1\omega_2\}(x)}\\
&-\left({}^\rho I^{\alpha,\beta}_{\eta,\kappa}f\right)(x)\left({}^\rho I^{\theta,\beta}_{\eta,\kappa}f\right)(x)\\
&=G_2(f,\omega_1,\omega_2)(x). \qquad (27)
\end{aligned}$$

Similarly, applying Lemma 1 with $\upsilon_1(\tau)=\upsilon_2(\tau)=f(\tau)=1$, we have

$$\Lambda_{x,\kappa}^{\rho,\beta}(\theta,\eta)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g^2\right)(x) - \left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}g\right)(x) \leq G_1(g,\omega_1,\omega_2)(x) \quad (28)$$

and

$$\Lambda_{x,\kappa}^{\rho,\beta}(\alpha,\eta)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}g^2\right)(x) - \left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\theta,\beta}g\right)(x) \leq G_2(g,\omega_1,\omega_2)(x). \quad (29)$$

Using (26)-(29), we conclude the desired result. □

Remark 5. *Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in Theorem 2, yields the inequality in [16, Theorem 3.6].*

Theorem 3. *Assume that all conditions of Theorem (2) are fulfilled. Then, we have:*

$$\left|\Lambda_{x,\kappa}^{\rho,\beta}(\alpha,\eta)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}fg\right)(x) - \left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f\right)(x)\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g\right)(x)\right| \leq |G(f,\upsilon_1,\upsilon_2)(x)G(g,\omega_1,\omega_2)(x)|^{1/2} \quad (30)$$

where

$$G(u,v,w)(x) = \frac{\Lambda_{x,\kappa}^{\rho,\beta}(\alpha,\eta)}{4} \times \frac{\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\{(v+w)u\}(x)\right)^2}{{}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}\{vw\}(x)} - \left(\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}u\right)(x)\right)^2.$$

Proof. Setting $\alpha = \theta$ in (22), we obtain (30). □

Corollary 4. *If $\upsilon_1 = m$, $\upsilon_2 = M$, $\omega_1 = n$ and $\omega_2 = N$, then we have*

$$G(f,m,M)(x) = \frac{(M-m)^2}{4Mm}\left(\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}f\right)(x)\right)^2,$$

$$G(g,n,N)(x) = \frac{(N-n)^2}{4Nn}\left(\left({}^{\rho}I_{\eta,\kappa}^{\alpha,\beta}g\right)(x)\right)^2.$$

Remark 6. We consider some particular cases of the result in Theorem 3.

(i) Setting $\kappa = 0$, $\eta = 0$ and $\rho = 1$ in the result in Theorem 3 yields the inequality in [16, Theorem 3.7],

(ii) Setting $\beta = 0$ and $\kappa = -\rho(\alpha+\eta)$ in the result in inequality 30 yields to

$$\left|\frac{\Gamma(\eta+1)}{\Gamma(\alpha+\eta+1)}\left(I_{0+,\rho,\eta}^{\alpha}fg\right)(x) - \left(I_{0+,\rho,\eta}^{\alpha}f\right)(x)\left(I_{0+,\rho,\eta}^{\alpha}g\right)(x)\right| \leq |G(f,\upsilon_1,\upsilon_2)(x)G(g,\omega_1,\omega_2)(x)|^{1/2}$$

where

$$G(u,v,w)(x) = \frac{\Gamma(\eta+1)}{4\Gamma(\alpha+\eta+1)} \times \frac{\left(I_{0+,\rho,\eta}^{\alpha}\{(v+w)u\}(x)\right)^2}{I_{0+,\rho,\eta}^{\alpha}\{vw\}(x)} - \left(\left(I_{0+,\rho,\eta}^{\alpha}u\right)(x)\right)^2$$

(iii) Setting $\beta = \alpha$, $\kappa = 0$ and $\eta = 0$ in the result in Theorem 3, under the corresponding reduced assumption, we obtain

$$\left|\frac{x^{\rho\alpha}}{\Gamma(\alpha+1)}\left({}^{\rho}I_{0+}^{\alpha}fg\right)(x) - \left({}^{\rho}I_{0+}^{\alpha}f\right)(x)\left({}^{\rho}I_{0+}^{\alpha}g\right)(x)\right| \leq |G(f,\upsilon_1,\upsilon_2)(x)G(g,\omega_1,\omega_2)(x)|^{1/2}$$

where

$$G(u,v,w)(x) = \frac{x^{\rho\alpha}}{4\Gamma(\alpha+1)} \times \frac{\left({}^{\rho}I_{0+}^{\alpha}\{(v+w)u\}(x)\right)^2}{{}^{\rho}I_{0+}^{\alpha}\{vw\}(x)} - \left(\left({}^{\rho}I_{0+}^{\alpha}u\right)(x)\right)^2.$$

References

[1] Čebyšev, P.L. (1982). Sur les expressions approximatives des integrales definies par les autres prises entre les memes limites. *Proc. Math. Soc. Charkov*, 2, 93-98.

[2] Belarbi, S. and Dahmani, Z. (2009). On some new fractional integral inequalities. *J. Ineq. Pure Appl. Math.*, 10(3) Article 86, 5 pages.

[3] Dahmani, Z., Mechouar, O. and Brahami, S. (2011). Certain inequalities related to the Chebyshevs functional involving a type Riemann-Liouville operator. *Bull. of Math. Anal. and Appl.*, 3(4), 38-44.

[4] Dahmani, Z. (2010). New inequalities in fractional integrals. *Int. J. of Nonlinear Sci.*, 9(4), 493-497.

[5] Set, E., Akdemir, O.A. and Mumcu, İ. Chebyshev type inequalities for conformable fractional integrals. Submitted.

[6] Özdemir, M.E., Set, E., Akdemir, O.A. and Sarıkaya, M.Z. (2015). Some new Chebyshev type inequalities for functions whose derivatives belongs to L_p spaces. *Afr. Mat.*, 26, 1609-1619.

[7] Polya, G. and Szegö, G. (1925). *Aufgaben und Lehrsatze aus der Analysis,* Band 1, Die Grundlehren der mathmatischen Wissenschaften 19, J. Springer, Berlin.

[8] Dragomir, S.S. and Diamond, N.T. (2003). Integral inequalities of Grüss type via Pólya-Szegö and Shisha-Mond results. *East Asian Math. J.,* 19, 27-39.

[9] Srivastava, H.M. and Choi, J. (2012). *Zeta and q-Zeta Functions and Associated Series and Integrals*, Elsevier Science Publishers, Amsterdam, London and New York.

[10] Kilbas, A.A., Srivastava, H.M. and Trujillo, J.J. (2006). *Theory and applications of fractional differential equations*, North-Holland Mathematical Studies, Vol.204, Elsevier (North-Holland) Science Publishers, Amsterdam, London and New York.

[11] Podlubny, I. (1999). *Fractional differential equations. An introduction to fractional derivatives, fractional differential equations, to methods of their Solution and some of their applications*, Academic Press, San Diego, CA.

[12] Katugampola, U.N. (2014). A new approach to generalized fractional derivatives. *Bull. Math. Anal. Appl.,* 6, 1-15.

[13] Katugampola, U.N. (2016). New fractional integral unifying six existing fractional integrals. arXiv:1612.08596v1 [math.CA].

[14] Sousa, J. and Oliviera, E.C. The Minkowski's inequality by means of a generalized fractional integral. arXiv preprint arXiv:1705.07191.

[15] Sousa, J., Oliveira, D.S and Oliveira, E.C. Grüss-type inequality by mean of a fractional integral. arXiv preprint arXiv:1705.00965.

[16] Ntouyas, S.K., Agarwal, P. and Tariboon, J. (2016). On Pólya-Szegö and Chebyshev type inequalities involving the Riemann-Liouville fractional integral operators. *J. Math. Inequal,* 10(2), 491-504.

Erhan Set *is a Associate Professor at the Department of Mathematics in Ordu University, Ordu, Turkey. He received his Ph.D. degree in Analysis in 2010 from Atatürk University, Erzurum, Turkey. He is editor on the Turkish Journal of Inequalities. He has a lot of awards. He has many research papers about convex functions, the theory of inequalities and fractional calculus.*

Zoubir Dahmani *is Professor, Labo. Pure and Appl. Maths., Faculty of Exact Sciences and Informatics, UMAB, University of Mostaganem, Algeria. He received his Ph.D. of Mathematics, USTHB (Algeria)and La ROCHELLE University (France), 2009. He has many research papers about the Evolution Equations,Inequality Theory, Fractional Differential Equations,Fixed Point Theory, Numerical Analysis, Probability Theory, Generalized Metric Spaces.*

İlker Mumcu *graduated from the department of mathematics teaching, Karadeniz Technical University, Trabzon, Turkey in 2000. He received his master degree in Mathematics from Ordu University in 2014. Her research interests focus mainly in inequality theory, especially Hermite- Hadamard, Grüss type inequalities and fractional integral inequalities.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.145-151 (2018)
http://doi.org/10.11121/ijocta.01.2018.00568

RESEARCH ARTICLE

Reproducing kernel Hilbert space method for solutions of a coefficient inverse problem for the kinetic equation

Esra Karatas Akgül

Department Mathematics, Faculty of Education, University of Siirt, Turkey
esrakaratas@siirt.edu.tr

ARTICLE INFO

Article History:
Received 16 December 2017
Accepted 08 February 2018
Available 11 April 2018

Keywords:
Reproducing kernel functions
Inverse problem for the kinetic equation
Reproducing kernel Hilbert space

AMS Classification 2010:
49J15, 46E22, 11P70

ABSTRACT

On the basis of a reproducing kernel Hilbert space, reproducing kernel functions for solving the coefficient inverse problem for the kinetic equation are given in this paper. Reproducing kernel functions found in the reproducing kernel Hilbert space imply that they can be considered for solving such inverse problems. We obtain approximate solutions by reproducing kernel functions. We show our results by a table. We prove the efficiency of the reproducing kernel Hilbert space method for solutions of a coefficient inverse problem for the kinetic equation.

1. Introduction

Kinetic theory emerged with Maxwell and Boltzmann, Hilbert, Enskog, Chapman, Vlasov, and Grad. Investigating for a form of matter which could clarify Saturn's rings, Maxwell considered that they were performed of rocks colliding and gravitating around the planet. The density of matter is then parameterized by the space position x and the velocity v of the rocks. Boltzmann modeled the operation, endowed a common representation of a dilute gas as particles undergoing collisions and with free motion between collisions, and he found the famous equation which is now named after him [1]. Vlasov obtained another kinetic equation (KE) for plasmas of charged particles. Kinetic equations (KEs) rise in a variety of sciences and implementations such as astrophysics, aerospace engineering, nuclear engineering, particle fluid interactions and semi-conductor technology recently. The general property of these models is that the underlying Partial Differential Equation is posed in the phase space $(x, v) \in R^n$ $n \geq 1$, [2].

We consider the problem of obtaining (f, σ) in Ω from the following equation [1]:

$$\begin{aligned} M_v(x,v)f_x(x,v) &- M_x(x,v)f_v(x,v) \\ &- \sigma(x,v)f = 0. \end{aligned} \quad (1)$$

with the boundary conditions:

$$f(a,v) = g(v), \quad f(b,v) = h(v) \quad (2)$$

$$f(x,c) = m(x), \quad f(x,d) = n(x). \quad (3)$$

In this work, the reproducing kernel functions for solving a coefficient inverse problem (IP) for the KE are given. Reproducing kernels were used for the first time at the beginning of the twentieth century by Zaremba in his work on boundary value problems for harmonic and biharmonic functions [3, 4]. The general theory of reproducing kernel Hilbert spaces was established simultaneously and independently by Aronszajn [5] and Bergman [6] in 1950. Mokhtari et al. have investigated an inverse problem for a parabolic equation with a nonlocal boundary condition in the reproducing kernel space [7]. Cui et al. have used

reproducing kernel method of solving the coefficient inverse problem [8]. Xu et al. used simplified reproducing kernel method for fractional differential equations with delay [9]. Tang et al. applied fitted reproducing kernel method for singularly perturbed delay initial value problems [10]. Fardi et al. implemented the reproducing kernel method for some variational problems depending on indefinite integrals [11]. Wahba used regression design for some equivalence classes of kernels [12]. Nashed et al. found regularization and approximation of linear operator equations in reproducing kernel spaces [13]. Al e'damat applied analytical-numerical method for solving a class of two-point boundary value problems [14]. For more details see [15-21].

2. Reproducing kernel functions

In this section, we give some important reproducing kernel functions.

Definition 1. *Hilbert function space H is a reproducing kernel space if and only if for any fixed $x \in X$, the linear functional $I(f) = f(x)$ is bounded [22].*

Definition 2. *We describe the space $T_2^2[1,2]$ as:*

$$T_2^2[1,2] = \{f \in AC[1,2] :\ f' \in AC[1,2],\ f'' \in L^2[1,2],\ f(1) = 0 = f(2)\}.$$

The inner product and the norm in $T_2^1[1,2]$ are obtained as follow:

$$\begin{aligned}\langle f, g\rangle_{T_2^2} &= \sum_{i=0}^{1} f^{(i)}(1)g^{(i)}(1) \\ &+ \int_1^2 f''(s)g''(s)\mathrm{d}s, \\ & f, g \in T_2^2[1,2]\end{aligned}$$

and

$$\|f\|_{T_2^2} = \sqrt{\langle f, f\rangle_{T_2^2}}, \quad f \in T_2^2[1,2].$$

Theorem 1. *Reproducing kernel function A_k of reproducing kernel space $T_2^2[1,2]$ is found as follow:*

$$A_k(s) = \begin{cases} \sum_{i=0}^{3} c_i(k)s^i, & s \le k, \\ \sum_{i=0}^{3} d_i(k)s^i, & s > k, \end{cases} \tag{4}$$

where

$$\begin{aligned} c_0(k) &= \frac{7}{12} - \frac{5}{8}k - \frac{1}{24}k^3 + \frac{1}{4}k^2, \\ c_1(k) &= -\frac{5}{8} - \frac{3}{16}k - \frac{1}{16}k^3 + \frac{3}{8}k^2, \\ c_2(k) &= \frac{1}{4} + \frac{7}{8}k + \frac{1}{8}k^3 - \frac{3}{4}k^2, \\ c_3(k) &= -\frac{5}{24} - \frac{1}{16}k - \frac{1}{48}k^3 + \frac{1}{8}k^2, \\ d_0(k) &= \frac{7}{12} - \frac{5}{8}k - \frac{5}{24}k^3 + \frac{1}{4}k^2, \\ d_1(k) &= -\frac{5}{8} - \frac{3}{16}k - \frac{1}{16}k^3 + \frac{7}{8}k^2, \\ d_2(k) &= \frac{1}{4} + \frac{3}{8}k + \frac{1}{8}k^3 - \frac{3}{4}k^2, \\ d_3(k) &= -\frac{1}{24} - \frac{1}{16}k - \frac{1}{48}k^3 + \frac{1}{8}k^2. \end{aligned}$$

Proof. Let $f \in T_2^2[1,2]$ and $1 \le k \le 2$. We have

$$\begin{aligned}\langle f, A_k\rangle_{T_2^2} &= \sum_{i=0}^{1} f^{(i)}(1)A_k^{(i)}(1) + \int_1^2 f''(x)A_k''(x)\mathrm{d}x \\ &= f(1)A_k(1) + f'(1)A_k'(1) \\ &\quad + f'(2)A_k''(2) - f'(1)A_k''(1) \\ &\quad - f(2)A_k^{(3)}(2) + f(1)A_k^{(3)}(1) \\ &+ \int_1^2 f(x)A_k^{(4)}(x)\mathrm{d}x.\end{aligned}$$

by integration by parts. Then, we get

$$\langle f, A_k\rangle_{T_2^2} = f(k).$$

This completes the proof. □

Definition 3. *We describe the space $M_2^2[-1,1]$ by:*

$$M_2^2[-1,1] = \{g \in AC[-1,1] :\ g' \in AC[-1,1],\ g'' \in L^2[-1,1],\ g(-1) = 0 = g(1)\}.$$

The inner product and the norm in $M_2^1[-1,1]$ are found as:

$$\begin{aligned}\langle g, h\rangle_{M_2^2} &= \sum_{i=0}^{1} g^{(i)}(-1)h^{(i)}(-1) \\ &+ \int_1^2 g''(z)h''(z)\mathrm{d}z, \\ & g, h \in M_2^2[-1,1]\end{aligned}$$

and

$$\begin{aligned}\|g\|_{M_2^2} &= \sqrt{\langle g, g\rangle_{M_2^2}}, \\ & g \in M_2^2[-1,1].\end{aligned}$$

Theorem 2. *Reproducing kernel function B_p of reproducing kernel space $M_2^2[-1,1]$ is acquired as:*

$$B_p(z) = \begin{cases} \sum_{i=0}^{3} c_i(p) z^i, & z \le p, \\ \sum_{i=0}^{3} d_i(p) z^i, & z > p, \end{cases} \tag{5}$$

where

$$\begin{aligned}
c_0(p) &= \frac{31}{240} + \frac{1}{80}p - \frac{17}{80}p^2 + \frac{17}{240}p^3, \\
c_1(p) &= \frac{1}{80} + \frac{13}{80}p - \frac{21}{80}p^2 + \frac{7}{80}p^3, \\
c_2(p) &= -\frac{17}{80} + \frac{19}{80}p - \frac{3}{80}p^2 + \frac{1}{80}p^3, \\
c_3(p) &= -\frac{23}{240} + \frac{7}{80}p + \frac{1}{80}p^3 - \frac{1}{20}p^3, \\
d_0(p) &= \frac{31}{240} + \frac{1}{80}p - \frac{17}{80}p^2 - \frac{23}{240}p^3, \\
d_1(p) &= \frac{1}{80} + \frac{13}{80}p + \frac{19}{80}p^2 + \frac{7}{80}p^3, \\
d_2(p) &= -\frac{17}{800} - \frac{21}{80}p - \frac{3}{80}p^2 + \frac{1}{80}p^3, \\
d_3(p) &= \frac{17}{240} + \frac{7}{80}p + \frac{1}{80}p^2 - \frac{1}{240}p^3.
\end{aligned}$$

Proof. Let $g \in M_2^2[-1,1]$ and $-1 \le p \le 1$. We obtain

$$\begin{aligned}
\langle g, B_p \rangle_{M_2^2} &= \sum_{i=0}^{1} g^{(i)}(-1) B_p^{(i)}(-1) \\
&+ \int_{-1}^{1} g''(x) B_p''(x) \mathrm{d}x \\
&= g(-1)B_p(-1) + g'(-1)B_p'(-1) \\
&\quad + g'(1)B_p''(1) - g'(-1)B_p''(-1) \\
&\quad - g(1)B_p^{(3)}(1) + g(-1)B_p^{(3)}(-1) \\
&+ \int_{-1}^{1} g(x) B_p^{(4)}(x) \mathrm{d}x.
\end{aligned}$$

by integration by parts. Then, we get

$$\langle g, B_p \rangle_{M_2^2} = g(p).$$

This completes the proof. □

Definition 4. *We describe the space $M_2^1[-1,1]$ as:*

$M_2^1[-1,1] = \{h \in AC[-1,1] : \ h' \in L^2[-1,1]\}$.

The inner product and the norm in $M_2^1[-1,1]$ are given as:

$$\begin{aligned}
\langle h, p \rangle_{V_2^1} &= h(-1)p(-1) \\
&+ \int_{-1}^{1} h'(t)p'(t)\mathrm{d}t, \\
h, p \in M_2^1[-1,1]&
\end{aligned}$$

and

$$\begin{aligned}
\|h\|_{T_2^1} &= \sqrt{\langle h, h \rangle_{T_2^1}}, \\
& h \in M_2^1[-1,1].
\end{aligned}$$

Lemma 1. *The space $M_2^1[-1,1]$ is a reproducing kernel space, and its reproducing kernel function E_k is given as [15]:*

$$E_k(s) = \begin{cases} 2 + s, & s \le k, \\ 2 + k, & s > k. \end{cases}$$

Definition 5. *We define the space $T_2^1[1,2]$ by*

$T_2^1[1,2] = \{h \in AC[1,2] : \ h' \in L^2[1,2]\}$.

The inner product and the norm in $T_2^1[1,2]$ are given as:

$$\langle h, p \rangle_{V_2^1} = h(1)p(1) + \int_1^2 h'(t)p'(t)\mathrm{d}t, \quad h, p \in T_2^1[1,2]$$

and

$$\|h\|_{T_2^1} = \sqrt{\langle h, h \rangle_{T_2^1}}, \quad h \in T_2^1[1,2].$$

Lemma 2. *The space $T_2^1[1,2]$ is a reproducing kernel space, and its reproducing kernel function F_p is given as [15]:*

$$F_p(z) = \begin{cases} z, & z \le p, \\ p, & z > p. \end{cases}$$

3. Main results

Definition 6. *If $m + n > 2$, define the binary space [22]*

$$W_2^{(m,n)}(\Omega) = \{u : \Omega \to R \,|$$

$Lu \in W_2^{(1,1)}(\Omega)$ if signature$(L) \preceq (m-1, n-1)\}$.

Equip $W_2^{(m,n)}(\Omega)$ with the inner product

$$\begin{aligned}
\langle u, v \rangle_{W_2^{(m,n)}(\Omega)} &= \sum_{i=0}^{m-1} \int_c^d \frac{\partial^n}{\partial t^n}\frac{\partial^i u}{\partial x^i}(a,t)\frac{\partial^n}{\partial t^n}\frac{\partial^i v}{\partial x^i}(a,t)dt \\
&+ \sum_{j=0}^{n-1} \left\langle \frac{\partial^j u}{\partial t^j}(\cdot, c), \frac{\partial^j v}{\partial t^j}(\cdot, c) \right\rangle_{W_2^m[a,b]} \\
&+ \int\int_\Omega \frac{\partial^2}{\partial x \partial t}\left(\frac{\partial^{m+n-2} u}{\partial x^{m-1}\partial t^{n-1}}\right)\frac{\partial^2}{\partial x \partial t}\left(\frac{\partial^{m+n-2} v}{\partial x^{m-1}\partial t^{n-1}}\right) dxdt.
\end{aligned}$$

We found the main reproducing kernel function for the problem in this section. We take $H(x,v) = x - \ln(v), a = -1, b = 1, c = 1, d = 2$, and

$$g(v) = \exp\left(-v + \frac{p}{8(4+v^2)} + \frac{\arctan(v/2)}{16}\right),$$

$$h(v) = \exp\left(v + \frac{v}{8(4+v^2)} + \frac{\arctan(v/2)}{16}\right),$$

$$m(x) = \exp\left(x^3 + \frac{1}{40} + \frac{\arctan(1/2)}{16}\right),$$

$$n(x) = \exp\left(2x^3 + \frac{1}{32} + \frac{\arctan(1)}{16}\right).$$

In our problem $m = 2, n = 2$. We obtain the main reproducing kernel function as: $J_{k,p}(s,z) = A_k(s)B_p(z)$.
where,

$$\begin{aligned} A_k(s) &= \frac{7}{12} - \frac{5}{8}k - \frac{1}{24}k^3 + \frac{1}{4}k^2 - \frac{5}{8}s \\ &- \frac{3}{16}sk - \frac{1}{16}sk^3 \\ &+ \frac{1}{4}s^2 + \frac{7}{8}s^2k + \frac{1}{8}s^2k^3 - \frac{3}{4}s^2k^2 \\ &- \frac{5}{24}s^3 - \frac{1}{16}s^3k \\ &+ \frac{3}{8}sk^2 - \frac{1}{48}s^3k^3 + \frac{1}{8}s^3k^2, \quad s \le k, \end{aligned}$$

$$\begin{aligned} A_k(s) &= \frac{7}{12} - \frac{5}{8}s - \frac{1}{24}s^3 + \frac{1}{4}s^2 - \frac{5}{8}k \\ &- \frac{3}{16}ks - \frac{1}{16}ks^3 \\ &+ \frac{1}{4}k^2 + \frac{7}{8}k^2s + \frac{1}{8}k^2s^3 - \frac{3}{4}k^2s^2 \\ &- \frac{5}{24}k^3 - \frac{1}{16}k^3s \\ &+ \frac{3}{8}ks^2 - \frac{1}{48}k^3s^3 + \frac{1}{8}k^3s^2, \quad s > k, \end{aligned}$$

$$\begin{aligned} B_p(z) &= \frac{7}{12} - \frac{5}{8}p - \frac{5}{24}p^3 + \frac{1}{4}p^2 - \frac{5}{8}z \\ &- \frac{3}{16}zp - \frac{1}{16}zp^3 \\ &+ \frac{7}{8}zp^2 + \frac{1}{4}z^2 + \frac{3}{8}z^2p + \frac{1}{8}z^2p^3 \\ &- \frac{3}{4}z^2p^2 - \frac{1}{24}z^3 \\ &- \frac{1}{16}z^3p - \frac{1}{48}z^3p^3 + \frac{1}{8}z^3p^2, \quad p \le z, \end{aligned}$$

$$\begin{aligned} B_p(z) &= \frac{7}{12} - \frac{5}{8}z - \frac{5}{24}z^3 + \frac{1}{4}z^2 - \frac{5}{8}p \\ &- \frac{3}{16}pz - \frac{1}{16}pz^3 \\ &+ \frac{7}{8}pz^2 + \frac{1}{4}p^2 + \frac{3}{8}p^2z + \frac{1}{8}p^2z^3 \\ &- \frac{3}{4}p^2z^2 - \frac{1}{24}p^3 \\ &- \frac{1}{16}p^3z - \frac{1}{48}p^3z^3 + \frac{1}{8}p^3z^2, \quad z > p. \end{aligned}$$

The reproducing kernel function $J_{(k,p)}(s,z)$ is in $W_2^{(2,2)}(\Omega)$.

Definition 7. *We say that a function $u : \Omega \to \mathbb{R}$ belongs to the binary space $W_2^{(1,1)}(\Omega)$ and write $u \in W_2^{(1,1)}(\Omega)$ provided $u \in AC(\Omega)$ and the following three square integrability conditions are satisfied [22]:*

(1) $u_x(\cdot, c) \in L^2[a,b]$;
(2) $u_t(a,\cdot) \in L^2[c,d]$;
(3) $u_{xt} \in L^2(\Omega)$.

Equip $W_2^{(1,1)}(\Omega)$ with the inner product

$$\begin{aligned} \langle u, v\rangle_{W_2^{(1,1)}} &= u(a,c)v(a,c) + \int_a^b u_x(x,c)v_x(x,c)dx \\ &+ \int_c^d u_t(a,t)v_t(a,t)dt \\ &+ \int_c^d \int_a^b u_{xt}(x,t)v_{xt}(x,t)dxdt. \end{aligned}$$

The binary space $W_2^{(1,1)}(\Omega)$ is a RKHS with reproducing kernel $G_{(k,p)}(s,z) = E_k(s)F_p(z)$.

4. Applications

The solution of (1)–(3) is given in the reproducing kernel space $W_2^{(2,2)}(\Omega)$ in this section. On defining the linear operator $N : W_2^{(2,2)}(\Omega) \to W_2^{(1,1)}(\Omega)$ by

$$\begin{aligned} Nf &= M_v(x,v)f_x(x,v) - M_x(x,v)f_v(x,v) \\ &- \sigma(x,v)f, \quad f \in W_2^{(2,2)}(\Omega), \end{aligned}$$

after homogenizing the boundary conditions, model problem (1)–(3) changes to the problem

$$\begin{aligned} Nf &= H(x,v,f(x,v)), \\ &(x,v) \in [-1,1] \times [1,2], \\ &f(a,v) = f(b,v) = f(x,c) = f(x,d) = 0. \end{aligned}$$

Lemma 3. *N is a bounded linear operator.*

Proof. Let $f \in W_2^{(2,2)}(\Omega)$ and $(x,v) \in \Omega$. We have

$$f(k,p) = \left\langle f, J_{(k,p)}\right\rangle_{W_2^{(2,2)}},$$

and

$$Nf(k,p) = \left\langle f, NJ_{(k,p)}\right\rangle_{W_2^{(2,2)}},$$

$$\frac{\partial}{\partial k}Nf(k,p) = \left\langle f, \frac{\partial}{\partial k}NJ_{(k,p)}\right\rangle_{W_2^{(2,2)}},$$

$$\frac{\partial}{\partial p}Nf(k,p) = \left\langle f, \frac{\partial}{\partial p}NJ_{(k,p)}\right\rangle_{W_2^{(2,2)}},$$

$$\frac{\partial}{\partial p}\frac{\partial}{\partial k}Nf(k,p) = \left\langle f, \frac{\partial}{\partial p}\frac{\partial}{\partial k}NJ_{(k,p)}\right\rangle_{W_2^{(2,2)}}.$$

Therefore, we have $a_0, b_0, a_1, b_1 > 0$ such that

$$|Nf(k,p)| \le a_0 \|f\|_{W_2^{(2,2)}},$$
$$\left|\frac{\partial}{\partial p}Nf(k,p)\right| \le b_0 \|f\|_{W_2^{(2,2)}},$$
$$\left|\frac{\partial}{\partial k}Nf(k,p)\right| \le a_1 \|f\|_{W_2^{(2,2)}},$$
$$\left|\frac{\partial}{\partial p}\frac{\partial}{\partial k}Nf(k,p)\right| \le b_1 \|f\|_{W_2^{(2,2)}}.$$

Thus, we get

$$\begin{aligned}
\|Nf\|^2_{W_2^{(1,1)}} &= \int_1^2 \left[\frac{\partial}{\partial p}Nf(-1,p)\right]^2 \mathrm{d}p \\
&+ \langle Nf(k,1), Nf(k,1)\rangle_{M_2^1} \\
&+ \int_{-1}^1\int_1^2 \left[\frac{\partial}{\partial k}\frac{\partial}{\partial p}Nf(k,p)\right]^2 \mathrm{d}k\mathrm{d}p \\
&= \int_1^2 \left[\frac{\partial}{\partial p}Nf(-1,p)\right]^2 \mathrm{d}p \\
&+ [Nf(-1,-1)]^2 \\
&+ \int_{-1}^1 \left[\frac{\partial}{\partial k}Nf(k,1)\right]^2 \mathrm{d}k \\
&+ \int_{-1}^1\int_1^2 \left[\frac{\partial}{\partial k}\frac{\partial}{\partial p}Nf(k,p)\right]^2 \mathrm{d}k\mathrm{d}p \\
&\le \left(a_0^2 + a_1^2 + T(b_0^2 + b_1^2)\right) \|f\|_{W_2^{(2,2)}}.
\end{aligned}$$

This completes the proof. □

Now, choose a countable dense subset $\{(x_1,v_1),(x_2,v_2),\ldots\}$ in Ω and define

$$\varphi_i = G_{(x_i,v_i)}, \quad \Psi_i = L^*\varphi_i,$$

where N^* is the adjoint operator of N. The orthonormal system $\{\widehat{\Psi}_i\}_{i=1}^\infty$ of $W_2^{(2,2)}$ can be derived from the process of Gram–Schmidt orthogonalization of $\{\Psi_i\}_{i=1}^\infty$ as

$$\widehat{\Psi}_i = \sum_{k=1}^i \beta_{ik}\Psi_k.$$

Theorem 3. *Let us assume $\{(x_i,v_i)\}_{i=1}^\infty$ is dense in Ω. Then $\{\Psi_i(x,v)\}_{i=1}^\infty$ is a complete system in $W_2^{(2,2)}$, and*

$$\Psi_i = NJ_{(x_i,v_i)}(x,v).$$

Proof. We acquire

$$\begin{aligned}
\Psi_i &= N^*\varphi_i = \left\langle N^*\varphi_i, J_{(x,v)}\right\rangle_{W_2^{(2,2)}} \\
&= \left\langle \varphi_i, NJ_{(x,v)}\right\rangle_{W_2^{(1,1)}} \\
&= \left\langle NJ_{(x,v)}, G_{(x_i,v_i)}\right\rangle_{W_2^{(1,1)}} \\
&= NJ_{(x,v)}(x_i,v_i) \\
&= NJ_{(x_i,v_i)}(x,v).
\end{aligned}$$

It is obvious that $\Psi_i \in W_2^{(2,2)}$. For each fixed $f \in W_2^{(2,2)}$, if

$$\langle f, \Psi_i\rangle_{W_2^{(2,2)}} = 0, \quad i = 1,2,\ldots,$$

then

$$\begin{aligned}
0 &= \langle f, \Psi_i\rangle_{W_2^{(1,1)}} \\
&= \langle f, N^*\varphi_i)\rangle_{W_2^{(2,2)}} \\
&= \langle Nf, \varphi_i\rangle_{W_2^{(1,1)}} \\
&= Nf(x_i,v_i), \quad i = 1,2,\ldots.
\end{aligned}$$

Note that $\{(x_i,v_i)\}_{i=1}^\infty$ is dense in Ω. Therefore, we obtain $Nf = 0$. From the existence of N^{-1}, it follows that $f = 0$. The proof is completed. □

Theorem 4. *If $\{(x_i,v_i)\}_{i=1}^\infty$ is dense in Ω, then the solution of the problem is obtained as:*

$$f(x,v) = \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} H(x_k,v_k,f(x_k,v_k))\widehat{\Psi}_i(x,v). \tag{6}$$

Proof. $\{\Psi_i(x,v)\}_{i=1}^\infty$ is a complete system in $W_2^{(2,2)}$. Therefore, we get

$$\begin{aligned}
f &= \sum_{i=1}^\infty \left\langle f, \widehat{\Psi}_i\right\rangle_{W_2^{(2,2)}} \widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} \langle f, \Psi_k\rangle_{W_2^{(2,2)}} \widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} \langle f, N^*\varphi_k\rangle_{W_2^{(2,2)}} \widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} \langle Nf, \varphi_k\rangle_{W_2^{(1,1)}} \widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} \left\langle Nf, G_{(x_k,v_k)}\right\rangle_{W_2^{(1,1)}} \widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} Nf(x_k,v_k)\widehat{\Psi}_i \\
&= \sum_{i=1}^\infty \sum_{k=1}^i \beta_{ik} H(x_k,v_k,f(x_k,v_k))\widehat{\Psi}_i(x,t).
\end{aligned}$$

This completes the proof. □

Now the approximate solution f_n can be obtained from the n-term intercept of the exact solution f and

$$f_n = \sum_{i=1}^n \sum_{k=1}^i \beta_{ik} H(x_k,v_k,f(x_k,v_k))\widehat{\Psi}_i. \tag{7}$$

Obviously

$$\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}} \to 0, \quad n \to \infty.$$

Theorem 5. *If $f \in W_2^{(2,2)}$ then, we have*

$$\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}} \to 0, \quad n \to \infty.$$

Moreover a sequence $\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}}$ is monotonically decreasing in n.

Proof. We have

$$\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}} = \left\| \sum_{i=n+1}^{\infty} \sum_{k=1}^{i} \beta_{ik} H(x_k, v_k, f(x_k, v_k)) \widehat{\Psi}_i(x,v) \right\|_{W_2^{(2,2)}}.$$

Therefore, we obtain

$$\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}} \to 0, \quad n \to \infty.$$

Furthermore, we have

$$\|f_n(x,v) - f(x,v)\|^2_{W_2^{(2,2)}} = \left\| \sum_{i=n+1}^{\infty} \sum_{k=1}^{i} \beta_{ik} H(x_k, v_k, f(x_k, v_k)) \widehat{\Psi}_i(x,v) \right\|^2_{W_2^{(2,2)}}$$

$$= \sum_{i=n+1}^{\infty} \left(\sum_{k=1}^{i} \beta_{ik} H(x_k, v_k, f(x_k, v_k)) \widehat{\Psi}_i(x,v) \right)^2.$$

It is obvious that $\|f_n(x,v) - f(x,v)\|_{W_2^{(2,2)}}$ is monotonically decreasing in n. □

To test the accuracy of the reproducing kernel Hilbert space mehod, an example has been given. The results are compared with the exact solutions. Let us take into consideration the problem of obtaining $(f(x,v), \sigma(x,v))$ in $\Omega = (-1,1) \times (1,2)$. The exact solution of the problem is given as [1]:

$$f(x,v) = \exp\left(x^3 v \frac{v}{8(4+v^2)} + \frac{\arctan(v/2)}{16} \right),$$

$$\sigma(x,v) = -3x^2 - x^3 - \frac{1}{(4+v^2)^2}.$$

Using our technique, we choose $25, 64$ and 100 points in the region $\Omega = [-1,1] \times [1,2]$ and obtain f_{25}, f_{64} and f_{100}. Numerical results are in good agreement with the exact solution. In order to prove the convergence of the exact solution we found absolute errors for different values of dense points n. We give the maximum absolute errors for different number of dense points in Table 1. The results demonstrate that the errors become smaller as n increases.

5. Conclusion

In this work, the reproducing kernel Hilbert space method was implemented for solving an inverse problem for the kinetic equation. Given technique is demonstrated to be of good convergence. It seems that this technique can also be applied to higher dimensional inverse problems. We found the reproducing kernel functions for solutions of a coefficient inverse problem for the kinetic equation. We concluded that these reproducing kernel functions can be used in much more complicated problems. We demonstrated our results by a table. These results proved the power of the reproducing kernel Hilbert space method.

Acknowledgments

The author thanks the anonymous authors whose work largely constitutes this sample file. This research was supported by 2017-SİUFEB-40.

References

[1] Golgeleyen, F., and Amirov, A. (2011). On the approximate solution of a coefficient inverse problem for the kinetic equation. *Math. Commun.* 16, 283-298.

[2] Perthame, B. (2014). Mathematical tools for kinetic equations, *Bull. Amer. Math. Soc.* 41, 205-244.

[3] Zaremba, S. (1907). L'equation biharmonique et une classe remarquable de fonctions fondamentales harmoniques. *Bulletin International l'Academia des Sciences de Cracovie*, pages 147-196.

[4] Zaremba, S. (1908). Sur le calcul numerique des fonctions demandees dan le probleme de dirichlet et le probleme hydrodynamique. *Bulletin International l'Academia des Sciences de Cracovie*, 125-195.

[5] Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68, 337-404.

[6] Bergman, S. (1950). The kernel function and conformal mapping. *American Mathematical Society, New York.*

[7] Mohammadi, M., Mokhtari, R., and Isfahani, F.T. (2014). Solving an inverse problem for a parabolic equation with a nonlocal boundary condition in the reproducing kernel space. *Iranian Journal of Numerical Analysis and Optimization* 4(1), 57-76.

[8] Cui, M., Lin, Y., and Yang, L. (2007). A new method of solving the coefficient inverse problem. *Science in China Series A: Mathematics Apr.*, 50(4), 561-572.

[9] Xu, M.Q., and Lin, Y. (2016). Simplified reproducing kernel method for fractional differential equations with delay. *Applied Mathematics Letters*, 52, 156-161.

[10] Tang, Z.Q. and Geng, F.Z. (2016). Fitted reproducing kernel method for singularly perturbed delay initial value problems. *Applied Mathematics and Computation*, 284, 169-174.

[11] Fardi, M., Ghaziani, R.K., and Ghasemi, M. (2016). The reproducing kernel method for some variational problems depending on indefinite integrals. *Mathematical Modelling and Analysis*, 21(3), 412-429.

[12] Wahba, G. (1974). Regression design for some equivalence classes of kernels. *The Annals of Statistics* 2(5), 925-934.

[13] Nashed, M. Z., and Wahba, G. (1974). Regularization and approximation of linear operator equations in reproducing kernel spaces. *Bulltetin of the American Mathematical Society*,80(6), 1213-1218.

[14] Al e'damat, A.H. (2015). Analytical-numerical method for solving a class of two-point boundary value problems. *International Journal of Mathematical Analysis* 9(40), 1987-2002.

[15] Inc, M., and Akgül, A. (2014). Approximate solutions for MHD squeezing fluid flow by a novel method. *Boundary Value Problems*, 2014:18.

[16] Inc, M., Akgül, A., and Kilicman, A. (2013). Numerical solutions of the second-order one-dimensional telegraph equation based on reproducing kernel hilbert space method. *Abstract and Applied Analysis*, Article ID 768963, 13 pages.

[17] Akgül, A., and Kilicman, A. (2015). Solving delay differential equations by an accurate method with interpolation. *Abstract and Applied Analysis*, Article ID 676939, 7 pages.

[18] Akgül, A. (2015). New reproducing kernel functions. *Mathematical Problems in Engineering*, Article ID 158134, 10 pages.

[19] Inc., M., and Akgül, A. (2014). Numerical solution of seventh-order boundary value problems by a novel method. *Abstract and Applied Analysis*, Article ID 745287, 9 pages.

[20] Boutarfa, B., Akgül, A., and Inc, M. (2017). New approach for the Fornberg-Whitham type equations. *Journal of Computational and Applied Mathematics*, 312, 13-26.

[21] Sakar, M.G., Akgül, A., and Baleanu, D. (2017). On solutions of fractional Riccati differential equations. *Advances in Difference Equations*. 2017:39.

[22] Akgül, A., and Grow, D. Existence of solutions to the Telegraph Equation in binary reproducing kernel Hilbert spaces, , Submitted.

Esra Karatas Akgül *is an Assistant Professor at the Department of Mathematics, Faculty of Education, Siirt University. He received his B. Sc. (2011), M. Sc. (2013) and Ph.D (2017) degrees from Department of Mathematics, Firat University, Turkey. His research areas includes differential equations and functional analysis.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.152-160 (2018)
http://doi.org/10.11121/ijocta.01.2018.00442

RESEARCH ARTICLE

Spectral tau algorithm for solving a class of fractional optimal control problems via Jacobi polynomials

Youssri H. Youssri* and Waleed M. Abd-Elhameed

Department of Mathematics, Faculty of Science, Cairo University, Giza, Egypt
youssri@sci.cu.edu.eg, waleed@sci.cu.edu.eg

ARTICLE INFO

Article History:
Received 21 January 2017
Accepted 06 March 2018
Available 11 April 2018

Keywords:
Jacobi polynomials
Tau method
Newton's iterative method
Optimal control problems
System of fractional differential equations

AMS Classification 2010:
65M70, 65N35, 35C10, 42C10

ABSTRACT

This paper is dedicated to presenting an efficient numerical algorithm for solving a class of fractional optimal control problems (FOCPs). The basic idea behind the suggested algorithm is based on transforming the FOCP under investigation into a coupled system of fractional-order differential equations whose solutions can be expanded in terms of the Jacobi basis. With the aid of the spectral-tau method, the problem can be reduced into a system of algebraic equations which can be solved via any suitable solver. Some illustrative examples and comparisons are presented aiming to demonstrate the accuracy, applicability and efficiency of the proposed algorithm.

1. Introduction

Fractional calculus is of course a very important branch of mathematical analysis. This branch of calculus is interested in generalizing the derivatives and integrals of integer order to include derivatives and integrals of an arbitrary order (real or complex). A large number of authors investigate fractional derivatives from both theoretical and practical points of view (see, for example, [1–4]). It is well-known that many physical phenomena in acoustics, damping laws, electroanalytical chemistry, neuron modeling, diffusion processing and material sciences (see for example, [5–7]) are described by fractional differential equations. Various algorithms are developed for handling different kinds of fractional differential equations. Some of these methods are, Adomian decomposition method [7, 8], variational iteration method [9] and fractional differential transform method [10, 11] and ultraspherical wavelets method [12].

Optimal control problems arise in various applied sciences such as mechanics, aerospace engineering, and economics. The exact solutions for most of these problems are not easy to implement, so it is natural to try to derive numerical solutions for such problems. The general definition of an optimal control problem requires the minimization of a criterion function of the states and control inputs of the system over a set of admissible control functions. This system is governed by constrained dynamics and control variables. Additional constraints such as final time constraints can be considered, see for example [13, 14].

Spectral and pseudospectral methods have become increasingly popular as higher order methods for the solution of various differential equations represent the solution of a certain problem in a basis set of orthogonal functions. The main advantage of these methods is that they can provide exponential convergence of the solutions. These methods are employed for solving many problems appear in various branches of science such as

*Corresponding Author

physics, chemistry and fluid mechanics. One can consult the monographs by Shen et al. [15], the books by Kopriva [16], and Shizgal [17] for some of these applications. There are three popular techniques for spectral methods, they are Galerkin, tau and collocation methods. Every method has its importance. The Galerkin approach requires to select suitable basis functions satisfying the boundary conditions and then enforcing the residual to be orthogonal with the basis functions. This method has been applied in a variety of papers. It has has been successfully applied to linear problems, see for example [18,19]. The collocation method is a suitable approach for treating nonlinear problems, see for example [20–23]. The tau method is a particular class of Petrov-Galerkin method. This method is often used for problems contain complicated boundary conditions.

Orthogonal polynomials in general and Jacobi polynomials in particular have distinguished parts in applied mathematical analysis. It is well-known that the class of the Jacobi polynomials includes six subclasses of orthogonal polynomials, namely, Legendre, ultraspherical and the four kinds of Chebyshev polynomials. The Jacobi polynomials and their special polynomials are extensively used along with spectral methods for solving ordinary and fractional differential equations, see for instance [24–29].

This paper introduces a general numerical algorithm for a class of FOCPs. A FOCP is an optimal control problem in which the criterion and/or the differential equations governing the dynamics of the system contain at least one fractional derivative operator.

FOCP can be described by different definitions of fractional derivatives. The most commonly used types of fractional derivatives are Riemann-Liouville and Caputo fractional derivatives. General necessary conditions of optimality have been developed for FOCPs. For instance, in [9, 11], the author has achieved Hamiltonian formulas for FOCPs with Riemann-Liouville fractional derivative, while the author in [30] has achieved some formulas with Caputo fractional derivative. Hamiltonian system of equations provides necessary conditions of optimization. The optimal solution of the FOCP should satisfy the system [31].

The main aim of this article is twofold:

- Reformulating fractional optimal control problem to a coupled system of fractional-order differential equations.
- Analyzing efficient spectral-tau algorithm for handling the resulting system of fractional-order differential equations via Jacobi polynomials.

The rest of the paper is arranged as follows. The next section is devoted to presenting mathematical preliminaries containing some basic definitions in the fractional calculus theory which are required for establishing our results. Also, some relevant properties of Jacobi polynomials and their shifted ones are presented. In Section 3, we present an efficient spectral tau algorithm for solving FOCPs. In Section 4, we give some numerical examples to ensure the efficiency, simplicity and applicability of the suggested algorithm. Finally, Section 5 is devoted to presenting concluding remarks.

2. Preliminaries

This section is devoted to presenting some basic definitions and properties in fractional calculus. In addition, some relevant properties of Jacobi polynomials and their shifted ones are presented.

2.1. Some definitions and properties of fractional calculus

Definition 1. *The Riemann-Liouville fractional integral operator I^α of order α on the usual Lebesgue space $L_1[0,1]$ is defined as*

$$I^\alpha f(t) = \begin{cases} \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau)\, d\tau, & \alpha > 0, \\ f(t), & \alpha = 0. \end{cases} \tag{1}$$

The operator I^α has the following properties:

$$(i)\, I^\alpha I^\beta = I^{\alpha+\beta},$$
$$(ii)\, I^\alpha I^\beta = I^\beta I^\alpha,$$
$$(iii) I^\alpha (t-a)^\nu = \frac{\Gamma(\nu+1)}{\Gamma(\nu+\alpha+1)} (t-a)^{\nu+\alpha},$$

where $f \in L_1[0,1]$, $\alpha, \beta \geqslant 0$, and $\nu > -1$.

Definition 2. *The Riemann-Liouville fractional derivative of order $\alpha > 0$ is defined by*

$$(D^\alpha f)(t) = \left(\frac{d}{dt}\right)^n (I^{n-\alpha} f)(t),\ n-1 \leqslant \alpha < n, n \in \mathbb{N}. \tag{2}$$

Definition 3. *The Caputo fractional left and right differential operators are defined respectively as*

$$({}_0D_t^\alpha f)(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau)\, d\tau, \tag{3}$$

$$({}_tD_1^{\alpha} f)(t) = \frac{1}{\Gamma(n-\alpha)} \int_t^1 (\tau - t)^{n-\alpha-1} f^{(n)}(\tau)\, d\tau, \tag{4}$$

where $n-1 \leqslant \alpha < n, n \in \mathbb{N}$.

The Caputo fractional differential operator satisfies the following basic formula, for $n-1 \leqslant \alpha < n$,

$$(I^{\alpha} D^{\alpha} f)(t) \quad = \quad f(t) - \sum_{k=0}^{n-1} \frac{f^{(k)}(0^+)}{k!} (t-a)^k,\ t > 0.$$

For more details on the mathematical properties of fractional derivatives and integrals, see for example, [3].

2.2. Classical Jacobi polynomials

For $\gamma, \delta \in \mathbb{R}, \gamma, \delta > -1$, and a nonnegative integer n, we denote by $P_n^{(\gamma,\delta)}(x)$ the Jacobi polynomial, which comprises all the polynomial solutions to singular Sturm-Liouville problems on $(-1, 1)$. $P_n^{(\gamma,\delta)}(x)$ has the following Gauss hypergeometric representation:

$$P_n^{(\gamma,\delta)}(x) = \frac{\Gamma(n+\gamma+1)}{n!\,\Gamma(\gamma+1)}\, {}_2F_1\left(-n, n+\gamma+1; \gamma+1; \frac{1-x}{2}\right).$$

The Jacobi polynomials are orthogonal on $[-1, 1]$ with the wight function $w^{(\gamma,\delta)}(x) = (1-x)^{\gamma}(1+x)^{\delta}$, in the sense that

$$\int_{-1}^{1} (1-x)^{\gamma}(1+x)^{\delta}\ P_m^{(\gamma,\delta)}(x) P_n^{(\gamma,\delta)}(x)\, dx = h_n^{(\gamma,\delta)}\, \delta_{mn},$$

where,

$$h_n^{(\gamma,\delta)} = \frac{2^{\gamma+\delta+1}\ \Gamma(n+\gamma+1)\,\Gamma(n+\delta+1)}{(2n+\gamma+\delta+1)\, n!\, \Gamma(n+\gamma+\delta+1)}, \tag{5}$$

and δ_{mn} is the well-known kronecker delta function. These polynomials are eigenfunctions of the following singular Sturm-Liouville equation

$$(1-x^2)u'' + \big(\delta - \gamma - (\gamma+\delta+2)x\big)u' + n(n+\gamma+\delta+1)u = 0.$$

Also, Jacobi polynomials may be generated by means of Rodrigue's formula

$$P_n^{(\gamma,\delta)}(x) = \frac{(-1)^n}{2^n\, n!\, w^{(\gamma,\delta)}(x)} \frac{d^n}{dx^n}\left[w^{(\gamma,\delta)}(x)(1-x^2)^n\right].$$

The polynomials, namely the ultraspherical, Legendre, first, second, third and fourth kinds Chebyshev polynomials, can be deduced as special cases of the Jacobi polynomials as shown in Table 1.

Table 1. Special cases of Jacobi polynomials.

Ultraspherical polynomials	$C_n^{(\lambda)} = \frac{(2\lambda)_n}{(\lambda+\frac{1}{2})_n} P_n^{(\lambda-\frac{1}{2},\lambda-\frac{1}{2})}(x)$
Legendre polynomials	$L_n(x) = P_n^{(0,0)}(x)$
Chebyshev polynomials of first kind	$T_n(x) = \frac{(\frac{1}{2})_n}{n!} P_n^{(-\frac{1}{2},-\frac{1}{2})}(x)$
Chebyshev polynomials of second kind	$U_n(x) = \frac{(\frac{3}{2})_n}{(n+1)!} P_n^{(\frac{1}{2},\frac{1}{2})}(x)$
Chebyshev polynomials of third kind	$V_n(x) = \frac{(2^n\, n!)^2}{(2n)!} P_n^{(-\frac{1}{2},\frac{1}{2})}(x)$
Chebyshev polynomials of fourth kind	$W_n(x) = \frac{(2^n\, n!)^2}{(2n)!} P_n^{(\frac{1}{2},-\frac{1}{2})}(x)$

Note that in this table, the symbol $(a)_n$ is the Pochhammer symbol, $(a)_n = \dfrac{\Gamma(a+n)}{\Gamma(n)}$. For more details about Jacobi polynomials, one can see [32–34].

2.3. Shifted Jacobi polynomials

In order to use the Jacobi polynomials on the interval $[0, 1]$, we define the so-called shifted Jacobi polynomials by introducing the change of variable $t = 2x - 1$. Let the shifted Jacobi polynomials $P_n^{(\gamma,\delta)}(2x-1)$ be denoted by $\rho_n^{(\gamma,\delta)}(x)$. Then $\rho_n^{(\gamma,\delta)}(x)$ can be obtained as follows:

$$\rho_n^{(\gamma,\delta)}(x) = \frac{(-1)^n}{n!\, x^{\delta}(1-x)^{\gamma}} \frac{d^n}{dx^n}\left[x^{\delta+n}(1-x)^{\gamma+n}\right].$$

The shifted Jacobi Polynomials $\rho_n^{(\gamma,\delta)}(x)$ has the orthogonality relation

$$\int_0^1 x^{\delta}(1-x)^{\gamma}\, \rho_m^{(\gamma,\delta)}(x)\, \rho_n^{(\gamma,\delta)}(x)\, dx = \frac{h_n^{(\gamma,\delta)}}{2^{\gamma+\delta+1}}\, \delta_{mn}. \tag{6}$$

The following special values will be of important use later

$$\rho_i^{(\gamma,\delta)}(0) = \frac{(-1)^i(\delta+1)_i}{i!}, \quad \rho_i^{(\gamma,\delta)}(1) = \frac{(\gamma+1)_i}{i!}, \tag{7}$$

$$D^q \rho_i^{(\gamma,\delta)}(0) = \frac{(-1)^{i-q}\,(\delta+q+1)_{i-q}(i+\gamma+\delta+1)_q}{(i-q)!}, \tag{8}$$

$$D^q \rho_i^{(\gamma,\delta)}(1) = \frac{(-1)^{i-q}\,(\gamma+q+1)_{i-q}(i+\gamma+\delta+1)_q}{(i-q)!}. \tag{9}$$

A function $f(t)$ defined over $[0, 1]$ may be expanded in terms of the shifted Jacobi polynomials as

$$f(t) = \sum_{i=0}^{\infty} c_i\, \rho_i^{(\gamma,\delta)}(t),$$

where,

$$c_i = \frac{2^{\gamma+\delta+1}}{h_i^{(\gamma,\delta)}} \int_0^1 t^{\delta}(1-t)^{\gamma} f(t)\, \rho_i^{(\gamma,\delta)}(t)\, dt. \quad (10)$$

The following theorem gives the differentiation and integration formulae of $\rho_i^{(\gamma,\delta)}(t)$

Theorem 1. *[35] The differentiation and integration of $\rho_i^{(\gamma,\delta)}(t)$ are given by the following formulae:*

$$\frac{d}{dt}\rho_i^{(\gamma,\delta)}(t) = (i+\gamma+\delta+1)\rho_{i-1}^{(\gamma+1,\delta+1)}(t), \quad (11)$$

$$\int_0^t \rho_i^{(\gamma,\delta)}(s)\, ds = \frac{1}{i+\gamma+\delta}\left(\rho_{i+1}^{(\gamma-1,\delta-1)}(t) + \frac{(-1)^i(\delta)_{i+1}}{(i+1)!}\right). \quad (12)$$

3. Numerical treatment for FOCP

In this section, we are interested in the reformulation of the fractional optimal control problem. In addition, we introduce numerical solutions for the proposed FOCP.

3.1. FOCP reformulation

The fractional optimal control problem (FOCP) can be defined as follows. Find the optimal control $u(t)$ for a fractional dynamical system (FDS) that minimizes the performance index

$$J(u) = \int_0^1 F(t, x(t), u(t))\, dt, \quad (13)$$

subject to the dynamical system

$$D_t^{\alpha} x(t) = G(t, x(t), u(t)), \alpha \in (n-1, n],\ t \in [0,1], \quad (14)$$

subject to the initial condition

$$x^{(i)}(0) = X_i, \quad i = 0, 1, \ldots, n-1, \quad (15)$$

where $x(t)$ is the state variable, t represents the time, F and G are two arbitrary functions and X_i are real known constants. The necessary optimality conditions (see, [30]) of FOCP (13-15) leads to

$${}_0D_t^{\alpha} x(t) = G(t, x(t), u(t)), \quad (16)$$

$${}_tD_1^{\alpha} \lambda(t) = \frac{\partial F}{\partial x} + \lambda \frac{\partial G}{\partial x}, \quad (17)$$

$$\frac{\partial F}{\partial u} + \lambda \frac{\partial G}{\partial u} = 0, \quad (18)$$

and

$$x^{(i)}(0) = X_i, \quad \lambda^{(i)}(1) = 0 \quad i = 0, 1, \ldots, n-1. \quad (19)$$

The necessary conditions for the optimality of the FOCP considered are those given in (16-19). These equations are similar to the Euler-Lagrange equations for classical optimal control problems except that the resulting differential equations contain the left and the right fractional derivatives. Moreover, the derivation of these equations is very similar to the derivation for an optimal control problem containing integral order derivatives. Determination of the optimal control for the fractional system requires solution of Eqs. (16)-(19).

3.2. Numerical algorithm for FOCP

In this section, we are concerned with the numerical solutions of the coupled system of equations (16)-(19), by applying shifted Jacobi tau method. First, we use Eq. (18), to eliminate $u(t)$ from Eqs. (16) and (17) to obtain a system of coupled fractional order differential equations of the form:

$${}_0D_t^{\alpha} x(t) = G_1(t, x(t), \lambda(t)), \quad (20)$$

$${}_tD_1^{\alpha} \lambda(t) = F_1(t, x(t), \lambda(t)), \quad (21)$$

subject to

$$x^{(i)}(0) = X_i, \quad \lambda^{(i)}(1) = 0 \quad i = 0, 1, \ldots, n-1. \quad (22)$$

We approximate $x(t), \lambda(t)$ as follows

$$x(t) \approx \sum_{i=0}^{N-1} a_i\, \rho_i^{(\gamma,\delta)}(t), \quad (23)$$

$$\lambda(t) \approx \sum_{i=0}^{M-1} b_i\, \rho_i^{(\gamma,\delta)}(t). \quad (24)$$

Now we can compute the residual of Eqs. (20) and (21) as follows:

$$R_1(t) = \sum_{i=0}^{N-1} a_i\, {}_0D_t^{\alpha} \rho_i^{(\gamma,\delta)}(t) - G_1\left(t, \sum_{i=0}^{N-1} a_i\, \rho_i^{(\gamma,\delta)}(t), \sum_{i=0}^{M-1} b_i\, \rho_i^{(\gamma,\delta)}(t)\right), \quad (25)$$

$$R_2(t) = \sum_{i=0}^{M-1} b_i\, {}_tD_1^{\alpha} \rho_i^{(\gamma,\delta)}(t) - F_1\left(t, \sum_{i=0}^{N-1} a_i\, \rho_i^{(\gamma,\delta)}(t), \sum_{i=0}^{M-1} b_i\, \rho_i^{(\gamma,\delta)}(t)\right). \quad (26)$$

The application of the typical tau method yields

$$\int_0^1 R_1(t)\rho_i^{(\gamma,\delta)}(t)\,dt = 0, \quad i = 0,1,\ldots,N-n-1, \tag{27}$$

$$\int_0^1 R_2(t)\rho_i^{(\gamma,\delta)}(t)\,dt = 0, \quad i = 0,1,\ldots,M-n-1. \tag{28}$$

The derivative and integral formulae of shifted Jacobi polynomials in (11) and (12) enable one to reduce the integrals in (27) and (28) into a set of algebraic equations in the unknown expansion coefficients. Moreover, the use of the initial conditions along with the identities (8) and (9) yield

$$\sum_{i=j}^{N-1} a_i \frac{(-1)^i(\delta+1)_i}{i!} = X_j, \quad j = 0,1,\ldots,n-1, \tag{29}$$

$$\sum_{i=j}^{M-1} b_i \frac{(\gamma+1)_i}{i!} = 0, \quad j = 0,1,\ldots,n-1. \tag{30}$$

Eqs. (27)-(30) generate a system of $(N+M)$ algebraic equations in the expansion coefficients, a_i and b_i, which can be solved with the aid of the well-known Newton's iterative method.

4. Numerical results and discussion

In this section, the spectral-tau Jacobi algorithm (STJA) is employed for handling some FOCPs accompanied with some comparisons hoping to demonstrate the efficiency and applicability of the proposed algorithm.

Example 1. *Consider the following minimization problem [31]:*

$$\min J = \frac{1}{2}\int_0^1 \left(x^2 + u^2 - 2t^2\,x - 2t^3\,u + t^4 + t^6\right) dt,$$
$${}_0D_t^\alpha x(t) = 2\left(t^{1+\alpha}\,\Gamma(3-\alpha)\right)^{-1} u(t), \quad \alpha \in (0,1),$$
$$x(0) = 0,$$

with the exact solution $\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}$.
This solution minimizes the performance index J, and the minimum value is $\mu = 0$. Applying the procedures explained in Section 3, we get the following coupled system:

$${}_0D_t^\alpha x(t) = 2\left(t^{1+\alpha}\,\Gamma(3-\alpha)\right)^{-1} \times \left[t^3 - 2\left(t^{1+\alpha}\,\Gamma(3-\alpha)\right)^{-1}\lambda(t)\right], \tag{31}$$

$${}_tD_1^\alpha \lambda(t) = x(t) - t^2, \tag{32}$$

$$x(0) = 0, \quad \lambda(1) = 1. \tag{33}$$

If we apply STJA with $N = M = 4$, then we have

$$x(t) \approx \sum_{i=0}^{3} a_i\,\rho_i^{(\gamma,\delta)}(t), \tag{34}$$

$$\lambda(t) \approx \sum_{i=0}^{3} b_i\,\rho_i^{(\gamma,\delta)}(t). \tag{35}$$

Now the residuals of (31), (32) *can be written in the following formulae*

$$R_1(t) = \sum_{i=0}^{3} a_i\,{}_0D_t^\alpha \rho_i^{(\gamma,\delta)}(t) - 2\left(t^{1+\alpha}\,\Gamma(3-\alpha)\right)^{-1} \times \left[t^3 - 2\left(t^{1+\alpha}\,\Gamma(3-\alpha)\right)^{-1}\sum_{i=0}^{3} b_i\,\rho_i^{(\gamma,\delta)}(t)\right], \tag{36}$$

$$R_2(t) = \sum_{i=0}^{3} b_i\,{}_tD_1^\alpha \rho_i^{(\gamma,\delta)}(t) - \sum_{i=0}^{3} a_i\,\rho_i^{(\gamma,\delta)}(t) + t^2. \tag{37}$$

The application of tau method yields

$$\int_0^1 R_1(t)\rho_j^{(\gamma,\delta)}(t)\,dt = 0, \quad j = 0,1,2, \tag{38}$$

$$\int_0^1 R_2(t)\rho_j^{(\gamma,\delta)}(t)\,dt = 0, \quad j = 0,1,2. \tag{39}$$

Moreover, the initial conditions lead to the following two equations:

$$\sum_{i=0}^{3} a_i \frac{(-1)^i(\delta+1)_i}{i!} = 0, \tag{40}$$

$$\sum_{i=0}^{3} b_i \frac{(\gamma+1)_i}{i!} = 1, \tag{41}$$

The system in (38)-(41) *can be solved to give*

$$a_0 = \frac{(1+\delta)_2}{(2+\gamma+\delta)_2}, a_1 = \frac{2(2+\delta)}{(2+\gamma+\delta)(4+\gamma+\delta)}, a_2 = \frac{2}{(3+\gamma+\delta)_2}, a_3 = 0,$$

and

$$b_i = 0, \quad 0 \le i \le 3.$$

Hence

$$x(t) = t^2, \quad \lambda(t) = 0.$$

Finally, and based on Eq. (18)*, we get*

$$u(t) = t^3,$$

which are the exact solutions.

Example 2. *Consider the following FOCP (see, [31]):*

$$\begin{aligned}
&\min J = \frac{1}{2}\int_0^1 \left[\left(x - t^2\right)^2 + \left(u + t^4 - \frac{20t^{\frac{9}{10}}}{9\Gamma(\frac{9}{10})}\right)^2 \right] dt,\\
&{}_0D_t^{1.1}x(t) = t^2\, x(t) + u(t),\\
&x(0) = x'(0) = 0,
\end{aligned}$$

with solution

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} t^2 \\ \frac{20t^{\frac{9}{10}}}{9\Gamma(\frac{9}{10})} - t^4 \end{pmatrix}.$$

This solution minimizes the performance index J and the minimum value is $\mu = 0$. In Table 2, we introduce the maximum absolute error

$$E := \max\left(\max_{t\in[0,1]} |x - x_N|, \max_{t\in[0,1]} |u - u_M|\right)$$

by using STJA for various choices of N, M, γ and δ.

Remark 1. *It is worthy to note that the best error obtained here is $3.28\ .\ 10^{-17}$ for $(N, M) = (5, 5)$ and $(\gamma, \delta) = (-\frac{1}{2}, \frac{1}{2})$, while the best error obtained in [31] is $7.03\ .\ 10^{-8}$, for $(M, N) = (8, 9)$.*

Example 3. *Consider the following time invariant FOCP: Find the control function $u(t)$ which minimizes the performance index (see, [9]):*

$$\begin{aligned}
min \quad J &= \frac{1}{2}\int_0^1 \left[\left(x(t)\right)^2 + \left(u(t)\right)^2 \right] dt,\\
{}_0D_t^{\alpha}x(t) &= -x(t) + u(t), \qquad 0 < \alpha \leqslant 1\\
x(0) &= 1.
\end{aligned}$$

The exact solution for this problem in case $\alpha = 1$, is

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} c_0 \sinh\left(\sqrt{2}t\right) + \cosh\left(\sqrt{2}t\right) \\ \left(c_0 + \sqrt{2}\right)\sinh\left(\sqrt{2}t\right) + \left(\sqrt{2}c_0 + 1\right)\cosh\left(\sqrt{2}t\right) \end{pmatrix},$$

where

$$c_0 = -\frac{\sqrt{2}\sinh\left(\sqrt{2}\right) + \cosh\left(\sqrt{2}\right)}{\sinh\left(\sqrt{2}\right) + \sqrt{2}\cosh\left(\sqrt{2}\right)}.$$

In Figure 1, we illustrate the exact solution in case of $\alpha = 1$, and the STJA solution for the case corresponds to $(N, M) = (7, 7)$, $(\gamma, \delta) = (-\frac{1}{2}, \frac{1}{2})$ (Chebyshev third kind case) and different values of α. Figure 1 shows the state and the control variables, respectively, as a function of time. From this figure it is clear that both the state and the control variables is close to the solution in the integral value of α as expected.

Example 4. *Consider a linear time varying system with the same performance index and the same initial condition as those considered in Example 3, except in this example, the system is subjected to the following dynamic constraint:*

$${}_0D_t^{\alpha}x(t) \quad = \quad t\, x(t) + u(t), \qquad 0 < \alpha \leqslant 1,$$

the exact solution for this problem in case $\alpha = 1$, is

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} c_1 e^{\frac{t^2}{2}}\sin(t) + e^{\frac{t^2}{2}}\cos(t) \\ c_1 e^{\frac{t^2}{2}}\cos(t) - e^{\frac{t^2}{2}}\sin(t) \end{pmatrix},$$

where $c_1 = \tan(1)$. In Figure 2, we illustrate the exact solution in case of $\alpha = 1$, and the STJA solution for the case corresponds to $(N, M) = (8, 8)$, $(\gamma, \delta) = (\frac{1}{2}, -\frac{1}{2})$ (Chebyshev fourth kind case) and different values of α. Figure 2 shows the state and the control variables, respectively, as a function of time. From this figure it is clear that both the state and the control variables is close to the solution in the integral value of α as expected.

Example 5. *Consider the following nonlinear FOCP (see, [36]):*

$$\begin{aligned}
\min \quad J = \int_0^1 \Big[& e^{t^{3/2}-t}u(t) - 2t^{3/2}x(t)\\
& + \frac{1}{4}e^{2t^{3/2}-2t} - \frac{3}{8}\sqrt{\pi}e^{t^{3/2}-2t}\\
& + t^3 + u(t)^2 - \frac{3}{4}\sqrt{\pi}e^{-t}u(t) + x(t)^2\\
& + e^{2t} + \frac{9}{64}\pi e^{-2t}\Big]\, dt,\\
& {}_0D_t^{1.5}x(t) = e^{x(t)} + 2e^t\, u(t),\\
& x(0) = x'(0) = 0.
\end{aligned}$$

The exact solution for this problem is

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} t^{\frac{3}{2}} \\ \frac{1}{2}\, e^{-t}(\frac{3\sqrt{\pi}}{4} - e^{t^{\frac{3}{2}}}) \end{pmatrix},$$

which minimizes the performance index J and the minimum value is $\frac{1}{2}(e^2 - 1) \approx 3.1945280494653251$. In Table 3, we introduce the maximum absolute error

Table 2. Maximum absolute error of Example 2.

(N,M)	(γ,δ)	E	(γ,δ)	E	(γ,δ)	E
(3,4)		$8.40 \cdot 10^{-10}$		$5.87 \cdot 10^{-10}$		$2.32 \cdot 10^{-10}$
(4,3)	$(0,0)$	$6.85 \cdot 10^{-10}$	$(-\frac{1}{2},-\frac{1}{2})$	$1.02 \cdot 10^{-10}$	$(\frac{1}{2},\frac{1}{2})$	$1.85 \cdot 10^{-10}$
(5,5)		$9.31 \cdot 10^{-16}$		$5.27 \cdot 10^{-16}$		$3.25 \cdot 10^{-16}$
(3,4)		$9.24 \cdot 10^{-10}$		$6.38 \cdot 10^{-10}$		$2.32 \cdot 10^{-10}$
(4,3)	$(-\frac{1}{2},\frac{1}{2})$	$4.28 \cdot 10^{-10}$	$(-\frac{1}{2},\frac{1}{2})$	$1.02 \cdot 10^{-10}$	$(1,2)$	$1.85 \cdot 10^{-10}$
(5,5)		$3.28 \cdot 10^{-17}$		$8.27 \cdot 10^{-17}$		$5.75 \cdot 10^{-16}$

$$E := \max\left(\max_{t\in[0,1]}|x-x_N|, \max_{t\in[0,1]}|u-u_M|\right)$$

by using STJA for various choices of N, M, γ *and* δ.

Remark 2. *It is worthy noting here, for* $M = N = 16$ *and* $(\gamma,\delta) = (\frac{1}{2},\frac{1}{2})$ *and* $(-\frac{1}{2},\frac{1}{2})$, *we get* $J = 3.1945280494653253$ *and* 3.1945280494653258 *which is almost the exact value.*

5. Concluding Remarks

In this paper, we presented an efficient numerical algorithm for some spectral solutions of a class of fractional optimal control problems. The proposed numerical solutions are expressed as expansions of Jacobi polynomials. The celebrated spectral tau method is utilized for this purpose. The numerical tests clarify that our method is applicable and it is more accurate than some other methods in literature.

Acknowledgments

The authors would like to thank the referees and the editor for their constructive comments; which helped substantially to improve the manuscript.

References

[1] Brunner, H., Pedas, A. and Vainikko, V. (2001). Piecewise polynomial collocation methods for linear Volterra integro-differential equations with weakly singular kernels. *SIAM J. Numer. Anal.*, 39(3),957–982.

[2] Kilbas, A.A., Trujillo, J.J. and Srivastava, H.M. (2006). *Theory and applications of fractional differential equations*, volume 204. Elsevier Science Limited.

[3] Podlubny, I. (1998). *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*, volume 198. Academic press.

[4] Stefan, S.G., Kilbas, A.A. and Marichev, O.T. (1993). Fractional integrals and derivatives. *Theory and Applications, Gordon and Breach, Yverdon.*

[5] Al-Mdallal, Q.M., Syam, M.I. and Anwar, M.N. (2010). A collocation-shooting method for solving fractional boundary value problems. *Commun. Nonlinear Sci. Numer. Simul.*, 15(12), 3814–3822.

[6] Çenesiz, Y., Keskin, Y. and Kurnaz, A. (2010). The solution of the Bagley–Torvik equation with the generalized Taylor collocation method. *J. Franklin Inst.*, 347(2), 452–466.

[7] Daftardar-Gejji, V. and Jafari, H. (2005). Adomian decomposition: a tool for solving a system of fractional differential equations. *J. Math. Anal. Appl.*, 301(2), 508–518.

[8] Momani, S. (2007). An algorithm for solving the fractional convection–diffusion equation with nonlinear source term. *Commun. Nonlinear Sci. Numer. Simul.*, 12(7), 1283–1290.

[9] Agrawal, O.P. (2004). A general formulation and solution scheme for fractional optimal control problems. *Nonlinear Dynam.*, 38(1-4), 323–337.

[10] Erturk, V.S., Momani, S. and Odibat, Z. (2008). Application of generalized differential transform method to multi-order fractional differential equations. *Commun. Nonlinear Sci. Numer. Simul.*, 13(8), 1642–1654.

[11] Agrawal, O.P. and Baleanu, D. (2007). A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems. *J. Vib. Control*, 13(9-10), 1269–1281.

[12] Abd-Elhameed, W.M. and Youssri, Y.H. (2015). New spectral solutions of multi-term fractional order initial value problems with error analysis. *CMES:Comp. Model. Eng. Sci*, 105, 375–398.

[13] Baleanu, D., Ozlem, D., and Agrawal, O.P. (2009). A central difference numerical scheme for fractional optimal control problems. *Journal of Vibration and Control*, 15(4), 583-597.

[14] Agrawal, O.P., Ozlem, D. and Baleanu, D. (2010). Fractional optimal control problems with several state and control variables. *Journal of Vibration and Control*, 16(13), 1967-1976.

[15] Shen, J., Tang, T. and Wang, L-L. (2011). *Spectral Methods: Algorithms, Analysis and Applications*, volume 41. Springer Science & Business Media.

[16] Kopriva, D.A. (2009). *Implementing spectral methods for partial differential equations: algorithms for scientists and engineers*. Springer Science & Business Media.

[17] Shizgal, S. (2015) *Spectral methods in chemistry and physics*. Springer.

[18] Doha, E.H., Abd-Elhameed, W.M. and Bhrawy, A.H. (2013). New spectral-Galerkin algorithms for direct solution of high even-order differential equations using symmetric generalized Jacobi polynomials. *Collect. Math.*, 64(3), 373–394.

Table 3. Maximum absolute error of Example 5.

$N = M$	(γ, δ)	E	(γ, δ)	E	(γ, δ)	E
4		$2.35 \cdot 10^{-4}$		$8.24 \cdot 10^{-4}$		$2.22 \cdot 10^{-4}$
10	$(0, 0)$	$5.91 \cdot 10^{-10}$	$(-\frac{1}{2}, -\frac{1}{2})$	$5.27 \cdot 10^{-10}$	$(\frac{1}{2}, \frac{1}{2})$	$5.86 \cdot 10^{-10}$
16		$3.29 \cdot 10^{-16}$		$8.61 \cdot 10^{-16}$		$1.11 \cdot 10^{-16}$
4		$4.19 \cdot 10^{-4}$		$3.61 \cdot 10^{-4}$		$2.54 \cdot 10^{-4}$
10	$(-\frac{1}{2}, \frac{1}{2})$	$7.72 \cdot 10^{-10}$	$(-\frac{1}{2}, \frac{1}{2})$	$2.58 \cdot 10^{-10}$	$(1, 2)$	$6.41 \cdot 10^{-10}$
16		$2.22 \cdot 10^{-17}$		$3.12 \cdot 10^{-17}$		$1.11 \cdot 10^{-16}$

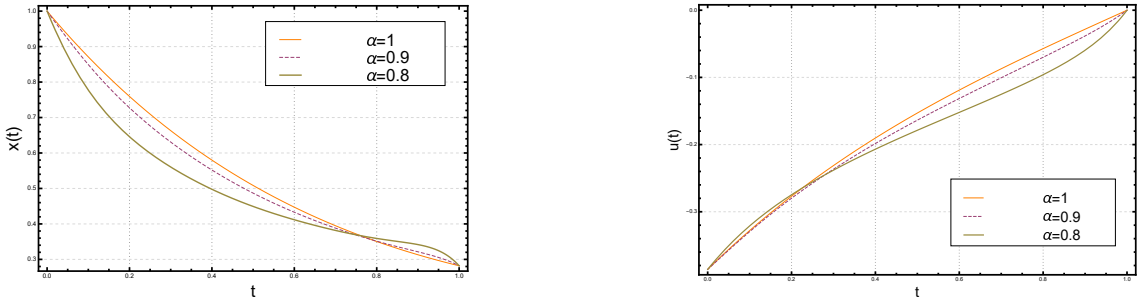


Figure 1. Different solutions of Example 3.

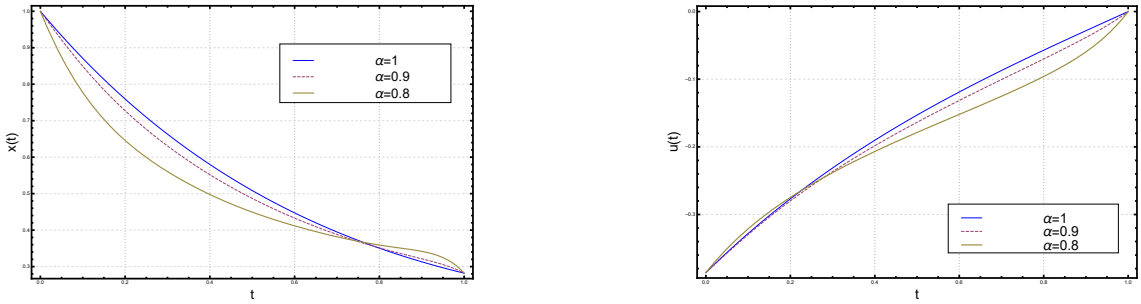


Figure 2. Different solutions of Example 4.

[19] Doha, E.H., Abd-Elhameed, W.M. and Bassuony, M.A. (2013). New algorithms for solving high even-order differential equations using third and fourth Chebyshev-Galerkin methods. *J. Comput. Phys.*, 236, 563–579.

[20] Abd-Elhameed, W.M., Doha, E.H. and Youssri, Y.H. (2013). New wavelets collocation method for solving second-order multipoint boundary value problems using Chebyshev polynomials of third and fourth kinds. *Abstr. Appl. Anal.*, Volume 2013:Article ID 542839, 9 pages.

[21] Abd-Elhameed, W.M., Doha, E.H., and Youssri, Y.H. (2013). New spectral second kind chebyshev wavelets algorithm for solving linear and nonlinear second-order differential equations involving singular and bratu type equations. *Abstr. Appl. Anal.*, Volume2013:Article ID 715756, 9 pages.

[22] Doha, E.H., Abd-Elhameed, W.M. and Youssri, Y.H. (2013). Second kind Chebyshev operational matrix algorithm for solving differential equations of Lane-Emden type. *New Astro.*, 23, 113–117.

[23] Abd-Elhameed, W.M. (2015). New Galerkin operational matrix of derivatives for solving Lane-Emden singular-type equations. *Eur. Phys. J. Plus*, 130(3), 52.

[24] Abd-Elhameed, W.M., Doha, E.H., and Youssri, Y.H. (2013). Efficient spectral-Petrov-Galerkin methods for third-and fifth-order differential equations using general parameters generalized jacobi polynomials. *Quaest. Math.*, 36(1), 15–38.

[25] Doha, E.H. and Abd-Elhameed, W.M. (2014). On the coefficients of integrated expansions and integrals of Chebyshev polynomials of third and fourth kinds. *Bull. Malays. Math. Sci. Soc.*, 37(2), 383–398.

[26] Bhrawy, A.H., and Zaky, M.A. (2016). Shifted fractional-order Jacobi orthogonal functions: application to a system of fractional differential equations. *Applied Mathematical Modelling*, 40(2), 832-845.

[27] Zaky, M.A., and Machado, J.A.T. (2017). On the formulation and numerical simulation of distributed-order fractional optimal control problems. *Communications in Nonlinear Science and Numerical Simulation*, 52, 177-189.

[28] Bhrawy, A.H., Zaky, M.A. and Machado, J.A.T. (2017). Numerical solution of the two-sided spacetime fractional telegraph equation via Chebyshev Tau approximation. *Journal of Optimization Theory and Applications*, 174(1), 321-341.

[29] Zaky, M.A. (2017). A Legendre spectral quadrature tau method for the multi-term time-fractional diffusion equations. *Computational and Applied Mathematics*, 1-14.

[30] Agrawal, O.P. (2008). A formulation and numerical scheme for fractional optimal control problems. *J. Vib. Control*, 14(9-10), 1291–1299.

[31] Lotfi, A., Yousefi, S.A. and Dehghan, M. (2013). Numerical solution of a class of fractional optimal control problems via the Legendre orthonormal basis combined with the operational matrix and the Gauss quadrature rule. *J. Comput. Appl. Math.*, 250, 143–160.

[32] Abramowitz, M. and Stegun, I.A. editors. (1970). *Handbook of Mathematical Functions (applied mathematics Series 55)*. National Bureau of Standards, New York.

[33] Andrews, G.E., Askey, R. and Roy, R. (1999) *Special functions*. Cambridge University Press, Cambridge.

[34] Szego, G. (1939). *Orthogonal polynomials*, volume 23. American Mathematical Soc.

[35] Rainville, E.D. (1960). *Special functions*. The Macmillan Company, New York.

[36] Keshavarz, E., Ordokhani, Y. and Razzaghi, M. (2015). A numerical solution for fractional optimal control problems via bernoulli polynomials. *Journal of Vibration and Control*, 22(18), 3889-3903.

Youssri H. Youssri *is Tenured assistant professor at Cairo University, he received a master's (2011) and doctorate (2014) in mathematics from Cairo University, Egypt. He has published more than 30 papers in different areas of Mathematics, including, orthogonal polynomials, spectral solutions of ordinary and fractional differential equations, boundary-value problems, initial-value problems and wavelets.*

Waleed M. Abd-Elhameed *is full professor of mathematics at Cairo University, he is a brilliant scientist in the field of spectral solution of differential equations and orthogonal polynomials. He has published more than 50 papers in different areas of Pure Mathematics.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.161-169 (2018)
http://doi.org/10.11121/ijocta.01.2018.00469

RESEARCH ARTICLE

Dynamic scheduling with cancellations: an application to chemotherapy appointment booking

Yasin Göçgün

Department of Industrial Engineering, Altinbas University,Turkey
yasin.gocgun@altinbas.edu.tr

ARTICLE INFO

Article History:
Received 27 April 2017
Accepted 02 March 2018
Available 22 April 2018

Keywords:
Dynamic scheduling
Markov decision processes
Approximate dynamic programming

AMS Classification 2010:
90-08, 41-04, 60JXX

ABSTRACT

We study a dynamic scheduling problem that has the feature of due dates and time windows. This problem arises in chemotherapy scheduling where patients from different types have specific target dates along with time windows for appointment. We consider cancellation of appointments. The problem is modeled as a Markov Decision Process (MDP) and approximately solved using a direct-search based approximate dynamic programming (ADP) technique. We compare the performance of the ADP technique against the myopic policy under diverse scenarios. Our computational results reveal that the ADP technique outperforms the myopic policy on majority of problem sets we generated.

1. Introduction

Dynamic scheduling problems arise in diverse fields such as manufacturing, transportation, project management, and healthcare. In these problems, unlike static scheduling problems, jobs arrive randomly at the system. Depending on the field, jobs could be orders, patients, and tasks. Arriving jobs are allocated to resources such as machines, operating rooms, and surgery rooms. Dynamic scheduling problems have received considerable attention in the operations research field in the last decade.

One of the recent variations of dynamic scheduling problems considers jobs with target dates and time windows. In these problems, jobs arriving randomly at a facility are ideally scheduled to specific target dates; if this is not possible, they are scheduled to days within a time window. Scheduling jobs outside their time windows results in penalty. These problems typically arise in chemotherapy appointment booking where jobs are considered as patients. Patients from each type have specific target dates and tolerance limits. In such settings, inefficient patient scheduling causes excessive wait listing, late patient appointment notifications, pharmacy congestion, unbalanced workload between nurses and considerable clerical work [1].Similar problems also arise in manufacturing where jobs from different types are scheduled for production considering target dates. In such settings, early scheduling causes inventory cost, whereas late scheduling results in penalty cost.

Chemotherapy appointment booking takes into account treatment protocols which are designed to maximize the efficacy of a chemotherapy treatment. Treatment protocols specify things such as the drugs to be administered, the dosage, appointment duration, the number of days between treatments, and tolerance limits [1].

In chemotherapy settings, appointments may be cancelled due to the fact that patients' protocols may change. It is fair to say that cancellations were generally ignored in the earlier work on chemotherapy appointment booking since the inclusion of cancellations makes the respective mathematical model significantly complex. However, cancellations should be among important features of realistic models.

Towards that end, we consider the chemotherapy appointment booking problem with cancellations.

In line with the literature, we model the problem as a discounted infinite horizon Markov Decision Process (MDP). Owing to intractability in state and action spaces, we resort to direct-search based approximate dynamic programming (ADP) to approximately solve the problem.

In this paper, we make the following contributions:

- We consider chemotherapy patient appointment booking problem with cancellations of treatments.
- We formulate the problem as an MDP.
- We employ a direct-search based ADP technique for approximately solving the underlying MDP model.
- We compare the performance of the direct-search based ADP technique against a myopic policy.
- Our computational results reveal that the direct-search based ADP improves the solution of the myopic policy on majority of problem sets we generated.

The paper is structured as follows. Section 2 reviews the relevant literature. In Section 3, the chemotherapy appointment booking problem with cancellations is described and its MDP formulation is provided. Section 4 includes the description of the Direct search-based ADP. Numerical results are provided in Section 5. Section 6 includes concluding remarks.

2. Literature review

Patient scheduling has been widely studied in the literature [2–4]. Green et al. [2] studied the issue of designing the outpatient appointment schedule and establishing dynamic priority rules for admitting patients into service. They formulated the problem as a finite-horizon dynamic program and identified the properties of the optimal policy. Cardoen et al. [3] reviewed operations research papers that discuss operating planning and scheduling. Hulshof et al. [4] reviewed research on resource capacity planning and control in health care.

Based on how patients are scheduled, patient scheduling can be divided into allocation scheduling and advance scheduling. Allocation scheduling requires that arriving patient requests be served or rejected immediately, whereas in advance scheduling, they are scheduled to specific slots/days in a booking horizon. Examples of research on allocation scheduling can be found in [2, 5–7]. Since the focus of this research is on advance scheduling, we review work on advance scheduling below.

Patrick et al. [8] studied problems where patients of different types are scheduled to future days in a diagnostic facility. They provided a Markov Decision Process (MDP) formulation of these problems; owing to intractability in state and action spaces, they employed a linear-programming based approximate dynamic programming (ADP) to obtain an approximate solution to underlying MDP.

Lamiri et al. [9] studied an operating room planning problem, considering two types of demand for surgery: elective surgery and emergency surgery. The planning problem is assigning elective cases to different periods over a planning horizon. The authors proposed a stochastic mathematical programming model to solve their problem.

Liu et al. [10] proposed a model for dynamic appointment scheduling problems, taking into account no-shows, cancellation, patient preferences, and overtime. Their results indicate that heuristic dynamic policies they proposed outperform benchmark policies.

Saure et al. [11] studied a dynamic patient scheduling problem in radiation therapy units. The authors provided an MDP formulation of their problem and employed linear programming-based ADP for obtaining an approximate solution. Their results reveal that the ADP technique outperforms the myopic policy in diverse problem sets.

Geng et al. [12] studied dynamic outpatient scheduling for a diagnostic facility with two waiting-time targets. The authors developed a finite-horizon MDP model for this problem, and characterized the optimal scheduling policy by proving the monotonicity and concavity properties of the components of the MDP model.

Tsai and Teng [13] proposed a stochastic appointment system for patients with a dynamic call-in sequence to outpatient clinics with multiple resources. In their model, the schedule for a single-service period includes a fixed number of blocks of equal length. Their results indicate that their stochastic model was able to schedule patients more efficiently as compared to traditional appointment systems.

There are several studies about patient scheduling on chemotherapy in the literature. These are summarized below.

Turkcan et al. [14] developed operations planning and scheduling methods for chemotherapy

patients, aiming to minimize deviation from optimal treatment plans due to limited availability of clinic resources. Their model differs from our model in that they used integer programming methods for solving their problem since they ignored random and dynamic arrival of patients.

Hahn-Goldberg et al. [15] studied dynamic optimization of chemotherapy outpatient scheduling with uncertainty. The authors developed dynamic template scheduling, combining proactive and online optimization.

Hahn-Goldberg et al. [16] modeled a deterministic chemotherapy outpatient scheduling problem using constraint programming. They compared the performance of their best constraint model against that of a mixed-integer programming (MIP) model. Their results revealed that constraint programming outperforms the MIP model.

In a recent study, Alvarado and Ntaimo [17] studied the problem of scheduling individual chemotherapy patient appointments and resources. The authors developed mean-risk stochastic integer programming models for solving their problem.

Our work differs from the abovementioned studies in that we consider dynamic scheduling with specific target dates and tolerance limits.

Gocgun and Puterman [1] extended the work in [8] by considering patients of different types with specific target dates and tolerance limits. They studied a chemotherapy patient appointment booking problem and formulated the problem as an MDP and employed linear-programming based ADP for approximately solving the MDP model. Their model differs from our model in that we consider cancellations of treatments.

3. Problem description and Markov decision process model

We consider the following dynamic patient scheduling problem (see [1] for a similar description).

- We consider an infinite time horizon and a finite rolling booking horizon. The booking horizon is a rolling period of N days.
- Patients are classified on the basis of their appointment tolerances.
- Each day patients of each type with specific target dates arrive at the facility randomly. Arrival distributions are assumed to be stationary, and arrivals across patient types and target dates are assumed to be independent.
- In line with the literature [1], we assume that each appointment requires one appointment slot.
- At the end of each day, scheduling to future days over a booking horizon is performed for arriving patients or those patients are diverted (i.e., they are sent to another hospital or served through overtime). (As stated in [1], diversion can be thought as overtime in chemotherapy settings or outsourcing in the manufacturing setting). Diversion/overtime capacity is assumed to be significantly higher than the maximum number of patients that need to be diverted on any day.
- There is no cost incurred for scheduling patients to days within their tolerance limits, whereas scheduling a patient to a day outside the tolerance limit results in a type-dependent scheduling cost per day.
- Diverting patients or serving them through overtime results in diversion cost, which is the same for patients of each type.
- We consider cancellation of appointments. In line with the related literature [18] we assumed that once an appointment is cancelled, it is not rescheduled to a future date. This is due to the fact that rescheduling cancelled appointments to a future day would further complicate the problem.
- The objective is to perform scheduling of arriving patients to available days (or divert them) in a booking horizon so as to minimize total discounted expected cost.

  In line with the literature [1, 2, 19], we assume that scheduling costs for patients of each type are linear in the number of days outside their time window. The function of these costs is to penalize the system for violating time windows. As stated in [1], higher scheduling costs are incurred for patients with lower tolerances. The function of diversion costs is to penalize the system for not scheduling a patient [1].

  It is worth noting that in our problem, the diversion is not necessarily considered as overtime. Utilizing overtime means the appointment can be dealt with through overtime on the day the decision is given. Hence, in line with the literature [1], we do not consider scheduling an appointment to a future day and at the same time performing it through overtime.

3.1. Markov decision process model

3.1.1. State Space

The state space $s \in S$ takes the following form:

$$s = (x_{ikn}, y_{11}, y_{12}, \ldots, y_{IN}),$$

where I is the number of patient types, x_{ikn} for $i = 1, \ldots, I$, $k = 1, \ldots, N$, $n = 1, \ldots, N$ is number of type-i patients with target date k scheduled to day n, y_{ik} for $i = 1, \ldots, I$, $k = 1, \ldots, N$ is number of type-i patients with target date k waiting to be scheduled.

3.1.2. Action sets

The decision to be made at the end of each day is to determine patients to book on specific days and diverted patients with specific target dates. $A_{x;y}$ denotes the set of available actions in state $s = (x; y)$. Any action in A_{xy} is represented by: $a = (a_{ikn}, z_{ik})$,
where a_{ikn} is the number of type-i patients with target date k to book on day n, z_{ik} is the number of diverted type-i patients with target date k.

The set of feasible actions must satisfy the following constraints:

$$\sum_{i=1}^{I}\sum_{k=1}^{N} x_{ikn} + \sum_{i=1}^{I}\sum_{k=1}^{N} a_{ikn} \leq C_1, n = 1, \ldots, N \quad (1)$$

$$\sum_{i=i}^{I}\sum_{k=1}^{N} z_{ik} \leq C_2 \quad (2)$$

$$\sum_{k=1}^{N}\sum_{n=1}^{N} a_{ikn} + \sum_{k=1}^{N} z_{ik} = \sum_{k=1}^{N} y_{ik}, i = 1, \ldots, I \quad (3)$$

where C_1 is daily resource capacity and C_2 is maximum number of patients diverted or served through overtime each day. Constraint 1 ensures that total number of patients scheduled on each day is limited by daily capacity. Constraint 2 dictates that total number of patients diverted (or served through overtime) is limited by diversion capacity. Constraint 3 ensures that all of the arriving patients must be either scheduled or diverted.

3.1.3. Transition probabilities

After all scheduling actions are taken, the number of new requests for each type of patient and number of treatments of each type that are cancelled affect the transition to the next state of the system.

Let d_{ikn} be the number of type-i patients with target date k booked on day n that are cancelled, and $Pr(q_{ik})$ be the probability that q_{ik} new requests for type-i patients arrived. The probability that d_{ikn} treatments for type-i patients with target date k scheduled to day n are cancelled is represented by the term $Q_{ikn}(d_{ikn})$, which is expressed as follows:

$$Q_{ikn}(d_{ikn}) = \binom{x_{ikn}}{d_{ikn}} p_{ikn}^{d_{ikn}} (1 - p_{ikn})^{x_{ikn} - d_{ikn}},$$

where p_{ikn} is the probability that a treatment of type i with target date k booked to day n is cancelled. On selecting decision a in state s, one component of the next state equals

$$(x_{ikn}^{'}, y_{ik}^{'}) = (x_{i,k,n+1} + a_{i,k,n+1} - d_{i,k,n+1}, y_{ik} - (\sum_{i=1}^{I}\sum_{n=1}^{N} a_{ikn} + \sum_{i=1}^{I}\sum_{k=1}^{N} z_{ik}) + q_{ik}))$$

with probability

$$Pr(q_{ik}) Q_{ikn}(d_{ikn})$$

Hence the state changes according to

$$P(s^{'}|s, a) = \prod_{i=1}^{I}\prod_{k=1}^{N} Pr(q_{ik}) \prod_{i=1}^{I}\prod_{n=1}^{N}\prod_{k=1}^{N} Q_{ikn}(d_{ikn}),$$

if $s^{'}$ satisfies Eqs. (4),(5), and (6).

$$x_{i,k,n}^{'} = x_{i,k,n+1} + a_{i,k,n+1} - d_{i,k,n+1}, \quad (4)$$
$$i = 1, \ldots, I; k = 1, \ldots, N; n = 1, \ldots, N-1$$

$$x_{i,k,N}^{'} = 0; \; i = 1, \ldots, I; k = 1, \ldots, N \quad (5)$$

$$y_{i,k}^{'} = y_{ik} - (\sum_{i=1}^{I}\sum_{n=1}^{N} a_{ikn} + \sum_{i=1}^{I}\sum_{k=1}^{N} z_{ik}) + q_{ik}, \quad (6)$$
$$i = 1, \ldots, I; k = 1, \ldots, N$$

Equations (4) and (5) show the new number of type-i patients with target date k booked on day n. Equation (6) defines the new number of type-i patients with target date k waiting to be scheduled.

3.1.4. Costs

As stated in [1], the cost of scheduling a type-i patient with target date k to day n is denoted $b(i, k, n)$. It is given by

$$b(i,k,n) = \begin{cases} 0, & \text{if } L_i \le n \le U_i \\ (L_i - n)c_1^i & \text{if } L_i \ge n \\ (n - U_i)c_2^i & \text{if } n > U_i, \end{cases} \quad (7)$$

where L_i for $i = 1, \ldots, I$ and U_i for $i = 1, \ldots, I$ are lower and upper tolerance limits for type-i patients, respectively. c_1^i for $i = 1, \ldots, I$ and c_2^i for $i = 1, \ldots, I$ are unit early and unit late costs for type-i patients, respectively. The immediate cost is expressed as

$$c(a,z) = \sum_{i,k,n} b(i,k,n)a_{ikn} + \sum_{i=1}^{I} d(i) \sum_{k=1}^{N} z_{ik} \quad (8)$$

where $d(i)$ for $i = 1, \ldots, I$ is a per unit penalty cost for diverting a type-i patient.

3.1.5. Bellman's equations

Discounting with discount factor λ is assumed in our model. Bellman's equations for finding a policy that minimizes the expected infinite horizon discounted cost are expressed as follows:

$$v(x,y) = \min_{(a,z)\in A_{x,y}} \left\{ c(a,z) + \lambda \textstyle\sum_{y' \in D}(y')v(x',y') \right\}, \quad (9)$$

where D is the set of all possible demand vectors. However, our MDP model is intractable owing to the fact that state and action spaces grow exponentially with the number of patient types and the length of the booking horizon. Therefore we resort to approximate dynamic programming (ADP) for approximately solving our model. The ADP technique we utilized in this research is described next.

4. Direct-search based approximate dynamic programming

ADP has been extensively used for solving intractable MDPs in diverse fields such as manufacturing, healthcare, and revenue management [20]. ADP techniques are mainly classified as linear programming (LP)-based ADPs and simulation-based ADPs. In the LP-based ADP approach, the underlying MDP model is transformed into the equivalent LP version of Bellman's equations, and then approximate value function is used to make the LP model tractable [21–23]. Whereas simulation-based ADP techniques find an approximate solution to the Bellman's equations by simulating the evolution of the system over a number of initial states in order to tune the parameters [24,25]. They employ simulation models such as statistical sampling and reinforcement learning methods for estimating the value functions [26].

In ADP, the value function is approximated through a combination of basis functions, which represent some important features of the state of the system. There are certain ways for doing so, and one of them is linear approximation, which takes the following form:

$$V(s) \approx \sum_{k=1}^{K} r_k \Phi_k(s),$$

where r_k for $k = 1, \ldots, K$ are tuning parameters and $\Phi_k(s)$ for $k = 1, \ldots, K$ are basis functions. After the value function is approximated, an ADP policy is obtained by tuning the approximation parameters iteratively. In particular, the goal of ADP techniques is to find the optimal parameter vector that minimizes a certain performance metric such as the sum of squared differences between the approximate cost-to-go function and the estimated cost-to-go function over sampled states [27]. The resulting optimization problem is generally solved using regression-based techniques [20, 24]. However, we utilize an approach that solves an optimization problem for tuning the ADP parameters to achieve the best policy resulting from those parameters. We then obtain the ADP policy using the approximate value functions.

4.1. Retrieving the ADP Policy from the Approximate Value Function

After the end of the parameter tuning phase which makes the approximate value of a given state available, the ADP policy is retrieved by computing a decision vector for any desired state of the system. As stated in [27], that decision vector is myopic with respect to the value function approximation of our MDP. The decision retrieval problem for a particular state s of our MDP model is given by

$$\min_{(a,z)\in A_{x,y}} \left\{ c(a,z) + \lambda \sum_{y' \in D} (y')\tilde{v}(x',y') \right\}, \quad (10)$$

where $\tilde{V}(x',y')$ is the approximate value of state (x',y').

Details about basis functions for the ADP implementation are discussed next.

4.2. Basis Functions

The basis function we chose for the ADP technique are as follows:

$$\Phi(s) = \sum_{k,n:n=k} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{1kn})^2/2$$
$$+\sum_{k=2}^{N}\sum_{n=1:n=k-1}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{2kn})^2/8$$
$$+\sum_{k=1}^{N}\sum_{n=1:n=k}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{2kn})^2/8$$
$$+\sum_{k=1}^{N-1}\sum_{n=1:n=k+1}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{2kn})^2/8$$
$$+\sum_{k=3}^{N}\sum_{n=1:n=k-2}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{3kn})^2/32$$
$$+\sum_{k=2}^{N}\sum_{n=1:n=k-1}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{3kn})^2/32$$
$$+\sum_{k=1}^{N}\sum_{n=1:n=k}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{3kn})^2/32$$
$$+\sum_{k=1}^{N-1}\sum_{n=1:n=k+1}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{3kn})^2/32$$
$$+\sum_{k=1}^{N-2}\sum_{n=1:n=k+2}^{N} (C_1 - \sum_{i=1}^{I}\sum_{kk=1}^{N} x_{i,kk,n} - a_{3kn})^2/32.$$

Our basis function utilizes available capacity for each day in retrieving the ADP policy and dictates that patients with lower tolerance limits have higher priority when scheduling is performed. The first term corresponds to available capacity for tolerance $(0,0)$, the next three terms correspond to available capacity for tolerance $(1,1)$, and the next five terms correspond to available capacity for tolerance $(2,2)$.

4.3. Direct Search

We tune parameters using direct search (see [27] and [28] for the implementation of direct search). Unlike regression-based techniques, direct search considers the ultimate goal of finding good policies. In particular, direct search deals with an optimization problem where the variables consist of feasible r's and the objective function value is the expected cost of the policy induced by the corresponding parameter vector [27]. The resulting optimization problem is expressed as follows:

$$\min_{r\in R^N} E\Big[\sum_{t=0}^{T} c(s_t, \pi_r(s_t))\Big], \qquad (11)$$

where T is a random variable denoting the final step of the search, s_t is the state at step t of the search, π_r is the policy obtained by the parameter vector r, $\pi_r(s_t)$ is the action dictated by the policy π_r in the state at stage t, and $c(s_t, \pi_r(s_t))$ is immediate cost incurred at step t as a result of choosing $\pi_r(s_t)$. Here, π_r is obtained by solving the aforementioned decision retrieval problem via the parameter vector r used to approximate the value function for each possible state visited during the search. The objective function in Eq. (11) is the expected cost of the policy π_r [27].

During the implementation of the direct search-based ADP, we let r range from 0 to 10000 in increments of 500. We choose the value based on the best policy performance when decisions are made by solving Equation 10.

5. Numerical results

This section focuses on the comparison of the performance of the direct search-based ADP policy with that of the myopic policy. The myopic policy is obtained by solving the problem $\min_{(a,z)\in A_{x,y}} c(a,z)$ for any state $s \in S$. We used AMPL with CPLEX 12 for solving all integer programs.

5.1. Experimental design

We consider a chemotherapy center where the length of the booking horizon is 10 (for larger problems the length of the booking horizon is set to 20). In line with the literature [1], we set the number of patient types to 3 and tolerance limits to (0,0), (1,1), and (2,2). We assume that arrival distribution is independent Poisson with mean 1 for each type for each day over the booking horizon. Probability that a scheduled treatment is cancelled has three levels: 0.01, 0.05, and 0.1. We assume capacity levels of 25 and 30; the former level corresponds to the low capacity case whereas the latter represents the medium capacity case owing to the fact that the total mean daily demand nearly equals 30 (for larger problems, those levels are set to 55 and 60.). In line with the literature [1], we set c_1^i, the unit early scheduling costs per day to 100, 75, and 50, whereas c_2^i, the unit late scheduling costs per day, is set to 125, 100, and 75.

We set diversion cost per patient for each patient type to 50 and 250. The case when the diversion cost equals 50 is referred by [1] as "rigid tolerance case", whereas other choices for the diversion costs are referred as "relaxed tolerance cases". We also consider the cases where early (late) scheduling costs are much higher than late (early) scheduling cost for the relaxed tolerance case. In line with the literature [1], we set the unit early scheduling cost per day for each patient type to 300, 225, and 150, respectively for

the high early scheduling cost case whereas the unit late scheduling cost per day for each patient type was set to 250, 200, and 150, respectively for the high late scheduling cost case. Discount factor is set to 0.99.

By changing the length of the booking horizon, cancellation probability, levels of capacity, diversion cost, and scheduling cost, we generated 32 problem sets. Simulation run lengths are set to 50. For each problem set, we ran 50 independent simulations.

5.2. Results

We present our results in Tables 1-8. Each row of the below tables represents a different problem set. For each problem, we utilized $Paired-t$ test to statistically identify whether the ADP policy or the myopic policy performs better. Total costs obtained by the ADP policy and the myopic policy are listed in the second and third column of each table, respectively. The percentage improvement obtained by the ADP policy over the myopic policy is given in the fourth column of each table. The final column of each table reveals whether there is a statistically significant improvement over the myopic policy.

We begin with the analysis of results for the rigid tolerance cases. As mentioned earlier, in the rigid tolerance case diversion cost per patient is 50. The results indicate that in reasonable-sized problems (i.e., $N = 10$), the ADP policy statistically performs better than the myopic policy except for the case where the level of capacity is 30 and cancellation probability is 0.5 (see Table 1). Additionally, the ADP policy generally performs better in low capacity situations as compared to medium capacity situations. It is also worth noting that the percentage difference between the two policies is significantly high in the case where the level of capacity is medium and cancellation probability is high. Finally, the superiority of the ADP policy over the myopic policy in the rigid tolerance case diminishes in large-sized problems (i.e., $N = 20$) (see Table 3).

In the relaxed tolerance case, the ADP policy significantly performs better than the myopic policy for reasonable-sized problems (see Table 2). The corresponding percentage improvement over the myopic policy is higher than that in the rigid tolerance case. Further, unlike in the rigid tolerance case, the performance of ADP increases when the level of capacity is increased from low to medium level. Yet, for large-sized problems, using ADP results in very small improvement when cancellation probability is small; if this probability is high, the performance of ADP and the myopic policy turns out to be the same (see Table 4).

The results for the high early scheduling cost case indicate that the performance of the ADP policy is significantly higher than that of the myopic policy for reasonable-sized problems (see Table 5). The corresponding percentage improvement over the myopic policy is higher as compared to that for the rigid tolerance case. However, the performance of the ADP policy for this case decreases when the problem size increases (see Table 6). We have similar observations for the high late scheduling cost case (see Tables 7 and 8).

Table 1. Results for the rigid tolerance case with $N = 10$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(25,0.01)	4141	4434	6.6	YES
(25,0.05)	2137	2266	5.7	YES
(25,0.1)	938	956	1.8	YES
(30,0.01)	1584	1655	4.3	YES
(30,0.05)	326	327	0.2	NO
(30,0.1)	5	17	72	YES

Table 2. Results for the relaxed tolerance case with $N = 10$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(25,0.01)	22911	25348	9.6	YES
(25,0.05)	9852	10625	7.3	YES
(25,0.1)	3719	3796	2	YES
(30,0.01)	6892	8141	15.3	YES
(30,0.05)	1110	1356	18.1	YES
(30,0.1)	10	36	72	YES

Table 3. Results for the rigid tolerance case with $N = 20$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(55,0.01)	16737	17110	2.2	YES
(55,0.05)	3064	3064	0	NO
(55,0.1)	1621	1621	0	NO
(60,0.01)	11936	12010	0.6	NO
(60,0.05)	214	214	0	NO
(60,0.1)	0.9	0.9	0	NO

Table 4. Results for the relaxed tolerance case with $N = 20$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(55,0.01)	86732	90672	4.3	YES
(55,0.05)	13310	13310	0	NO
(55,0.1)	7499	7499	0	NO
(60,0.01)	59620	62248	4.2	YES
(60,0.05)	856	856	0	NO
(60,0.1)	2	2	0	NO

Table 5. Results for the relaxed tolerance high early scheduling case with $N = 10$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(25,0.01)	22995	25360	9.3	YES
(25,0.05)	9947	10693	7.0	YES
(25,0.1)	3824	3899	1.9	YES
(30,0.01)	6969	8019	13.1	YES
(30,0.05)	1289	1365	5.6	YES
(30,0.1)	12	45	73	YES

Table 6. Results for the relaxed tolerance high early scheduling case with $N = 20$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(55,0.01)	87078	91287	4.6	YES
(55,0.05)	13976	13976	0	NO
(55,0.1)	7526	7526	0	NO
(60,0.01)	59976	62650	4.3	YES
(60,0.05)	1087	1087	0	NO
(60,0.1)	0.9	0.9	0	NO

Table 7. Results for the relaxed tolerance high late scheduling case with $N = 10$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(25,0.01)	21797	23610	7.7	YES
(25,0.05)	10436	11621	10.2	YES
(25,0.1)	4288	4394	2.4	YES
(30,0.01)	7672	8477	9.5	YES
(30,0.05)	1401	1782	21.4	YES
(30,0.1)	12	38	67	YES

Table 8. Results for the relaxed tolerance high late scheduling case with $N = 20$.

(C_1, p_{ikn})	ADP	Myopic	% impr.	Sign.?
(55,0.01)	84263	87521	3.7	YES
(55,0.05)	14320	14320	0	NO
(55,0.1)	7866	7866	0	NO
(60,0.01)	59242	61145	3.1	YES
(60,0.05)	990	990	0	NO
(60,0.1)	0.9	0.9	0	NO

6. Conclusion

We studied a chemotherapy appointment booking problem where patients have specific target dates, are classified based on their tolerance limits, and are scheduled to days in advance. Unlike the relevant literature, we considered cancellations of treatments. We provided an MDP formulation of this problem and because of huge state and action spaces, we approximately solved the problem using a direct search-based ADP.

Direct search-based ADP can be considered as a relatively new technique among other ADP techniques. In this research, we demonstrated that it can be a viable method for solving chemotherapy appointment booking problems. In particular, our work revealed that the performance of the myopic policy can be significantly improved through the implementation of direct search-based ADP. It is worth noting that further improvements may be achieved by trying various basis functions as part of direct search-based ADP.

Cancellation of the treatments in the chemotherapy appointment booking problem with target dates and tolerance limits were not considered in the relevant literature because of resulting computational challenges. By including cancellations into our model, we filled this gap in this area. As a future research, overbooking and no-shows can be incorporated into our model and direct search-based ADP can be utilized for solving other dynamic scheduling problems.

References

[1] Gocgun, Y., & Puterman, M. L. (2014). Dynamic scheduling with due dates and time windows: an application to chemotherapy patient appointment booking. *Health Care Manag Sci.*, 17, 60-76.

[2] Green, L. V., Savin, S., & Wang, B. (2006). Managing Patient Service in a Diagnostic Medical Facility. *Operations Research*, 54, 11-25.

[3] Cardoen, B., Demeulemeester, E., & Belien, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201, 921-932.

[4] Hulshof, P., Kortbeek, N., Boucherie, R., Hans, E., & Bakker, P. (2012). Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. *Health Systems*, 1, 129-175.

[5] Gocgun, Y., Bresnahan, B., Ghate, A., & Gunn, M. (2011). A markov decision process approach to multi-category patient scheduling in a diagnostic facility. *Artificial Intelligence in Medicine*, 53, 73-81.

[6] Kolisch,R., & Sickinger, S. (2008). Providing radiology health care services to stochastic demand of different customer classes. *OR Spectrum*, 30, 375-395.

[7] Min, D., & Yih, Y. (2010). An elective surgery scheduling problem considering patient priority. *Computers and Operations Research*, 37, 1091-1099.

[8] Patrick, J., Puterman, M. L., & Queyranne, M. (2008). Dynamic multi-priority patient scheduling for a diagnostic resource. *Operations Research*, 56, 1507-1525.

[9] Lamiri M., Xie X., Dolgui, A., & Grimaud, F. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research*, 185, 1026-1037.

[10] Liu, N., Ziya, S. & Kulkarni, V. G. (2010). Dynamic scheduling of outpatient appointments under patient no-shows and cancellation. *Manufacturing and Service Operations Management*, 12, 347-364.

[11] Saure, A., Patrick J., Tyldesley, S., & Puterman, M. L. (2012). Dynamic multi-appointment patient scheduling for radiation therapy. *European Journal of Operational Research*, 223, 573-584.

[12] Geng, N., & Xiaolan, X. (2016). Optimal dynamic outpatient scheduling for a diagnostic facility with two waiting time targets. *IEEE Transactions on Automatic Control*, 61, 3725-3739.

[13] Tsai, P. J., & Teng, G. (2014). A stochastic appointment scheduling system on multiple resources with dynamic call-in sequence and patient no-shows for an outpatient clinic. *European Journal of Operational Research*, 239, 427-436.

[14] Turkcan, A., Zeng, B., & Lawley, M. (2012). Chemotherapy operations planning and scheduling. *IIE Transactions on Healthcare Systems Engineering*, 2, 31-49.

[15] Goldberg, S., Carter, M., Beck, J., Trudeau, M., Sousa, P., & Beattie, K. (2014). Dynamic optimization of chemotherapy outpatient scheduling with uncertainty. *Healthcare Management Science*, 17, 379-392.

[16] Goldberg, S., Beck, J., Carter, M., Trudeau, M., Sousa, P., & Beattie, K. (2014). Solving the chemotherapy outpatient scheduling problem with constraint programming. *Journal of Applied Operational Research*, 6, 135-144.

[17] Alvarado, M. & Ntaimo, L. (2018). Chemotherapy appointment scheduling under uncertainty using mean-risk stochastic integer programming. *Healthcare Management Science*, 21, 87-104.

[18] Parizi, M. S., & Ghate, A. (2016). Multi-class, multi-resource advance scheduling with no-shows, cancellations and overbooking. Computers Operations Research, 67, 90-101.

[19] Kleywegt, A. J., & Papastavrou, J. D. (1998). The Dynamic and Stochastic Knapsack Problem. *Operations Research*, 46, 17-35.

[20] Powell W. B. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley and Sons.

[21] Adelman, D. (2003). Price-directed replenishment of subsets: methodology and its application to inventory routing. *Manufacturing and Service Operations Management*, 5, 348-371.

[22] Adelman, D. (2004). A price-directed approach to stochastic inventory routing. *Operations Research*, 52, 499-514.

[23] De Farias, D. P., & Roy, B. V. (2004). The linear programming approach to Approximate Dynamic Programming. *Operations Research*, 51, 850-865.

[24] Bertsekas, D., & Tsitsiklis, J. (1996). *Neuro-Dynamic Programming*. Athena Scientific.

[25] Sutton, R. S. & Barto, A. G. (1998). *Reinforcement Learning*. MIT Press.

[26] Chang, H. S., Fu, M. C., Hu, J., & Marcus, S. I. (2007). *Simulation-based algorithms for Markov Decision Processes*. Springer.

[27] Gocgun, Y. (2018). Approximate dynamic programming for optimal search with an obstacle (Submitted).

[28] Maxwell, M. S., Henderson, S. G., & Topaloglu, H. (2013). Tuning Approximate Dynamic Programming Policies for Ambulance Redeployment via Direct Search. *Stochastic Systems*, 3, 1-40.

Yasin Göçgün *earned his B.S. degree in industrial engineering from Bilkent University in Turkey in 2003. He earned his M.S. degree in industrial engineering from the same university in 2005. Dr. Gocgun earned his Ph.D. degree in industrial engineering from the University of Washington in 2010. He then worked as a postdoctoral fellow at the University of British Columbia in Vancouver, Canada for two years. After his first post-doc, he worked as a postdoctoral fellow at the University of Toronto in Toronto, Canada for two years. He has been working as an assistant professor in the Department of Industrial Engineering at Altinbas University in Istanbul, Turkey since September 2014. His research interests are stochastic dynamic optimization, operations research in healthcare, Markov decision processes, approximate dynamic programming, scheduling, and simulation.*

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.2, pp.170-175 (2018)
https://doi.org/10.11121/ijocta.01.2018.00567

RESEARCH ARTICLE

Heartbeat type classification with optimized feature vectors

Özal Yıldırım*, Ulas Baran Baloglu

Department of Computer Engineering, Munzur University, Turkey
oyildirim@munzur.edu.tr, ulasbaloglu@gmail.com

ARTICLE INFO

Article history:
Received: 16 December 2017
Accepted: 7 March 2018
Available Online: 12 April 2018

Keywords:
ECG signals
Feature optimization
Feature vectors
Classification
Machine learning

AMS Classification 2010:
68TXX, 62-07, 92-08

ABSTRACT

In this study, a feature vector optimization based method has been proposed for classification of the heartbeat types. Electrocardiogram (ECG) signals of five different heartbeat type were used for this aim. Firstly, wavelet transform (WT) method were applied on these ECG signals to generate all feature vectors. Optimizing these feature vectors is provided by performing particle swarm optimization (PSO), genetic search, best first, greedy stepwise and multi objective evoluationary algorithms on these vectors. These optimized feature vectors are later applied to the classifier inputs for performance evaluation. A comprehensive assessment was presented for the determination of optimized feature vectors for ECG signals and best-performing classifier for these optimized feature vectors was determined.

1. Introduction

Electrocardiogram (ECG) signals have a critical importance for determining abnormal heart conditions. Computer-aided analysis of these data is particularly considerable regarding the development of smart medical platforms. For this purpose, it is necessary to automatically detecting and recognizing heartbeats from ECG signals.

Heartbeat detection process from the ECG signals is not easy to realize because of the various types of noise, which exist in the ECG signals [1]. Different techniques have been comprehensively analyzed for beat detection such as digital filter usage [2, 3], the wavelet transform (WT) [4-7], neural networks (NN) [8, 9], hidden markov model [10] and particle swarm optimization (PSO) [11]. Detected heartbeat signals are segmented to be used in the classification systems.

There are two main stages in the ECG signal classification. These are feature extraction and classification stages. In the feature extraction stage, distinguishing features of ECG signals are revealed. These features are used to generate feature vectors for each signal. Feature extraction is required to remove unnecessary, noisy or corrupted inputs. This stage improves the accuracy of classifiers which are used in the consequent classification stage. In the classification stage, a suitable classifier was trained on the obtained features.

WT methods such as multiresolution and discrete types [12-16] are frequently used for feature extraction process. There are also other studies where higher order statistic [17] and mathematical morphology method [18] were employed. In the classification stage, different classifiers can be applied such as extreme learning machines [19] and support vector machines [20-22]. NN classifier also was widely used in classification studies [23-25].

In this study, feature vector optimization approach was used to classify heartbeat signals. The PSO search [26], genetic search [27], best first-greedy stepwise [28] and multi-objective evolutionary search [29] methods are applied on these feature vectors for reducing the computational complexity of the overall process. The optimized vectors were passed as inputs to the classifier algorithms. Random forests (RF) [30] and least square support vector machines (LS-SVM) [31] classifiers were selected for determining the heartbeat types. Experimental results were conducted on an ECG dataset which includes five different heartbeat type. These ECG signals are collected from the MIT-BIH arrhythmia database.

The rest of this paper is organized as follows: Section 2 briefly presents the feature optimization problem. Section 3 describes the proposed method in detail. Section 4 presents the experimental results and Section 5 concludes the paper.

*Corresponding author

2. Feature optimization

Feature optimization is one of the most significant challenges in data analysis studies due to the huge volume of data to be processed. Feature selection and optimization reduce the dimension of the data by removing unnecessary features so that these processes improve the performance of algorithms. There are variously supervised, semi-supervised and unsupervised feature optimization techniques in the literature.

The main idea of this optimization process is gathering a subset of existing features by eliminating features which are containing relatively little information. The relevance relation of a set and the target class should be defined to facilitate the feature optimization. Let X set denotes the input features and Y set denotes the relevant classes.

$$X = \{x_1, x_2, ..., x_m\},\ Y = \{y_1, y_2, ..., y_n\} \quad (1)$$

Among $\{x, y\}$ pairs, the objective function of feature selection is finding a subset of pairs which can be defined as follows,

$$S(\overline{X}) = \max_{|\Upsilon|<m,(\Upsilon\subset X)} O(\Upsilon) \quad (2)$$

In equation (2), O function is the feature optimization function which calculates the accuracy of a feature subset. In feature optimization, the search space contains all the possible subsets of features so that feature optimization can significantly affect the performance of ECG signal classification. Consequently, feature selection and optimization problem is an NP-hard problem [32] so that meta-heuristics such as evolutionary algorithms are frequently considered in creating a solution space when there is a large number of features [33]. It is essential to determine the best techniques for specific tasks such as beat detection and recognition. In this study, we analyze ECG signal with various optimization methods and classifier algorithms.

3. The proposed method

3.1 Feature vectors

The first step in the process of determining the optimal feature vectors for ECG signals is to make feature deductions on these signals. For this purpose, 6 level Dabuchies (Db6) [34] wavelet transform method was used in the study. By using wavelet transform, coefficient matrices are obtained for the signals separated into lower frequency bands. The steps of this process are shown in detail in Figure 1.

In the wavelet transform process, input signals are passed through high-pass and low-pass filters after each conversion. These filters provide a detailed analysis of high and low-frequency components of the signal. As a result of the wavelet transform, approximation (A^n) and detail (D^n) coefficients are formed at each level of the input signal.

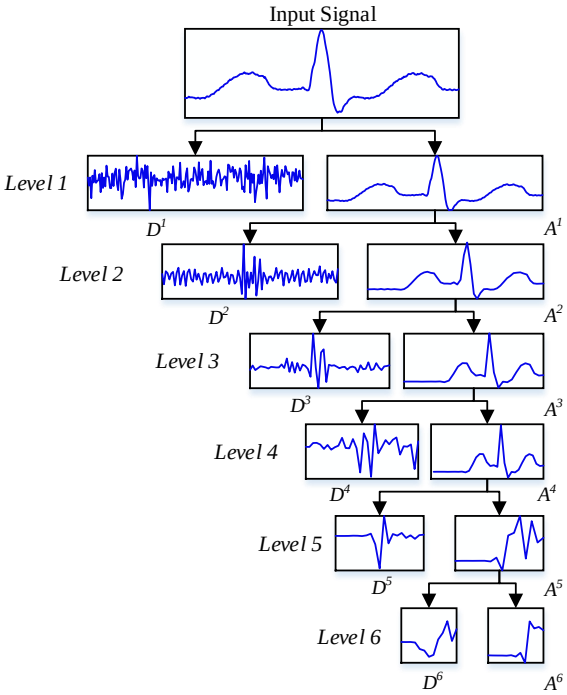


Figure 1. Wavelet transform steps for 6 levels.

Coefficient matrices do not have an appropriate use because of the large size data classifiers they contain. For this reason, data in these coefficient matrices have to be reduced to lower dimensional representing data. Statistical methods were commonly employed for this purpose. In this study, energy, mean, standard deviation and norm entropy methods were used. Energy calculation on coefficient matrices can be defined as:

$$E^i = \sum_{k=1}^{N} \left(C_k^i\right)^2 \quad \text{for } i=1...M \quad (3)$$

where C denotes coefficient matrices, and N is the size of these matrices. Finally, M denotes the number of the sub-bands. The average of coefficient matrices is calculated as follows,

$$\mu^i = \frac{1}{N}\sum_{k=1}^{N} C_k^i \quad (4)$$

Thus, an average value is obtained for the coefficient matrices of each frequency sub-band. Another employed method, standard deviation calculation, is as follows,

$$\sigma^i = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(C_k^i - \mu^i\right)^2} \quad (5)$$

Lastly, the norm entropy calculation for each coefficient matrix is obtained by the following equation.

$$\rho^i = \sum_{k=1}^{N}\left(C_k^i\right)^P 1 \le P \quad (6)$$

The coefficient matrix for each input signal consists of a total of 7 coefficient matrices as:

$C^i=[D_1,D_2,D_3,D_4,D_5,D_6,A_6]$

All of these calculations generate a 28-dimensional feature vector for each signal. This 28-dimensional feature vector can be defined as:

$$V^k = \left[E_j^k, \mu_j^k, \sigma_j^k, \rho_j^k \right] \text{ for } j=1 \ldots 7 \text{ and } k=1 \ldots N \qquad (7)$$

where k is number of feature vectors, and j is number of sub-bands respectively.

3.2. ECG dataset

MIT-BIH arrhythmia [35] database, which is widely used in the literature, was selected as ECG data source in experimental studies. Within this database, there are 48 different records and beats of these records were labelled by experts. Normal, Premature Ventricular Contraction (PVC), Paced, Left Bundle Branch Block (LBBB), and Right Bundle Branch Block (RBBB) heartbeats are selected from the data. In Figure 2 the signal form of a normal heartbeat is given.

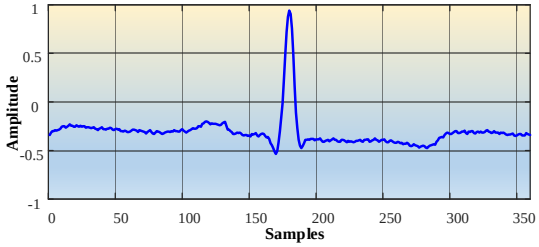


Figure 2. Normal heartbeat ECG signal example.

PVC is an abnormal condition that the heartbeat is initiated by ventricular Purkinje fibers rather than by the sinoatrial node, which is the normal heartbeat initiator. As a result, extra contractions occur, and the regular heart rhythm breaks down. An illustration of the PVC signal is given in Figure 3.

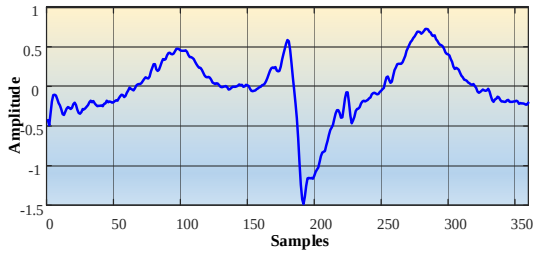


Figure 3. PVC signal form example.

An example ECG signal for paced, another abnormal heartbeat, is shown in Figure 4.

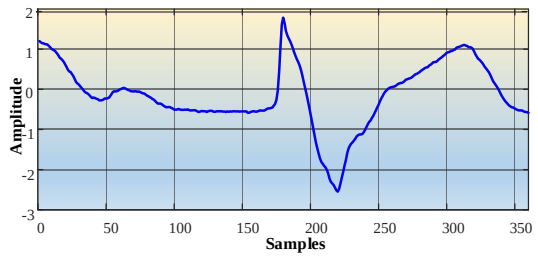


Figure 4. Paced arrhythmia signal example.

The Bundle branch block is a delay in the way of electrical impulses which are ejected to provide a heartbeat. This delay can occur in the right or left ventricles of the heart. If this delay happens in the right ventricles, the RBBB heartbeat shown in Figure 5 occurs, and if this delay happens in the left ventricle, the LBBB heartbeat shown in Figure 6 occurs.

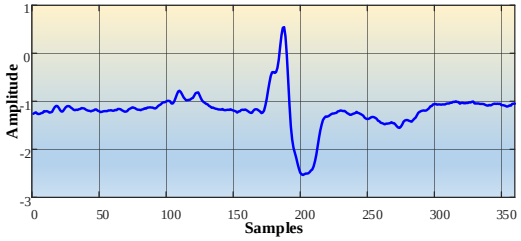


Figure 5. RBBB arrhythmia signal example.

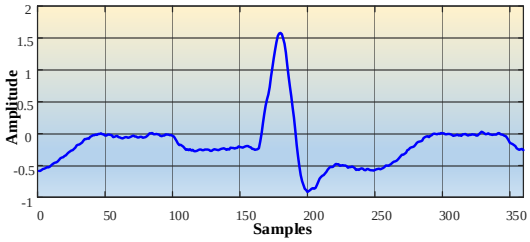


Figure 6. LBBB arrhythmia signal example.

4. Experimental results

In order to determine the optimal features for ECG signals, five different heartbeat classes were selected from the MIT-BIH arrhythmia database. The 50% of the data were used in the training phase while the rest were used for testing. The numerical distributions of data classes are given in Table 1.

Table 1. Number of used ECG signals.

Classes	Train Data	Test Data	Total
Normal	2500	2500	5000
Paced	950	950	1900
PVC	250	250	500
RBBB	2250	2250	4500
LBBB	950	950	1900

Statistical methods have been used on 6-levels Db6 wavelet transform coefficients to generate 28-dimensional feature vectors of ECG signals. The properties and definitions in feature vectors are given in Table 2.

Feature vectors that have 28-dimensional for each heartbeat signal are available as input data. PSO, genetic search, best first, greedy stepwise and multi objective evolutionary algorithms are used for feature optimization. The optimal features obtained after applying these methods to the feature vectors are given in Table 3.

PSO, best first and greedy stepwise methods were

determined 17 feature for this signals. The lowest number of features is determined by the genetic search method while the multi objective evolutionary search algorithm determines nine features. We have tested the classification accuracy with feature vectors which are optimized for evaluating the performance of the optimization algorithms. For this purpose, LS-SVM and RF classifiers were used in the classification of ECG signals. The recognition accuracy is determined by applying the feature vectors obtained from the optimization algorithms to these classifier inputs.

Table 2. Representation of the features.

Number	Feature	Number	Feature
1…6	E (D_1…D_6)	*7*	E (A_6)
8…13	μ (D_1…D_6)	*14*	μ (A_6)
15…20	σ (D_1…D_6)	*21*	σ (A_6)
22…27	ρ (D_1…D_6)	*28*	ρ (A_6)

E: Energy, μ: Mean, σ: Standard deviation, ρ: Norm entropy, D: Detail band, A: Approximation band.

Table 3. Optimized feature vectors and selected features.

Method	Optimized Feature Vectors	Vector Size
Particle Swarm Optimization	3-5-8-9-10-11-12-13-15-16-18-23-24-25-26-27-29	17
Best First and Greedy Stepwise	2-3-6-9-10-11-12-13-15-17-18-23-24-25-26-27-29	17
Genetic Search	13-15-21-23-25-26-27-28	8
Multi Objective Evolutionary Search	2-3-4-6-10-23-25-26-28	9

Table 4. Performances of the classifiers on the feature vectors.

Feature Vector	Number of Features	Classifiers	
		LS-SVM	RF
All Features	28	97.30%	98.82%
Particle Swarm Optimization	17	97.34%	98.82%
Best First and Greedy Stepwise	17	97.00%	96.79%
Genetic Search	8	88.52%	**98.95%**
Multi Objective Evoluationary Search	9	80.84%	97.39%

As seen in Table 4, the best performance regarding both feature size and classification performance has been obtained by the genetic search method. RF classifier provides 98.5% performance on ECG datasets having eight features and being optimized by using genetic search. The RF classifier achieved 98.82% performance when the 28-dimensional feature vector containing all the features was given as input. As a result, both feature size has been drastically reduced, and the performance has increased. The PSO algorithm increased the performance from 97.30% to 97.34% with the LS-SVM classifier and reduced 11 features from the feature vector. Other optimization techniques have reduced feature size, but at the same time, they also reduced the performance.

5. Conclusion

In this study, feature vector optimization methods were analyzed in the classification of ECG signals. Five different heartbeat classes from the MIT-BIH arrhythmia dataset were used in experimental studies. Feature vectors containing distinctive features of signals are obtained by using wavelet transform and statistical methods on heartbeat signals. Various optimization algorithms have been used to optimize the 28-dimensional feature vector. Feature vectors obtained from these optimization algorithms are given as input to LS-SVM and RF classifiers. The feature vector, which is a total of 28 dimensions, was reduced to 8 dimensions as a result of genetic search optimization algorithm, resulting in 98.95% performance. When 28 features are used, it is seen that this performance is 98.82%. With the genetic search

optimization algorithm, both the feature vector is reduced, and the recognition performance is improved. In addition, the number of feature vectors is reduced by the PSO algorithm, and the recognition performance is preserved. Another important point is the selection of the classifier. The success achieved by the genetic search algorithm with the LS-SVM classifier was as low as 88.52%, but the success rate with the RF classifier increased to 98.95%.

As a result of experimental studies, it has been observed that the feature vectors significantly affect the performance in recognizing ECG signals. Further, it was shown that how the selected classifier can lead to a better performance on optimized feature vectors.

References

[1] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE transactions on biomedical engineering*, (3), 230-236.

[2] Okada, M. (1979). A digital filter for the ors complex detection. *IEEE Transactions on Biomedical Engineering*, (12), 700-703.

[3] Afonso, V. X., Tompkins, W. J., Nguyen, T. Q., & Luo, S. (1999). ECG beat detection using filter banks. *IEEE transactions on biomedical engineering*, 46(2), 192-202.

[4] Li, C., Zheng, C., & Tai, C. (1995). Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on biomedical Engineering*, 42(1), 21-28.

[5] Rekik, S., & Ellouze, N. (2016). QRS detection combining entropic criterion and wavelet transform. *International Journal of Signal and Imaging Systems Engineering*, 9(4-5), 299-304.

[6] Rani, R., Chouhan, V. S., & Sinha, H. P. (2015). Automated detection of qrs complex in ECG signal using wavelet transform. *International Journal of Computer Science and Network Security (IJCSNS)*, 15(1), 1.

[7] Kaur, I., Rajni, R., & Marwaha, A. (2016). ECG Signal Analysis and Arrhythmia Detection using Wavelet Transform. *Journal of The Institution of Engineers (India): Series B*, 97(4), 499-507.

[8] Chen, T. H., Zheng, Y., Han, L. Q., Guo, P. Y., & He, X. Y. (2008). The Sorting Method of ECG Signals Based on Neural Network. *In Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on* (pp. 543-546). *IEEE*.

[9] Dokur, Z., Ölmez, T., Yazgan, E., & Ersoy, O. K. (1997). Detection of ECG waveforms by neural networks. *Medical engineering & physics*, 19(8), 738-741.

[10] Coast, D. A., Stern, R. M., Cano, G. G., & Briller, S. A. (1990). An approach to cardiac arrhythmia analysis using hidden Markov models. *IEEE Transactions on biomedical Engineering*, 37(9), 826-836.

[11] Jain, S., Kumar, A., & Bajaj, V. (2016). Technique for QRS complex detection using particle swarm optimisation. *IET Science, Measurement & Technology*, 10(6), 626-636.

[12] Thomas, M., Das, M. K., & Ari, S. (2015). Automatic ECG arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69(4), 715-721..

[13] Inan, O. T., Giovangrandi, L., & Kovacs, G. T. (2006). Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE Transactions on Biomedical Engineering*, 53(12), 2507-2515.

[14] Sahoo, S., Kanungo, B., Behera, S., & Sabut, S. (2017). Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement*, 108, 55-66.

[15] Martis, R. J., Acharya, U. R., & Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomedical Signal Processing and Control*, 8(5), 437-448.

[16] Ince, T., Kiranyaz, S., & Gabbouj, M. (2009). A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Transactions on Biomedical Engineering*, 56(5), 1415-1426.

[17] Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2013). Cardiac decision making using higher order spectra. *Biomedical Signal Processing and Control*, 8(2), 193-203.

[18] Tadejko, P., & Rakowski, W. (2007). Mathematical morphology based ECG feature extraction for the purpose of heartbeat classification. *In Computer Information Systems and Industrial Management Applications, 6th International Conference on* (pp. 322-327). *IEEE*.

[19] Kim, J., Shin, H., Lee, Y., & Lee, M. (2007). Algorithm for classifying arrhythmia using Extreme Learning Machine and principal component analysis. *In Engineering in Medicine and Biology Society, EMBS 2007. 29th Annual International Conference of the IEEE* (pp. 3257-3260). *IEEE*.

[20] Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2012). Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. *Expert Systems with Applications*, 39(14), 11792-11800.

[21] Mehta, S. S., & Lingayat, N. S. (2008). Development of SVM based ECG Pattern Recognition Technique. *IETE Journal of Research*, 54(1), 5-11.

[22] Raman, P., & Ghosh, S. (2016). Classification of Heart Diseases based on ECG analysis using FCM and SVM Methods. *International Journal of*

Engineering Science, 6739.

[23] Ceylan, R., & Özbay, Y. (2007). Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. *Expert Systems with Applications*, 33(2), 286-295.

[24] Shadmand, S., & Mashoufi, B. (2016). A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization. Biomedical Signal Processing and Control, 25, 12-23.

[25] Güler, İ., & Übeylı, E. D. (2005). ECG beat classifier designed by combined neural network model. *Pattern recognition*, 38(2), 199-208.

[26] Moraglio, A., Di Chio, C., & Poli, R. (2007). Geometric particle swarm optimisation. *In European conference on genetic programming* (pp. 125-136). Springer, Berlin, Heidelberg.

[27] Gutlein, M., Frank, E., Hall, M., & Karwath, A. (2009). Large-scale attribute selection using wrappers. *In Computational Intelligence and Data Mining, IEEE Symposium on* (pp. 332-339). *IEEE*.

[28] Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine learning*, 3(2), 95-99.

[29] Jiménez, F., Sánchez, G., García, J. M., Sciavicco, G., & Miralles, L. (2017). Multi-objective evolutionary feature selection for online sales forecasting. *Neurocomputing*, 234, 75-92.

[30] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[31] Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.

[32] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182..

[33] Martín-Smith, P., Ortega, J., Asensio-Cubero, J., Gan, J. Q., & Ortiz, A. (2017). A supervised filter method for multi-objective feature selection in EEG classification based on multi-resolution analysis for BCI. *Neurocomputing*, 250, 45-56.

[34] Mark, R. Moody, G. (1997). MIT-BIH Arrhythmia Database, *http://ecg.mit.edu/dbinfo.html*

Özal Yıldırım *received BSc degree in Electrical and Computer Education, Fırat University, Turkey, in 2006. He received MSc degree in Computer Engineering Department from Fırat University, in 2010. He was awarded PhD degree from Electrical and Electronics Engineering from Fırat University, in 2015. Dr. Yıldırım is working as assistant professor in Computer Engineering Department at Munzur University, Turkey. His main research interests include signal processing, power quality, artificial intelligence and deep learning.*

Ulas Baran Baloglu *received BSc degree in Computer Engineering from I.D. Bilkent University, Ankara, Turkey and the MSc degree in Computer Engineering from Firat University, Elazig, Turkey. He was awarded PhD degree from Electrical and Electronics Engineering from Firat University in 2017. He is working as an assistant professor at the department of Computer Engineering, Munzur University, Turkey. His research interests include multi-agent systems, smart grids, data structures and optimization algorithms. He is IEEE volunteer member.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.176-182 (2018)
http://doi.org/10.11121/ijocta.01.2018.00544

RESEARCH ARTICLE

A conformable calculus of radial basis functions and its applications

Fuat Usta

Department of Mathematics, Faculty of Science and Arts, Düzce University, Düzce, Turkey
fuatusta@duzce.edu.tr

ARTICLE INFO

Article History:
Received 03 October 2017
Accepted 06 April 2018
Available 22 April 2018

Keywords:
Conformable fractional derivative
Radial basis functions
Kansa collocation technique

AMS Classification 2010:
65L60, 26A33

ABSTRACT

In this paper we introduced the conformable derivatives and integrals of radial basis functions (RBF) to solve conformable fractional differential equations via RBF collocation method. For that, firstly, we found the conformable derivatives and integrals of power, Gaussian and multiquadric basis functions utilizing the rule of conformable fractional calculus. Then by using these derivatives and integrals we provide a numerical scheme to solve conformable fractional differential equations. Finally we presents some numerical results to confirmed our method.

1. Introduction

Recently, the question of how to take non-integer order of derivative or integration was phenomenon among the scientists. However together with the development of mathematics knowledge, this question was answered via Fractional Calculus which is a generalization of ordinary differentiation and integration to arbitrary (non-integer) order. Then In conjunction with the development of theoretical progress of fractional calculus, a number of mathematicians have started to applied the obtained results to real world problems consist of fractional derivatives and integrals [1, 2].

An significant point is that the fractional derivative at a point x is a local property only when a is an integer; in non-integer cases we cannot say that the fractional derivative at x of a function f depends only on values of f very near x, in the way that integer-power derivatives certainly do. Therefore it is expected that the theory involves some sort of boundary conditions, involving information on the function further out. To use a metaphor, the fractional derivative requires some peripheral vision. As far as the existence of such a theory is concerned, the foundations of the subject were laid by Liouville in a paper from 1832. The fractional derivative of a function to order a is often now defined by means of the Fourier or Mellin integral transforms. Various types of fractional derivatives were introduced: Riemann- Liouville, Caputo, Hadamard, Erdelyi-Kober, Grunwald-Letnikov, Marchaud and Riesz are just a few to name [3, 4].

Now, all these definitions satisfy the property that the fractional derivative is linear. This is the only property inherited from the first derivative by all of the definitions. However, all definitions do not provide some properties such as Product Rule (Leibniz Rule), Quotient Rule, Chain Rule, Rolls Theorem and Mean Value Theorem. In addition most of the fractional derivatives except Caputo-type derivatives, do not satisfy $D^{\alpha}(f)(1) = 0$ if α is not a natural number.

Recently, a new local, limit-based definition of a so-called conformable derivative has been formulated in [5, 6], with several follow-up papers [1, 2, 7–16]. This new idea was quickly generalized by Katugampola [17, 18]. This new definition forms the basis for this work and is referred to here as the Conformable derivative (D_{α} will henceforth be referring to the Conformable derivative). This definition has several practical properties which are summarized below.

Note that if f is fully differentiable at t; then the derivative is $D_\alpha(f)(t) = t^{1-\alpha}f'(t)$. (Here, operators of a very similar form, $t^\alpha D_1$, have been applied in combinatorial theory [18]). Of course, for $t = 0$ this is not valid and it would be useful to deal with equations and solutions with singularities. Additionally it must be noted that conformable derivative is conformable at $\alpha = 1$, as

$$\lim_{\alpha\to 1} D^\alpha(f) = f',$$

but

$$\lim_{\alpha\to 0^+} D^\alpha(f) \neq f'.$$

On the other hand radial basis functions method is one of the more practical ways of solving fractional order of models. The most significant property of an RBF technique is that there is no need to generate any mesh so it called meshfree method. One only requires the pairwise distance between points for an RBF approximation. Therefore it can be easily applied to high dimensional problems since the computation of distance in any dimensions is straightforward. On the other hand in order to solve partial differential equations (PDEs) in [19, 20] Kansa proposed RBF collocation method which is meshfree and easy-to-handle in comparison with the other methods. Not only integer order PDEs [21] but also Kansa's approach has been used fractional order of PDEs [22].

In this paper we find the conformable derivatives and integrals of needed function of RBF interpolation such as powers, Gaussians and multiquadric. This derivatives play a significant role in the numerical solution of conformable differential equations by the help of RBF method.

The remainder of this work is organized as follows: In Section 2, the related definitions and theorems are summarised. In Section 3, the conformable derivative and integrals have been obtained for the radial basis functions which will use in the RBF computations. Numerical experiments are given in Section 4, while some conclusions and further directions of research are discussed in Section 5.

2. Preliminaries

2.1. Review of fractional derivatives and integrals

Here we review the Riemann-Liouville fractional derivatives and integrals introduced in [3, 4, 23].

Definition 1. *The left-sided Riemann-Liouville fractional derivative of order α of function $u(t)$ is described as*

$${}^\alpha\mathcal{D}_a^t u(t) = \frac{1}{\Gamma(\tau-\alpha)}\int_a^t (t-\xi)^{\tau-\alpha-1}u(\xi)d\xi, t > a$$

where $\tau = \lceil\alpha\rceil$.

Definition 2. *The right-sided Riemann-Liouville fractional derivative of order α of function $u(t)$ is described as*

$${}^\alpha\mathcal{D}_t^b u(t) = \frac{(-1)^\tau}{\Gamma(\tau-\alpha)}\int_t^b (\xi-t)^{\tau-\alpha-1}u(\xi)d\xi, t < b$$

where $\tau = \lceil\alpha\rceil$.

Definition 3. *The left-sided Riemann-Liouville fractional integral of order α of function $u(t)$ is described as*

$${}^\alpha\mathcal{I}_a^t u(t) = \frac{1}{\Gamma(\alpha)}\int_a^t (t-\xi)^{\alpha-1}u(\xi)d\xi, \quad t > a$$

Definition 4. *The right-sided Riemann-Liouville fractional integral of order α of function $u(t)$ is described as*

$${}^\alpha\mathcal{I}_t^b u(t) = \frac{1}{\Gamma(\alpha)}\int_t^b (\xi-t)^{\alpha-1}u(\xi)d\xi, \quad t < b$$

Then Khalil et.al. [6] have introduced the conformable fractional derivative and integrals by following definition.

Definition 5. *Let $u : [0,\infty) \to \mathbb{R}$. The conformable derivative of $u(t)$ of order α described by*

$${}^\alpha\mathfrak{D}u(t) = \lim_{\eta\to 0}\frac{u(t+\eta t^{1-\alpha}) - u(t)}{\eta}$$

where $\alpha \in (0,1)$ and for all $t > 0$. In other words if $u(t)$ is differentiable, then

$${}^\alpha\mathfrak{D}u(t) = t^{1-\alpha}f'(t),$$

where prime denotes the classical derivative operator.

Similarly, one can define the conformable fractional integral operator.

Definition 6. *Let $u : [0,\infty) \to \mathbb{R}$. The left sided conformable integral of $u(t)$ of order α described by*

$${}^\alpha\mathfrak{I}_a^t u(t) = \int_a^t t^{\alpha-1}u(t)dt, \qquad t > a$$

where $\alpha \in (0,1)$ and the integral is classical integral operator.

Definition 7. *Let $u : [0,\infty) \to \mathbb{R}$. The right sided conformable integral of $u(t)$ of order α described by*

$${}^\alpha\mathfrak{I}_t^b u(t) = \int_t^b (-t)^{\alpha-1}u(t)dt, \qquad t < b$$

where $\alpha \in (0,1)$ and the integral is classical integral operator.

2.2. Radial basis function method

One of the properly approach to solving PDE is radial basis functions (RBFs). The main idea of the RBFs is to calculate distance to any fixed center points x_i with the form $\varphi(\|x-x_i\|_2)$. Additionally RBF may also have scaling parameter called shape parameter ε. This can be done in the manner that $\varphi(r)$ is replaced by $\varphi(\varepsilon r)$. Generally shape parameter have been chosen arbitrarily because there are no exact consequence about how to choose best shape parameter. Some of the RBFs are listed in Table 1.

Table 1. Radial basis functions.

RBFs	$\varphi(r)$
Multiquadric (MQ)	$\sqrt{1+r^2}$
Inverse Multiquadric (IMQ)	$\frac{1}{\sqrt{1+r^2}}$
Inverse Quadratic (IQ)	$\frac{1}{1+r^2}$
Gaussian (GA)	e^{-r^2}

The main advantageous of RBF technique is that it does not require any mesh hence it called meshfree method. Therefore the RBF interpolation can be represent as a linear combination of RBFs as follows:

$$s=\sum_{i=1}^{N} a_i\varphi(\|x-x_i\|_2)$$

where the a_i's the coefficients which are usually calculated by collocation technique. Some of the greatest advantages of RBF interpolation method lies in its practicality in almost any dimension and their fast convergence to the approximated target function.

3. Conformable derivatives of RBFs in one dimension

In order to construct conformable derivatives and integrals we will make use of the fractional calculus. Namely the relationship between Riemann-Liouville fractional integral and conformable fractional integral can be given as follows:

Definition 8. *Let $\alpha \in (\epsilon, \epsilon+1]$, then the left sided relationship between Riemann-Liouville fractional integral and conformable fractional integral is*

$${}^{\alpha}\mathfrak{I}_a^t u(t) = {}^{\epsilon+1}\mathcal{I}_a^t\{(t-a)^{\theta-1}u(t)\}$$

Here if $\alpha=\epsilon+1$ then $\theta=1$ since $\theta=\alpha-\epsilon$.

Theorem 1. *Let $\theta>-1$ and $t>a$*

$${}^{\alpha}\mathfrak{I}_a^t(t-a)^{\gamma}=\frac{\Gamma(\alpha-\epsilon+\gamma)}{\Gamma(\alpha+1+\gamma)}(t-a)^{\alpha+\gamma}$$

Proof.

$$\begin{aligned}{}^{\alpha}\mathfrak{I}_a^t(t-a)^{\gamma} &= {}^{\epsilon+1}\mathcal{I}_a^t\{(t-a)^{\theta-1}(t-a)^{\gamma}\}\\ &= {}^{\epsilon+1}\mathcal{I}_a^t(t-a)^{\gamma+\alpha-\epsilon-1}\\ &= \frac{1}{\Gamma(\epsilon+1)}\int_a^t (t-\xi)^{\epsilon}\\ &\quad\times(\xi-a)^{\gamma+\alpha-\epsilon-1}d\xi\\ &= \frac{\Gamma(\gamma+\alpha-\epsilon)}{\Gamma(\gamma+\alpha+1)}(t-a)^{\gamma+\alpha}\end{aligned}$$

□

Theorem 2. *Let $\theta>-1$ and $t>a$*

$${}^{\alpha}\mathfrak{I}_t^b(b-t)^{\gamma}=\frac{\Gamma(\alpha-\epsilon+\gamma)}{\Gamma(\alpha+1+\gamma)}(b-t)^{\alpha+\gamma}$$

Proof. The proof is similar to Theorem 1. □

For instance if we take $a=0$ and $b=0$ for the above results, we obtain ${}^{\alpha}\mathfrak{I}_0^t(t)^{\gamma}=\frac{\Gamma(\alpha-\epsilon+\gamma)}{\Gamma(\alpha+1+\gamma)}(t)^{\alpha+\gamma}$ and ${}^{\alpha}\mathfrak{I}_t^0(-t)^{\gamma}=\frac{\Gamma(\alpha-\epsilon+\gamma)}{\Gamma(\alpha+1+\gamma)}(-t)^{\alpha+\gamma}$ respectively. Now, similarly, we can get the conformable derivative of function $(t-a)^{\gamma}$. Namely, the derivative of $(t-a)^{\gamma}$ is

$${}^{\alpha}\mathfrak{D}(t-a)^{\gamma}=\gamma t^{1-\alpha}(t-a)^{\gamma-1}.$$

and again if we choose $a=0$, we get

$${}^{\alpha}\mathfrak{D}(t)^{\gamma}=\gamma t^{\gamma-\alpha}.$$

Now, by using the above results, one can find the conformable derivatives and integration of radial basis functions. Additionally throughout this and next sections ${}^nC_k$ denotes the combination of n and k such that ${}^nC_k=\frac{n!}{(n-k)!k!}$.

3.1. For $\varphi(t)=t^m$ (power basis function)

Theorem 3. *For $a\neq 0$, $t>a$ and $m\in\mathbb{N}$*

$${}^{\alpha}\mathfrak{I}_a^t t^m=(t-a)^{\alpha}\sum_{k=0}^{m}{}^mC_k a^{m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.$$

Proof. In order to prove the above theorem we use the Taylor expansion of t^m about the point $t=a$. Namely,

$$t^m=\sum_{k=0}^{m}{}^mC_k a^{m-k}(t-a)^k. \qquad (1)$$

If we substitute the equation (1) into conformable integration definition, we have

$$\begin{aligned}
{}^{\alpha}\mathfrak{J}_a^t t^m &= {}^{\alpha}\ \mathfrak{J}_a^t \sum_{k=0}^{m} {}^mC_k a^{m-k}(t-a)^k \\
&= \sum_{k=0}^{m} {}^mC_k a^{m-k}\,{}^{\alpha}\mathfrak{J}_a^t (t-a)^k \\
&= \sum_{k=0}^{m} {}^mC_k a^{m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^{\alpha+k} \\
&= (t-a)^{\alpha}\sum_{k=0}^{m} {}^mC_k a^{m-k} \\
&\times \frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.
\end{aligned}$$

□

Theorem 4. *For $b \neq 0$, $b > t$ and $m \in \mathbb{N}$*

$$\begin{aligned}
{}^{\alpha}\mathfrak{J}_t^b t^m &= (b-t)^{\alpha}\sum_{k=0}^{m}\sum_{k=0}^{m} {}^mC_k a^{m-k} \\
&\times \frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-b)^k.
\end{aligned}$$

Proof. The proof is similar to Theorem 3. □

Theorem 5. *For $a \neq 0$, $t > a$ and $m \in \mathbb{N}$*

$${}^{\alpha}\mathfrak{D}(t)^m = \frac{t^{1-\alpha}}{t-a}\sum_{k=0}^{m} {}^mC_k k a^{m-k}(t-a)^k. \quad (2)$$

Proof. In order to prove the above theorem we use the Taylor expansion of t^m about the point $t = a$ again. In other words if we substitute the equation (1) into conformable integration definition, we have

$$\begin{aligned}
{}^{\alpha}\mathfrak{D}t^m &= {}^{\alpha}\ \mathfrak{D}\sum_{k=0}^{m} {}^mC_k a^{m-k}(t-a)^k \\
&= \sum_{k=0}^{m} {}^mC_k a^{m-k}\,{}^{\alpha}\mathfrak{D}(t-a)^k \\
&= \sum_{k=0}^{m} {}^mC_k a^{m-k} t^{1-\alpha} k(t-a)^{k-1} \\
&= \frac{t^{1-\alpha}}{t-a}\sum_{k=0}^{m} {}^mC_k k a^{m-k}(t-a)^k.
\end{aligned}$$

□

3.2. For $\varphi(t) = e^{(-t^2/2)}$ (Gaussian basis function)

Now we can make use of the conformable derivatives and integration of power basis function, we are able to find out the Gaussian basis function derivatives and integrations.

Theorem 6. *For $a \neq 0$, $t > a$ and $m \in \mathbb{N}$*

$$\begin{aligned}
{}^{\alpha}\mathfrak{J}_a^t e^{-t^2/2} &= (t-a)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}\sum_{k=0}^{2m} {}^{2m}C_k a^{2m-k} \\
&\times \frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.
\end{aligned}$$

Proof. In order to prove the above theorem we use the Taylor expansion of $e^{-t^2/2}$ about the point $t = 0$. Namely,

$$e^{-t^2/2} = \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}(t)^{2m}. \quad (3)$$

If we substitute the equation (3) into conformable integration definition, we have

$$\begin{aligned}
{}^{\alpha}\mathfrak{J}_a^t e^{(-t^2/2)} &= {}^{\alpha}\ \mathfrak{J}_a^t \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}(t)^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}\,{}^{\alpha}\mathfrak{J}_a^t t^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}\left[(t-a)^{\alpha}\sum_{k=0}^{2m}\right. \\
&\left.\times\ {}^{2m}C_k a^{2m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k\right] \\
&= (t-a)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}\sum_{k=0}^{2m} \\
&\times\ {}^{2m}C_k a^{2m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.
\end{aligned}$$

□

Theorem 7. *For $a \neq 0$, $b > t$ and $m \in \mathbb{N}$*

$$\begin{aligned}
{}^{\alpha}\mathfrak{J}_t^b e^{-t^2/2} &= (b-t)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}\sum_{k=0}^{2m} {}^{2m}C_k a^{2m-k} \\
&\times \frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-b)^k.
\end{aligned}$$

Proof. The proof is similar to Theorem 7. □

Theorem 8. *For $a \neq 0$, $t > a$ and $m \in \mathbb{N}$*

$${}^{\alpha}\mathfrak{D}e^{-t^2/2} = t^{1-\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m 2m}{2^m m!}(t)^{2m-1}.$$

Proof. Similarly by using the Taylor expansion of Gaussian function about $t = 0$ we can calculate

the conformable derivative of it. That is,

$$
\begin{aligned}
{}^{\alpha}\mathfrak{D}e^{-t^2/2} &= {}^{\alpha}\mathfrak{D}\sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}(t)^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}{}^{\alpha}\mathfrak{D}t^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m}{2^m m!}t^{1-\alpha}2mt^{2m-1} \\
&= t^{1-\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m 2m}{2^m m!}(t)^{2m-1}.
\end{aligned}
$$

□

3.3. For $\varphi(t)=\sqrt{1+t^2}$ (Multiquadric basis function)

Similarly one can compute the conformable derivatives and integrations.

Theorem 9. *For $a\neq 0$, $t>a$ and $m\in\mathbb{N}$*

$$
\begin{aligned}
{}^{\alpha}\mathfrak{I}_a^t\sqrt{1+t^2} &= (t-a)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m} \\
&\times\sum_{k=0}^{2m}{}^{2m}C_k a^{2m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.
\end{aligned}
$$

Proof. In order to prove the above theorem we use the Taylor expansion of $\sqrt{1+t^2}$ about the point $t=0$. Namely,

$$
\sqrt{1+t^2}=\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}(t)^{2m}. \qquad (4)
$$

If we substitute the equation (4) into conformable integration definition, we have

$$
\begin{aligned}
{}^{\alpha}\mathfrak{I}_a^t\sqrt{1+t^2} &= {}^{\alpha}\mathfrak{I}_a^t\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}(t)^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}{}^{\alpha}\mathfrak{I}_a^t t^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m} \\
&\times\left[(t-a)^{\alpha}\sum_{k=0}^{2m}{}^{2m}C_k a^{2m-k}\right. \\
&\left.\times\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k\right] \\
&= (t-a)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m} \\
&\times\sum_{k=0}^{2m}{}^{2m}C_k a^{2m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(t-a)^k.
\end{aligned}
$$

□

Theorem 10. *For $a\neq 0$, $b>t$ and $m\in\mathbb{N}$*

$$
\begin{aligned}
{}^{\alpha}\mathfrak{I}_t^b\sqrt{1+t^2} &= (b-t)^{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m} \\
&\times\sum_{k=0}^{2m}{}^{2m}C_k a^{2m-k}\frac{\Gamma(\alpha-\epsilon+k)}{\Gamma(\alpha+1+k)}(b-t)^k.
\end{aligned}
$$

Proof. The proof is similar to Theorem 9. □

Theorem 11. *For $a\neq 0$, $t>a$ and $m\in\mathbb{N}$*

$$
{}^{\alpha}\mathfrak{D}\sqrt{1+t^2}=t^{1-\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m 2m}{(1-2m)4^m}t^{2m-1}.
$$

Proof. Similarly by using the Taylor expansion of multiquadric basis function about $t=0$ we can calculate the conformable derivative of it. That is,

$$
\begin{aligned}
{}^{\alpha}\mathfrak{D}\sqrt{1+t^2} &= {}^{\alpha}\mathfrak{D}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}(t)^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}t^{2m} \\
&= \sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m}{(1-2m)4^m}t^{1-\alpha}2mt^{2m-1} \\
&= t^{1-\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m\,{}^{2m}C_m 2m}{(1-2m)4^m}t^{2m-1}.
\end{aligned}
$$

□

4. Numerical example

In this section we will give some results of numerical solution of conformable differential equations to validate our numerical scheme. For that we will use RBF interpolation method by the help of collocation technique. Consider the general form of following conformable differential equation:

$$
{}^{\alpha}\mathfrak{D}y(t)+p(t)y(t)=q(t), \qquad y_0(t)=y(t_0). \quad (5)
$$

Let t_j be equally spaced grid points in the interval $0\leq t_j\leq K$ such that $1\leq j\leq L$, $t_1=0$ and $t_L=K$. Additionally, because collocation

approach has been used we not only require an expression for the value of the function

$$y(t) = \sum_{k=1}^{L} a_k \psi(\|x - x_k\|) \tag{6}$$

but also for the conformal derivative given in (5). Thus, by conformal differentiating (6), we get

$${}^{\alpha}\mathfrak{D}y(t) = \sum_{k=1}^{L} a_k^{\alpha} \mathfrak{D}\psi(\|t - t_k\|)$$

where ${}^{\alpha}\mathfrak{D}$ denotes the conformable derivative the with respect to t. In order to compute conformable derivative of radial basis functions we take the advantage of formulas which are derived in the previous section. Then using the RBF collocation method, one can compute the unknown coefficients a_k's by solving following matrix system:

$$\sum_{k=1}^{L} a_k^{\alpha} \mathfrak{D}\psi(\|x_j - x_k\|) + p(t) \sum_{k=1}^{L} a_k \psi(\|x_j - x_k\|) = q(t), \qquad j = 2, \ldots, L.$$

with boundary condition. In order to illustrate this scheme by numerically we take the following conformable differential equations:

(1)

$${}^{\alpha}\mathfrak{D}y(t) + y(t) = 0$$
$$y_0(t) = 1, \quad y_{exact}(t) = e^{-\frac{1}{\alpha}t^{\alpha}}$$

(2)

$${}^{\alpha}\mathfrak{D}y(t) + \alpha y(t) = 1 + t^{\alpha}$$
$$y_0(t) = 0, \quad y_{exact}(t) = \frac{t^{\alpha}}{\alpha}$$

(3)

$${}^{\alpha}\mathfrak{D}y(t) + y(t) = \sqrt{1 + \sin\left(\frac{2t^{\alpha}}{\alpha}\right)}$$
$$y_0(t) = 0, \quad y_{exact}(t) = \sin\left(\frac{t^{\alpha}}{\alpha}\right)$$

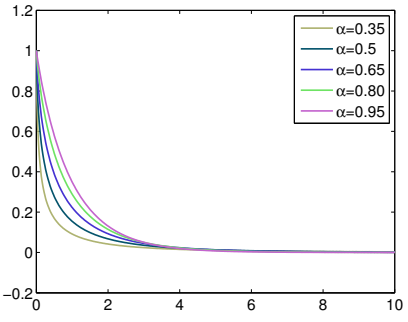


Figure 1. $y(t)$ versus t using multiquadric basis function with $\varepsilon = 10^{-4}$ for $p(t) = 1$ and $q(t) = 0$ for different value of α.

Here we use the multiquadric basis function with $\varepsilon = 10^{-4}$. In Figures 1, 2 and 3, we present the numerical solutions of given conformable differential equations with different α values. These results are in accord with the exact solutions of them.

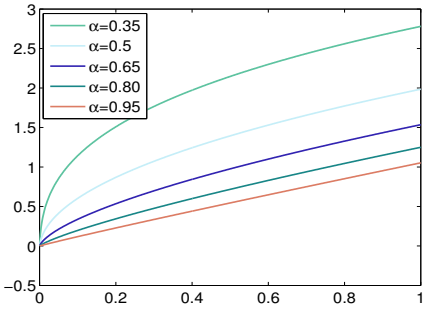


Figure 2. $y(t)$ versus t using multiquadric basis function with $\varepsilon = 10^{-4}$ for $p(t) = \alpha$ and $q(t) = 1 + t^{\alpha}$ for different value of α.

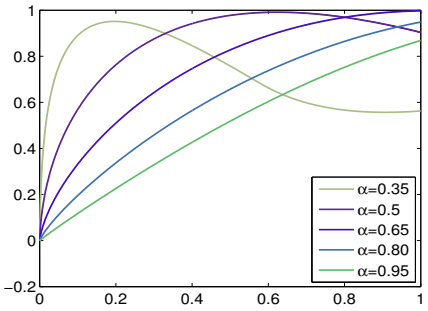


Figure 3. $y(t)$ versus t using Multiquadric basis function with $\varepsilon = 10^{-4}$ for $p(t) = 1$ and $q(t) = \sqrt{1 + \sin\left(\frac{2t^{\alpha}}{\alpha}\right)}$ for different value of α.

5. Conclusion

In this paper we gave the derivatives and integrals of three kinds of radial basis functions such as powers, Gaussians and multiquadric by using the conformable derivatives and integrals which are new type of fractional calculus. These findings allow to solve conformable differential equations by the RBF's. Then we gave three differential equations to show that this technique is applicable. These differential equations are solved by the help of RBF collocation method.

References

[1] Hammad, M.A. and Khalil, R. (2014). Conformable fractional heat differential equations, *International*

Journal of Pure and Applied Mathematics, 94(2), 215-221.

[2] Hammad, M.A. and Khalil, R. (2014). Abel's formula and wronskian for conformable fractional differential equations, *International Journal of Differential Equations and Applications*, 13(3), 177-183.

[3] Kilbas, A.A., Srivastava, H.M. and Trujillo, J.J. (2006). *Theory and applications of fractional differential equations*, North-Holland Mathematical Studies, Vol.204, Elsevier (North-Holland) Science Publishers, Amsterdam, London and New York.

[4] Samko, S.G., Kilbas, A.A. and Marichev, O.I. (1993). Fractional Integrals and Derivatives: Theory and Applications, *Gordonand Breach, Yverdon et alibi*.

[5] Abdeljawad, T. (2015). On conformable fractional calculus, *Journal of Computational and Applied Mathematics*, 279, 57-66.

[6] Khalil, R., Al horani, M, Yousef, A. and Sababheh, M. (2014). A new definition of fractional derivative, *Journal of Computational Apllied Mathematics*, 264, 65-70.

[7] Anderson, D.R. and Ulness, D.J. (2016). Results for conformable differential equations, *preprint*.

[8] Atangana, A., Baleanu, D. and Alsaedi, A. (2015) New properties of conformable derivative, *Open Mathematics*, 13, 889-898.

[9] Iyiola, O.S. and Nwaeze, E.R. (2016). Some new results on the new conformable fractional calculus with application using D'Alambert approach, *Progress in Fractional Differentiation and Applications*, 2(2), 115-122.

[10] Zheng, A., Feng, Y. and Wang, W. (2015). The Hyers-Ulam stability of the conformable fractional differential equation. *Mathematica Aeterna*, 5(3), 485-492.

[11] Usta, F. and Sarkaya, M.Z. (2017). Some improvements of conformable fractional integral. *Inequalities International Journal of Analysis and Applications*, 14(2), 162-166.

[12] Sarkaya, M.Z., Yaldz, H. and Budak, H. (2017). On weighted Iyengar-type inequalities for conformable fractional integrals, *Mathematical Sciences*, 11(4), 327-331.

[13] Akkurt, A., Yldrm, M.E. and Yldrm, H. (2017). On Some Integral Inequalities for Conformable Fractional Integrals, *Asian Journal of Mathematics and Computer Research*, 15(3), 205-212.

[14] Usta, F. (2016). A mesh-free technique of numerical solution of newly defined conformable differential equations. *Konuralp Journal of Mathematics*, 4(2), 149-157.

[15] Usta, F. (2016). Fractional type Poisson equations by radial basis functions Kansa approach. *Journal of Inequalities and Special Functions*, 7(4), 143-149.

[16] Usta, F. (2017). Numerical solution of fractional elliptic PDE's by the collocation method. *Applications and Applied Mathematics: An International Journal (AAM)*, 12 (1), 470-478.

[17] Katugampola, U.N. (2014). A new fractional derivative with classical properties, *ArXiv:1410.6535v2*.

[18] Katugampola, U.N. (2015). Mellin transforms of generalized fractional integrals and derivatives, *Applied Mathematics and Computation*, 257, 566580.

[19] Kansa, E.J. (1990), Multiquadrics a scattered data approximation scheme with applications to computational filuid-dynamics. I. Surface approximations and partial derivative estimates, *Computers and Mathematics with Applications*, 19(8-9), 127-145.

[20] Kansa, E.J. (1990), Multiquadrics a scattered data approximation scheme with applications to computational filuid-dynamics. II. Solutions to parabolic, hyperbolic and elliptic partial differential equations, *Computers and Mathematics with Applications*, 19(8-9), 147-161.

[21] Zerroukat, M., Power, H. and Chen, C.S. (1998). A numerical method for heat transfer problems using collocation and radial basis functions. *International Journal for Numerical Methods in Engineering*, 42, 12631278.

[22] Chen, W., Ye, L. and Sun, H. (2010). Fractional diffusion equations by the Kansa method, *Computers and Mathematics with Applications*, 59, 16141620.

[23] Podlubny, I. (1999), *Fractional differential equations. An introduction to fractional derivatives, fractional differential equations, to methods of their Solution and some of their applications*, Academic Press, San Diego, CA.

Fuat Usta *received his BSc (Mathematical Engineering) degree from Istanbul Technical University, Turkey in 2009 and MSc (Mathematical Finance) from University of Birmingham, UK in 2011 and PhD (Applied Mathematics) from University of Leicester, UK in 2015. At present, he is working as a Assistant Professor in the Department of Mathematics at Düzce University (Turkey). He is interested in Approximation Theory, Multivariate approximation using Quasi Interpolation, Radial Basis Functions and Hierarchical/Wavelet Bases, High-Dimensional Approximation using Sparse Grids. Financial Mathematics, Integral Equations, Fractional Calculus, Partial Differantial Equations.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.183-194 (2018)
http://doi.org/10.11121/ijocta.01.2018.00399

RESEARCH ARTICLE

Multiobjective PID controller design for active suspension system: scalarization approach

O. Tolga Altınöz

Department of Electrical and Electronics Engineering, Ankara University, Turkey
taltinoz@ankara.edu.tr

ARTICLE INFO

Article History:
Received 06 October 2016
Accepted 14 October 2017
Available 22 April 2018

Keywords:
PID controller
Multiobjective
Scalarization
Particle swarm optimization
Differential evolution
Partitioned method

AMS Classification 2010:
90C26, 80M50, 78M50
49N05, 93C05

ABSTRACT

In this study, the PID tuning method (controller design scheme) is proposed for a linear quarter model of active suspension system installed on the vehicles. The PID tuning scheme is considered as a multiobjective problem which is solved by converting this multiobjective problem into single objective problem with the aid of scalarization approaches. In the study, three different scalarization approaches are used and compared to each other. These approaches are called linear scalarization (weighted sum), epsilon-constraint and Bensons methods. The objectives of multiobjective optimization are selected from the time-domain properties of the transient response of the system which are overshoot, rise time, peak time and error (in total there are four objectives). The aim of each objective is to minimize the corresponding property of the time response of the system. First, these four objective is applied to the scalarization functions and then single objective problem is obtained. Finally, these single objective problems are solved with the aid of heuristic optimization algorithms. For this purpose, four optimization algorithms are selected, which are called Particle Swarm Optimization, Differential Evolution, Firefly, and Cultural Algorithms. In total, twelve implementations are evaluated with the same number of iterations. In this study, the aim is to compare the scalarization approaches and optimization algorithm on active suspension control problem. The performance of the corresponding cases (implementations) are numerically and graphically demonstrated on transient responses of the system.

1. Introduction

Car suspension system has been installed to the vehicle for omitting the undesired low-frequency high magnitude disturbance/vibration due to the imperfect conditions of the road. The main aims of the suspension system are to the comfort of the passengers and absorb the vibrations from ground to car body. Suspension systems can be categorized as depended and independent systems with respect to the connection between suspension and car body. Depended systems are relatively heavy structures which are generally considered as the rear suspension system. Left and right suspensions are (not directly) connected to each other. Hence change at the dynamics of one of the suspension has a direct effect on the other suspension, which causes discomfort to the passengers. However, since the number of moving parts are less than independent systems, the lifetime of these systems are longer than independent systems under the same conditions. Independent systems are discrete systems. Modern cars prefer independent suspensions, especially for the front wheels.

Since there are many different road conditions, it isn't possible to mention about single vibration frequency and magnitude. Therefore, a particular structure to absorb occurred vibration is needed. There are three main structures can be proposed;

which are passive, semi-active and active suspension systems. Passive systems are pre-design systems where the conditions (relatively protected environment) of the road are known. Therefore, well-designed damper and spring can be answer the suspension problem. They do not have to interconnected devices such as actuator, pneumatics, and hydraulics related to suspension system. Semi-active has at least a damper with a narrow range of variable damping force and damping coefficients. They can be changed before the journey, they don't have any control actions (may be a very simple one). Active systems are improved version of the passive systems with controllable actuators, which can be activated with current and/or voltage. In this paper, active independent suspension system is considered and quarter-car model is evaluated for this purpose.

The control action of the active suspension can be produced vary from classical control approaches, robust control, optimal control, to intelligent methods. Among them, PID controller is the simplest and relatively cheaper controller which is applied to this problem. However, even today the tuning of PID controller is considered as an important aspect. Astrom and Hagglund [1] discussed the future of PID controller in their paper. It is imposed from the study that since there are many aspects which are needed to be considered for tuning, there is a need for design frameworks (section 4). In this study, the PID coefficients for the active suspension control problem is tunned by converting into the multiobjective optimization problem.

In literature, there are some attempts for multiobjective PID tuning problem. In the study of [2], the authors proposed a niche-based multiobjective optimization algorithm and Genetic Algorithm (GA) for tuning of PID controller which is applied to nonlinear MIMO process. As the objective parameters, settling time, overshoot and rise time are weighted and summed. In other words, the authors of that study prefer weighted sum scalarization method. The results showed that proposed method presents better performance when compared with Ziegler-Nichols. Similar results are obtained by Neath et al. [3]. They applied GA-based PID tuning methodology with the aid of weighted sum scalarization method (rise time, settling time and overshoot). In the paper, both simulation and implementation results are demonstrated, and both showed the performance of the tuning method. In [4], Lin et al. applied similar GA-based multiobjective PID design ([5]) method with the aid of weighted sum scalarization. As a difference three time-domain specifications are considered as objectives which are the rise time, overshoot and steady-state error. A more general contribution of the multiobjective PID tuning is presented by Reynoso-Meza et. al. [6]. Both PID and state space feedback controllers are tunned by using multiobjective optimization algorithm. However, instead of time-domain properties the integral of the absolute value of the derivative the control signal and integral of the absolute value of the error is considered as objectives. This idea is applied to twin rotor MIMO system (as a PI controller tuning algorithm similar study evaluated on [7] and for aircraft in [8]) and results also support the better performance of the multiobjective tuning methodology. Also in [9], two similar objectives which are integral time absolute error and control effort is taken to tune PID controller for the plastic injection molding process, and in [10] same objectives are evaluated on weighted sum scalarization function. Artificial Bee Colony algorithm and weighted sum approach was also applied to load frequency control problem [11], and almost same improvement was reached. Hung et. al. [12] presents the same methodology, but on the contrary the authors preferred multiobjective simulated annealing algorithm and improved strength Pareto algorithms. The other difference is the selection of objective function. In that study, three objectives are considered which are robust stability, disturbance attenuation and integral of square error. Simulation results demonstrate the high performance of the proposed framework. In the study of Tseng et al. [13], the sufficient performance of the multiobjective PID control design in plant uncertainties and under external disturbance, also parametric uncertainties are considered in [14]. In [15], Tang et al. gives the multiobjective optimization scheme (multiobjective generic algorithm (MOGA)) for Fuzzy PID controller. Liu and Daley [16] proposed a three-layered study, in a way that first time-domain optimal tuning of PID control is presented. Following that consequently, frequency-domain optimal-tuning PID and multiobjective tuning PID controllers are discussed. These PID tuning algorithms are applied to three industrial systems. Results showed that optimal PID significantly improve performance. In Ayala and Coelho paper [17], a multiobjective optimization algorithm (NSGA-II) is applied without using scalarization functions (the similar study is presented in [18]). The objectives are selected as position error and torque.

In active suspension problem, the aim is to improve the transient response of the system under

change in the road profile. At the steady state of the system, the damper (or spring) in the system converges/remains in a stable state and this state is not changed if the road remains the same. However, if the road is altered due to the imperfect conditions of the road the suspension system response to the change at the road. If this change is relatively small and the mass differences between body and suspension system including the spring property of the tire is large that this change cannot be perceived. However, if this change is large enough, then the car body moves upwards, and oscillation begins. If this oscillation could not be damped or car body moves too far from the suspension system than safety and comfort of the passengers is reduced dramatically. Therefore a control structure is needed to apply a force on the damper to damp this effect. Hence in this paper, PID controller is applied to the quarter car suspension model. The PID coefficients are the key parameters which are the direct effect on the performance of the system. However, the conventional PID tuning methods could not be applied due to the not satisfactory performance. Hence, in this study, a methodology for PID tuning for active suspension system is proposed. Initially, the tuning scheme of active suspension control is converted into the multiobjective problem. However, since it is desired to get a single solution, instead of multiobjective optimization algorithms, scalarization approaches are evaluated for this purpose. Three scalarization methods are applied in this paper which are Weighted sum method, Epsilon method, and Benson's method. As a result of scalarization, the multiobjective problem is reduced to single-objective one. Then this problem is solved by using heuristic optimization algorithms. Four optimization algorithms are applied to solve the problem, which are Particle Swarm Optimization, Differential Evolution, Firefly Algorithm and Cultural Algorithm. In total, three scalarization approaches with four heuristic optimization algorithm are compared to each other.

This paper is organized into five main sections including the introduction. After the presentation of the aim and literature search at the introduction section, problem definition is given in the second section. The mathematical model of the car suspension system and corresponding controller algorithm with objectives are also presented in this section. The third section is written for briefly explaining the toolsets which are used in this paper. Two sub-sections are given in this section which are scalarization approaches and optimization algorithms. The fourth section is given for implementation and obtained results. In this section, all the information explained in the previous sections are evaluated on the simulation environment. The last section is the conclusion of this study.

2. Problem definition

The graphical description of the quarter-car passive suspension model [19] is illustrated in Figure 1 , and the following equations give the mathematical description of this model [20].

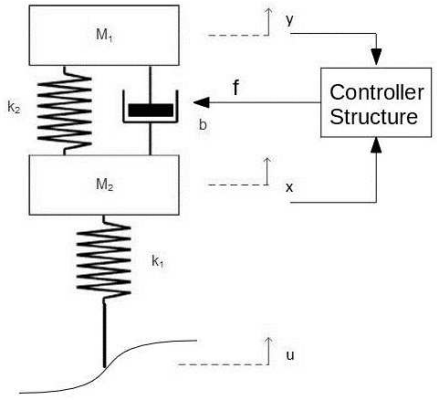


Figure 1. Quarter car suspension system.

$$M_1\ddot{y} = -b(\dot{y} - \dot{x}) - k_2(y - x) - f + M_2 g \quad (1)$$

$$\begin{aligned} M_2\ddot{x} = &-b(\dot{x} - \dot{y}) - k_2(x - y) \\ &-k_1(x - u) + f + M_1 g \end{aligned} \quad (2)$$

where g is the gravitational constant, f is the force (control input) at the damper under the car body, k is the spring constants and b is the damper constant. The masses m_1 and m_2 are corresponded to bar body and tire masses. The parameters are selected as: $m_1 = 60kg$, $m_2 = 300kg$, $k_1 = 160000kg/s$, $k_2 = 16000kg/s$, $g = 9.8m/s^2$, and $b = 1400kgm/s^2$.

2.1. PID controller

In this study, PID controllers are designed and parameters are optimized [21]. The Laplace transform of the PID controller is given below.

$$G(s) = K_P + \frac{K_I}{s} + K_D s \quad (3)$$

For proper usage of the PID controller three parameters (proportional parameter (K_P), integral term (K_I), derivative term (K_D)) are needed to

be optimized for the desired performance. As a general perspective, K_P decreases the rise time and steady-state error, but for large steady-state error properly selected K_I is needed to eliminate this large steady-state error.However, this term increases the overshoot. Generally, the increase in overshoot and settling time is relatively small (depended on the structure of the system). If it is large enough, derivative term K_D is added for decreasing the overshoot and settling time. In this study, three values corresponding to these three parameters are determined with the aid of optimization algorithms.

The framework of the controller system is given in Figure 2. The PID controller is applied to the active suspension system, and the road disturbance is changes the system dynamic. By collecting the transient response parameters are converted into objective functions. Nest, this objective functions is formed to the multiobjective problem and converted to single objective one with the aid of scalarization approach. Finally, the optimization algorithm is applied to this offline PID parameter tuning framework.

3. Methods and techniques

The study begins with the conversion of the multiobjective problem which is defined in the previous section (Problem Definition) into single objective problem. For this purpose three different scalarization methods are applied to the multiobjective problem. The scalarization methods are called i) weighted sum, ii) ϵ-constrained, and iii) Benson's methods. After conversion to the single-objective problem, two different optimization algorithms are applied to solve this problem. The algorithms are called Particle Swarm Optimization and Differential Evolution algorithms. At the next section, the results of these implementations are compared with each other.

3.1. Scalarization approaches

3.1.1. Weighted sum method

Weighted (linear) sum method is the oldest and best known approach for solving multiobjective optimization problems [22]. The scalarization formula of weighted sum method is given at the following equation.

$$J = \sum_{m=1}^{M} (w_m f_m(x)) \tag{4}$$

where w are the positive weights of each objective. It is assumed that the sum of all weight are equal to one because of the necessity of normalize values at the objective space.

$$\sum_{m=1}^{M} (w_m) = 1 \tag{5}$$

The distribution of the solutions obtained from optimization algorithm is generally non-uniform. Also, the weighted sum method can perform better in convex regions of the search space. Since the sum of weights equals to one, w_m parameters are selected as equal to each other, which has the value of $w_m = \frac{1}{m}$. In our study there are four objectives (in the implementation section these objectives are explained). Therefore all of the weights are equal to 0.25.

3.1.2. ϵ-constraint Method

This method is proposed to get rid of the convexity problem of the weighted sum method; ϵ-constrained method was introduced by Haimes et al. (1971) [23]. The idea is based on minimizing the single objective while considering the other objective as constraint in the form of inequality. The scalarization function is presented below.

$$\begin{aligned} J &= f_n(x) \\ f_m(x) &\leq \epsilon, \\ m = 1, ..., M; m &\neq n \end{aligned} \tag{6}$$

where ϵ is the variable such that with a properly selected ϵ, feasible solution can be obtained. This method can be applied for general problems, no convexity assumption is desired. This scalarization approach converts the unconstrained multiobjective problem into a constrained single objective problem. Therefore, a mechanism is needed to handle these constrained. For our study, there are three constrains are defined for this approach (since we have four objectives). For this purpose penalty function idea is applied to ϵ-constraint method. The penalty function definition is given below.

Definition 1. *A function $p(x)$ is said to be penalty function for the vector x if penalty function satisfies two conditions where $g(x) \leq 0$ is the constrained i) $p(x) = 0$ if $g(x) \leq 0$ and ii) $p(x) > 0$ if $g(x) > 0$*

For this study, $\sum_{i=i}^{m} max(\epsilon_i, f_i(x))$ function is selected as penalty function. As a result the overall

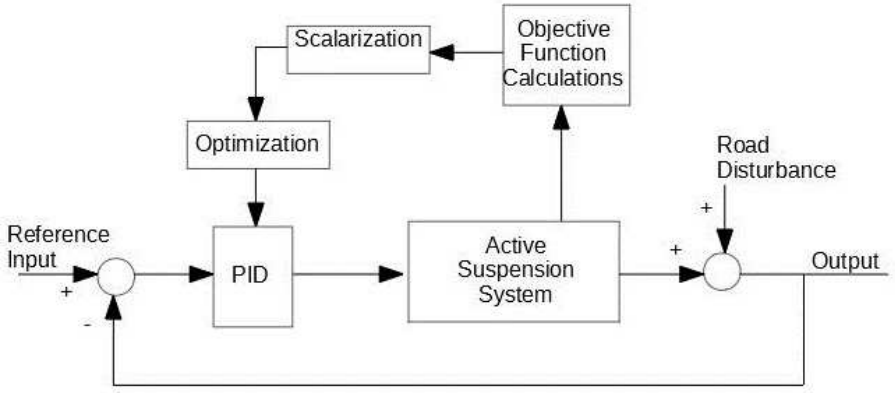


Figure 2. Control structure for quarter car suspension system.

scalarization equation is changed to the following form.

$$J = f_n(x) + \sum_{i=i}^{m} max(\epsilon_i, f_i(x)) \tag{7}$$

3.1.3. Benson's method

Benson's method, as one of the scalarization methods, was introduced by Benson in 1978 [24] which is an extended and improved version of their study on vector maximization [25]. In [26], it is shown that the scalarization method is valid in general case without convexity assumption if the reference point is selected properly on the feasible solution set. The idea is based on the determination of the properly or improperly efficient solution, which are defined below.

Definition 2. *x_0 is said to be efficient solution if $f_i(x) > f_i(x_0)$ for some x, and there exists at least one j such that $f_j(x) < f_j(x_0)$.*

Definition 3. *x_0 is said to be properly efficient solution if it is efficient and if there exist a scalar named as $M > 0$, such that for each i, we have*

$$\frac{f_i(x) - f_i(x_0)}{f_j(x_0) - f_j(x)} \leq M \tag{8}$$

for some j such that $f_j(x) < f_j(x_0)$ whenever $f_i(x) > f_i(x_0)$.

Definition 4. *x_0 is said to be improperly efficient solution if it is efficient and if there exist a scalar named as $M > 0$, such that there is a point x and an i such that $f_i(x) > f_i(x_0)$ and*

$$\frac{f_i(x) - f_i(x_0)}{f_j(x_0) - f_j(x)} > M \tag{9}$$

for all j such that $f_j(x) < f_j(x_0)$.

Since the originally proposed method has drawbacks against differentiation, Ehrgott [27] was proposed a modified formulation to make easier of the differentiation process at classical optimization/search algorithms. The scalarization method is given at the following formula.

$$J = \sum_{m=1}^{M} max(0, (z_m - f_m(x))) \qquad f_m(x) \leq z_m, \qquad m = 1, ..., M; m \neq n \tag{10}$$

where z is the chosen solution in the feasible region. First the nonnegative difference between each objective value is calculated and summarized. The maximization is similar to find a cube with the largest perimeter [28]. A set of constraints are added to the formulation. similarly to ϵ-constrained approach, this constraints are handled by using the penalty function.

3.2. Optimization algorithms

In the previous section, scalarization approaches, which are converted multiobjective PID tuning problem into the single objective problem; are explained. In this section, two single objective optimization algorithms are explained, which are Particle Swarm Optimization (PSO) and Differential Evolution (DE).

3.2.1. Particle swarm optimization

Particle Swarm Optimization (PSO) is an optimization algorithm which is proposed by Kennedy and Eberhart [29] in 1995 inspired from the behaviors of the animal swarms. Figure 3 graphically illustrates the idea of the PSO algorithm.

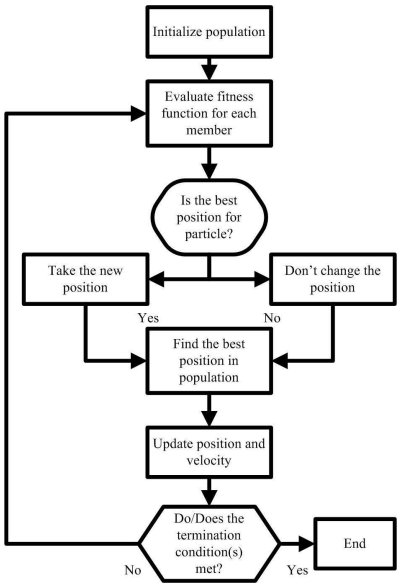


Figure 3. Flow-diagram of particle swarm optimization.

Each member of the population has two properties; position $(x_1, x_2, ..., x_D)$ and velocity $(v_1, v_2, ..., v_D)$, where D is the dimension of the search space. At the beginning of the algorithm, positions are randomly assigned inside the borders of the search space, and similarly velocities are taken value inside $[0, 1]$. At the first phase of each iteration, objective values of each member in population is calculated. There are two kinds of memory defined for each member. The first one is the best position among the population (*gbest*). Same position is recorded for all members at each iteration. The second one is belongs to the each member. It stores the best location which has been ever visited by the corresponding member (*pbest*). Therefore, at the second phase of the algorithm these two memories are updated by using the objective values. Then as the last phase of the algorithm the position and velocities are updated by using the equations given below.

$$v_i[k+1] = v_i[k] + c_1 rand()(pbest_i - x_i[k]) + c_2 rand()(gbest_i - x_i[k]) \quad (11)$$

$$x_i[k+1] = x_i[k] + v_i[k+1] \quad (12)$$

where $c_1 = c_2 = 2.05$ are the algorithm parameters and *rand* is the random number generator. These steps are repeated until the termination condition is met.

3.2.2. Differential evolution

Differential Evolution (DE) was proposed by Storn and Price in 1995 [30]. Since DE uses operators mutation, crossover and selection; it is considered as a part of evolutionary algorithms. Figure 4 gives the flow diagram of the DE algorithm.

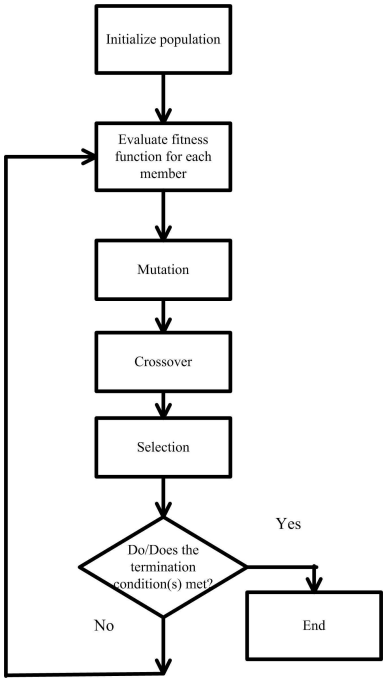


Figure 4. Flow-diagram of differential evolution.

The algorithm begins with the randomly initialization of the population $(x_1, x_2, ..., x_D)$ on search space, where D is the dimension of the problem. After the initialization of the population is completed, then as the first operator, Mutation, is evaluated. The idea of the mutation operator is to form a new population $(v_1, v_2, ..., v_D)$. One of the possible mutation operation (also used in this study) is given below.

$$v_i[k+1] = x_{r_1}[k] + F(x_{r_2}[k] - x_{r_3}[k]) \quad (13)$$

where r_1, r_2, and r_3 are the randomly selected index from the population, and F is the algorithm coefficient which is selected as $F = 0.8$.

After the mutation operator is completed, then Crossover operation is evaluated. By using this operator a new set of solution is obtained $(u_1, u_2, ..., u_D)$, by using the formulation (called binomial operator) given below.

$$u_i[k] = v_i[k] \quad \text{if} \quad rand < CR$$
$$u_i[k] = x_i[k] \quad \text{otherwise} \quad (14)$$

where CR is the second algorithm parameter, and it has the value of $CR = 0.9$. The final operator is called the Selection. This operator is just compared the two vector set X and U, and best of these two sets are selected with respect to the objective value. This new set survives to the next iteration.

3.2.3. Firefly algorithm

As a population based optimization algorithm, Firefly Algorithm (FFA) was proposed by Yang [31], [32] in 2008. The algorithm is designed based on the light intensity and attractiveness between fireflies in their nest. Light intensity of fireflies gives the warning ability against predators and attractiveness for mating. The light intensity is changed with respect to the distance between light source and fireflies. The algorithm is designed based on light intensity property, where Figure 5 gives the flow diagram of the FFA algorithm

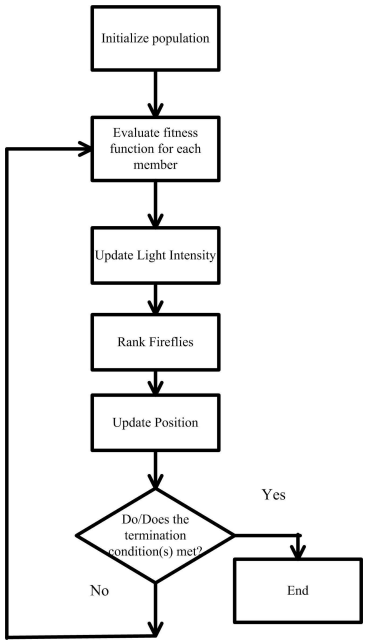


Figure 5. Flow-diagram of firefly algorithm.

In the FFA perspective, the search space is corresponding to the light distribution. Therefore, objective value becomes the light distribution. The performance of the FFA is evaluated on benchmark problems in [33]. The results showed that FFA presents acceptable performance. Then the algorithm is improved for multimodal problems [34] and applied to real-world problems [32], [35]. Fireflies in the algorithm assumes to have position $(x_1, x_2, ..., x_D)$ on search space. Then, light intensity is calculated by using the position of the member and the distance to other firefights as formulated in below.

$$I(r) = I_0 e^{\gamma r_{ij}^2} \tag{15}$$

where γ is the constant light absorption coefficient, I_0 is the initial light intensity at the distance r, which is defined below.

$$r_{ij} = \|x_i - x_j\| \tag{16}$$

As the last operator of the algorithm the position of each firefly is changed by using the light intensity (attractiveness) and a random movement. Even the algorithm is well organized and proposed acceptable performance, the number of algorithmic control parameters is relatively many in number. Therefore, in this paper the parameters reported in [31] is preferred.

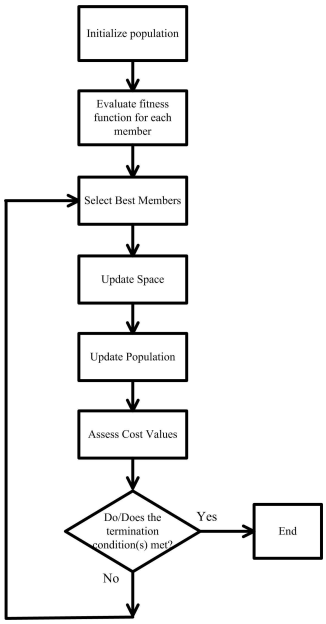


Figure 6. Flow-diagram of cultural algorithm.

3.2.4. Cultural algorithm

Cultural Algorithm (CA) is an optimization algorithm based on social learning and evolution. It was introduced by Reynold [36] and matured in [37]. The algorithm is successfully applied to industrial problems [38]. In CA, the population shares the information pool (in other words belief space). This space contains normative [39], spatial, temporal [40], domain and exemplar knowledge [37]. In interaction of these knowledge is the source of the algorithm [41]. Figure 6 gives the flow diagram of the CA algorithm

In CA, two spaces named as Belief Space and Population space are interacted with each other. The

members in population space are used to calculate the fitness (cost) functions. Based on the cost values the individuals are selected to impact (update) the Belief Space. Then beliefs in the Belief Space have influence the evolution (generation) of the populations in population space. New members are generator with the aid of Belief space and the best members among the joint space remains for the next generator.

4. Implementations and results

The purpose of this study is to design PID controller for active suspension control under the road change. In addition to this purpose, scalarization and heuristic optimization algorithms are compared with each other. For these purposes, first it is assumed that the road change is the step, in other words the disturbance of the system is the step input. All of the implementations are compared with each other with respect to the time response transient properties numerically. The results are discussed and one of from each scalarization methods and optimization algorithms are selected. Then, as the second phase of the study, different road change as ramp input is applied the same system and the performance of the proposed method is discussed with a ramp input, which is applied different ramps.

Initially, the car suspension is controlled under a rapid change at the contact between road and the wheel. It is assumed that there is $1m$ change at this contact, in other words, a step input is applied to the system. It is desired that the suspension is rapidly and smoothly absorb the vibration and return its equilibrium point. First, the problem is defined as a multiobjective problem. The objectives are selected from the time-response of the system. For using at objectives; overshoot, rise time, peak time and error are taken as variables.

Time domain response of the system is divided into two part. The initial part is called the transient response and the rest of it is the steady-state response. For a steady state response the system output (car body (M_1) position change) must be settled within %2 of the desired output which is selected as 1 for this study. Since the mechanical properties of the suspension system is based on damper and springs, it is expected to settle on a certain level. Therefore, even the steady state error is important as an objective, it is not necessary to add as a comparison factor.

The time where the output reaches the desired output (steady state level) for the first time is called rise time (tr), the time where the response reaches the peak is called peak time (tp). The percentage of the maximum value with respect to the desired response is called overshoot (OS). In addition to these time properties, also error (e) (difference between desired level and the output) is evaluated as the objective, which is important as an objective (as explained previously). The mean integral of absolute difference between desired signal and the output (mae) is calculated and considered as one of the objective.

Corresponding four objectives ($f_i, \quad i = 1, 2, 3, 4$) are given below;

$$f_1 = min(tr) \quad f_2 = min(tp)$$
$$f_3 = min(OS/100) \quad f_4 = min(mae(e)) \tag{17}$$

As the next step, this four objective problem is converted into single objective one. Three scalarization functions are preferred, which are weighted-sum, ϵ-constrained, and Benson's methods. In weighted sum method, all of these objectives are summed to each other since all of the weights are same with each other. The single objective of the weighted sum method is given below.

$$f = \frac{1}{4}[tr + tp + (OS/100) + mae(e)] \tag{18}$$

Like weighted sum method, ϵ-constrained method also evaluated for the problem. However, since this scalarization approach is considered, the constrained part of the approach should be evaluated. For this purpose, penalty function is preferred and constrained problem converted into unconstrained problem. The final ϵ-constrained approach formulation is presented below.

$$f = mae(e) + [max(\epsilon_1, tr)$$
$$+max(\epsilon_2, tp) + max(\epsilon_3, OS/100)] \tag{19}$$

where $\epsilon_1 = 0.05$, $\epsilon_2 = 0.05$, and $\epsilon_3 = 0.2$ are selected. In other words, (as an example) if the overshoot is decreased under 20%, it is considered as the desired performance is reached and the contribution of this property is becomes zero. However, if it is larger than 20%, only the overshoot value is added to the objective value. If all of the given values are above the given ϵ values, than ϵ-constrained method is almost the same as weighted sum method. Similarly, the overall function for Benson's method is given below.

Table 1. Time response parameter comparison with respect to the step input.

Time Response	PSO-WS	PSO-Eps	PSO-Benson	DE-WS	DE-Eps	DE-Benson
Rise Time	0.1637	0.1624	0.1624	0.1366	0.1368	0.1368
Overshoot (%)	48.03	49.99	49.99	50.04	50.04	50.49
Settling Time	0.4389	0.4585	0.4392	0.3542	0.3542	0.3543
	FFA-WS	FFA-Eps	FFA-Benson	CA-WS	CA-Eps	CA-Benson
Rise Time	0.1376	0.1367	0.1365	0.0458	0.0796	0.0796
Overshoot (%)	49.29	49.58	50	10.81	0.0431	0.0431
Settling Time	0.3539	0.3541	0.3541	0.1	0.1224	0.1224

$$f = max(0,(z_1 - mae(e))) + max(0,(z_2 - tr)) + max(0,(z_3 - tp)) + max(0,(z_4 - OS/100)) + [max(z_1, mae(e)) + [max(z_2, tr) + max(z_3, tp) + max(z_4, OS/100)] \quad (20)$$

where $z_1 = 10$, $z_2 = 0.5$, $z_3 = 0.5$, and $z_4 = 0.2$ are selected. Even this method looks different from other two scalarization functions, it is almost the united version of weighted sum and ϵ-constrained methods such that last objective (mae) also considered inside the maximum function and difference between a reference point and objective is calculated. This scalarization equation can be considered as two parts. In the first part if the obtained property is larger than the z then first part ($max(0,(z-P))$) become zero. However the second part (the constrained part) is equals to the properties value. On contrary, if z is larger that obtained property, than first part equals to the difference between $(z-P)$ and the second part equals to z. When compared to the ϵ-constrained method, a residue from ($max(0,(z-P))$) is added, which increases when the undesired response obtained.

Since the change at the relative position between ground and car body forms an undesired vibration on the car body. This vibration continues at a certain time if any control action didn't apply. Figure 7 gives the obtained uncontrolled signal. This figure also demonstrates the necessity of the control action. In general, the fastest and low (preferably zero) overshoot is desired. In real world application it corresponds to that after the tire falls into the hole on the road, it is expected to return the car body to its initial height as fast as possible. In addition, it is not desired to move the car body to a higher height of the initial position.

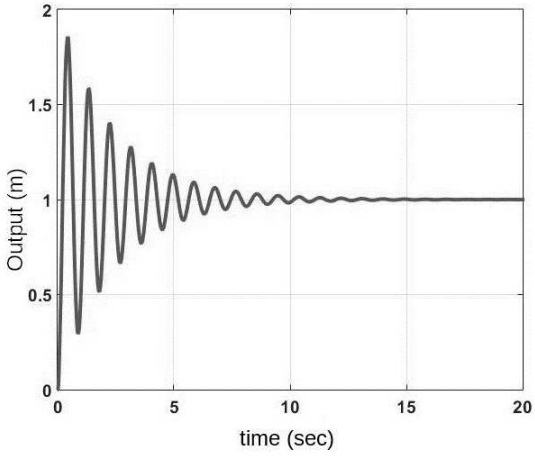


Figure 7. Change the relative position of the car body without control action.

From this figure, the car body travels almost $\times 2$ (100 %) more than desired level, and it needs almost $15sec$ to absorb and reach to the equilibrium point. Also from the figure the number of cycles is more than 10. This figure graphically demonstrates the necessity of the control algorithm for a better drive characteristics. As the final step, the tunned PID controllers are applied to this problem and tried to get a better response from figure 7.

In this study, at first, three scalarization functions are evaluated on four optimization algorithms on car suspension problem which is explained in section 2. All of the optimization algorithms are run 100 iterations with 100 members, also a termination condition is defined to be sure that all of the optimization algorithms calculate the same number of functions evaluations. Table 1 presents the transient response parameters numerically, and the PID parameters are also reported in Table 2.

The results in Table 1 shows that, PSO, DE and FFA are both presents almost the same performance for different PID parameters (Table 2).

Table 2. PID controller parameters, where WS: Weighted Sum, and Eps: Epsilon.

PID Parameters	PSO-WS	PSO-Eps	PSO-Benson	DE-WS	DE-Eps	DE-Benson
KP	696	700	701	988	988	987
KI	266	18.85	206	10^{-3}	10^{-3}	498
KD	927	942	942	978	978	970
	FFA-WS	FFA-Eps	FFA-Benson	CA-WS	CA-Eps	CA-Benson
KP	743	995	999	141	439	141
KI	10^{-3}	1.306	748	159	160	159
KD	989	998	996	716	131	716

However, CA gives the best performance with respect to the transient performance among all heuristic algorithms. It can be concluded that the meta-heuristics of PSO, DE and FFA converges to local optimum, and remains in that solution. It should be also noted that some mechanisms may help to move that from local solution. CA gives the best performance among all other algorithms such that overshoot is almost 10% which is the lowest level and presents fastest rise and settling times. When the results are discussed with respect to the scalarization approaches, it is clear to see that three scalarization approaches are almost similar to each other; however, weighted sum method can able to give the best result in overall. But when compared inside CA algorithm epsilon and Benson's methods give the best results. Even there is a very slight difference between two scalarization methods, it can be seen that the Benson's approach perform better that ϵ-constrained approach.

Table 2 gives the PID parameters obtained from optimization algorithms. In general, PSO, DE, FFA and CE gives almost same parameters for scalarization approaches. For PSO algorithm, K_P and K_D are almost same with each other. DE algorithm seraches the solution especially at the boarder of the search space such that parameters are from lower and higher boarder values. Like PSO, FFA and CA give same values for K_D and K_I parameters, respectively.

As the second half of the implementation, the optimized PID controller is applied to the system under ramp input. The ramp input corresponds to a slightly change at the level of the road. In real world the roach change may happens rapidly, like holes at the road. In addition, speed bump like changes may happen on the road. Therefore, in this part of the study, the performance of the controller under speed bump like changes on the road is investigated. For this purpose, ramp input at different slopes are applied to the problem and obtained results are graphically demonstrated in Figure 8.

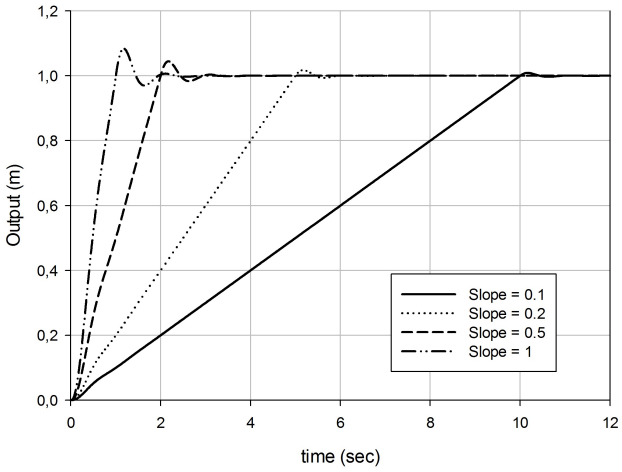


Figure 8. Car body position change for ramp input with different slopes.

Figure 8 gives the transient response of the CA-optimized PID controller for different slope ramp input. The figure shows that for slopes 0.1, 0.2, 0.5, and the overshoots becomes 8%, 5%, 2%, and 1.8% respectively. As the slope of the ramp increases, the overshoot and settling time decreases. In other words, as the change of level of the road is more slowly, then the response to this change becomes much better. Since the same optimized parameters are preferred for both implementations, the overall solution also supports the performance of the CA-based PID tuning methodology.

5. Conclusion

In this paper the active independent quarter-car suspension problem is solved by using PID controller. The PID parameter tuning is considered as the multiobjective optimization problem. First this problem is converted into single objective problem by using three scalarization approaches which are weighted sum, ϵ-constrained and Benson's methods. Then the obtained single objective problem is solved with four optimization algorithms which are called Particle Swarm Optimization (PSO), Differential Evolution (DE), Firefly Algorithm and Cultural Algorithm.

There are three aims of this study: i) propose a multiobjective PID tuning scheme for active independent car suspension system ii) compare three scalarization functions and show the implementation of these approaches and iii) compare four optimization algorithms. In total twelve implementations are made and compared with each other. The results showed that even the scalarization approaches present similar performances, among all optimization algorithms, CA gives the best performance. Also, it can be stated that, even a slight difference weighted sum still can be preferred as an efficient scalarization approach.

After CA is selected as the optimizer for PID parameters, a different input as the road change is applied. The ramp input with different slopes are implemented and results are discussed. The results indicates that as the slope of the ramp increases the transient response performance of the overall system is also increases due to the slow change at the level of the road.

In conclusion, CA algorithm presents the best performance among all other optimization algorithms discussed in this paper. The main reason is the local optimum problem of the other algorithms for this problem. Also, scalarization methods can be able to successfully applied to the PID tuning algorithm to convert the multiobjective problem into single objective one. The results indicate that weighted sum presents the best performance overall. The ϵ-constrained and Benson's methods give almost the same performance as PSO, DE, and FFA with a slight difference. As the future study, constrained-based scalarization approaches are deeply investigated with a different set of ϵ and reference points.

References

[1] Astrom K.J., and Hagglund T., The future of PID control, Control Engineering Practice, 9, 11, 1163-1175 (2001).

[2] Arruda L.V.R, Swiech M.C.S, Delgado M.R.B, and Neves-Jr F., PID control of MIMO process based on rank niching genetic algorithm, Applied Intelligence, 29, 3, 290-305 (2008).

[3] Neath M.J., Swain A.K., Madawala U.K., and Thrimawithana D.J., An Optimal PID Controller for a Bidirectional Inductive Power Transfer System Using Multiobjective Genetic Algorithm, IEEE Transactions on Power Electronics, 29, 3, 1523-1531 (2014).

[4] Lin C.-L., Jan H.-Y., and Shieh N.-C, GA-based multiobjective PID control for a linear brushless DC motor, IEEE/ASME Transactions on Mechatronics, 8, 1, 56-65 (2003).

[5] Herreros A., Baeyens E., and Pern J.R., Design of PID-type controllers using multiobjective genetic algorithms, ISA Transactions, 41, 4, 457-472 (2002).

[6] Reynoso-Meza G., Garcia-Nieto S., Sanchis J., and Blasco F.X., Controller Tuning by Means of Multi-Objective Optimization Algorithms: A Global Tuning Framework, IEEE Transactions on Control Systems Technology, 21, 2, 445-458 (2013).

[7] Reynoso-Meza G., Sanchis J., Blasco X., and Herrero J.M., Multiobjective evolutionary algorithms for multivariable PI controller design, Expert Systems with Applications, 39, 9, 7895-7907 (2012).

[8] Zhaoa S.-Z, Willjuice Iruthayarajanb M., Baskarc S., and Suganthan P.N., Multi-objective robust PID controller tuning using two lbests multi-objective particle swarm optimization, Information Sciences, 181, 16, 3323-3335 (2011).

[9] C.M. Seaman, A.A. Desrochers, and G.F. List, Multiobjective optimization of a plastic injection molding process, IEEE Transactions on Control Systems Technology, 2, 3, 157-168 (1994).

[10] Gambier A., MPC and PID control based on Multi-Objective Optimization, American Control Conference, 4727-4732 (2008).

[11] Naidua K., Mokhlisa H., and Bakar A.H.A., Multiobjective optimization using weighted sum Artificial Bee Colony algorithm for Load Frequency Control, International Journal of Electrical Power and Energy Systems, 55, 657-667 (2014).

[12] Hung M.-H., Shu L.-S., Ho S.-J., and Hwang S.-F., A Novel Intelligent Multiobjective Simulated Annealing Algorithm for Designing Robust PID Controllers, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 38, 2, 319-330 (2008).

[13] Tseng C.-S., and Chen B.-S., Multiobjective PID control design in uncertain robotic systems using neural network elimination scheme, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 31, 6, 632-644 (2001).

[14] Takahashi R.H.C., Peres P.L.D., and Ferreira P.A.V., Multiobjective H_2/H_∞ guaranteed cost PID design , IEEE Control Systems, 17, 5, 37-47 (2002).

[15] Tang K.S., Man K.F., Chen G., and Kwong S., An optimal fuzzy PID controller, IEEE Transactions on Industrial Electronics, 48, 4, 757-765 (2001).

[16] Liua G.P., and Daley S., Optimal-tuning PID control for industrial systems, Control Engineering Practice, 9, 11, 1185-1194 (2001).

[17] Ayalaa H.V.H., and Coelho L.S., Tuning of PID controller based on a multiobjective genetic algorithm applied to a robotic manipulator, Expert Systems with Applications, 49, 10, 8968-8974 (2012).

[18] Panda S., Multi-objective PID controller tuning for a FACTS-based damping stabilizer using Non-dominated Sorting Genetic Algorithm-II, International Journal of Electrical Power and Energy Systems, 33, 7, 12961308 (2011).

[19] Bourmistrova, A., Storey, I., and Subic A., Multi-objective optimisation of active and semi-active suspension systems with application of evolutionary algorithm, International Conference on Modelling and Simulation, 1217-1223 (2005).

[20] Xue X.D., Cheng K.W.E., and Xu C.D., Optimization of spring stiffness in automotive and rail active suspension systems, International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles, International Transportation Electrification, 1-6 (2016).

[21] Aldair A.A., Abdalla T.Y., and Alsaedee E.B., Design of PID controller using ABC for full vehicle nonlinear active suspension system with passenger seat, International Journal of Computer Applications, 167, 18-27 (2017).
[22] Rastegar N., and Khorram E., A combined scalarizing method for multiobjective programming problems, European Journal of Operational Research, 236, 1, 229237 (2014).
[23] Haimes Y., Ladson L., and Wismer D., On a bi criterion formulation of the problems of integrated system identification and system optimization, IEEE Transactions on System, Man, and Cybernetics, 1, 296-297 (1971).
[24] Benson H.P., Existence of efficient solutions for vector maximization problems, Journal of Optimization Theory and Applications, 26, 4, 569-580 (1978).
[25] Benson H.P., and Morin T.L., The vector maximization problem: Proper Efficiency and Stability, SIAM Journal of Applied Mathematics, 1-18 (1974).
[26] Soleimani-Damaneh M., and Zamani M., On Bensons scalarization in multiobjective optimization, Optimization Letters, Early Access, 1-6 (2016).
[27] Ehrgott M., Approximation algorithms for combinatorial multicriteria optimization problems, International Transactions in Operational Research, 7, 531 (2000).
[28] Deb K., Multi-Objective Optimization Using Evolutionary Algorithms, Publisher : Chichester, UK : Wiley (2002).
[29] Kennedy J., and Eberhart R.C., Particle Swarm Optimization, Proceedings of IEEE International Conference on Neural Networks, 1942-1948 (1995).
[30] Storn R., and Price K., Differential Evolution - a simple and efficient adaptive scheme for global optimization over continuous space, International Computer Science Institute, Technical Report, TR-95-012 (1995).
[31] Yang X.-S., Firefly Algorithm, Nature Inspired Metaheuristic Algorithms,79-90 (2008).
[32] Fister I., Yang X.-S., and Brest J., A Comprehensive Review of Firefly Algorithm, Swarm and Evolut,onary Computation, 13, 34-46 (2013).
[33] Yang X.-S., Firefly Algorithm, Stochastic Test Functions and Design Optimization, International Journal of Bio-Inspired Computation, 2, 78-84 (2010).
[34] Yang X.-S., Firefly Algorithms for Multimodal Optimization, Lecture Notes in Computer Science, 5792 LNCS, 169-178 (2009).
[35] Yang X.-S., Hosseini S., and Gandomi A.H., Firefly Algorithm for Solving Non-convex Economic Dispatch Problems with Valve Loading Effect, Applied Soft Computing Journal, 12, 1180-1186 (2012).
[36] Reynolds R.G., An Adaptive Computer Model of the Evolution of Agriculture for Hunter-Gatherers in the Valley of Oaxaca Mexico, PH.D. Thesis, Department of Computer Science, University of Michigan, (1976).
[37] Reynolds R.G., and Peng B. Cultural Algorithms: Modeling of How Cultures Learn to Solve Problems, International Conference on Tools with Artificial Intelligence, 1-7 (2004).
[38] Rychtyckyj, N., Using Cultural Algorithm in Industry, IEEE Swarm Intelligence Symposium, 1-8 (2003).
[39] Chung C., and Reynolds R.G., CAEP: An Evolution-based tool for Real-Valued Function Optimization using Cultural Algorithms, International Journal on Artificial Intelligence Tools, 7, 239-291 (1998).
[40] Jin X., and Reynolds R.G., Using Knowledge-based Evolutionary Computation to Solve Nonlinear Constraint Optimization Problems: A Cultural Algorithm Approach, Congress on Evolutionary Computation, 1672-1678 (1999).
[41] Reynolds R.G., and Saleem S. The Impact of Environmental Dynamics on Cultural Emergence, Oxford University Press, 1-10 (2003).

O. Tolga Altınöz *received his B.S. and M.S degrees in Department of Electrical and Electronics Engineering from Baskent University in 2003 and 2009 respectively, and Ph.D. degree from Ankara University in 2015 from Department of Electrical and Electronics Engineering. He is currently with the same department in Ankara University. His current research interests include evolutionary computation, optimization, control systems, power electronics, machine learning, and biomedical systems.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.195-200 (2018)
http://doi.org/10.11121/ijocta.01.2018.00601

RESEARCH ARTICLE

A simulation algorithm with uncertain random variables

Hasan Dalman

Department of Computer Engineering, Istanbul Gelisim University,Turkey
hdalman@gmail.com

ARTICLE INFO

Article History:
Received 05 April 2018
Accepted 22 April 2018
Available 25 April 2018

Keywords:
α- optimistic value
α- pessimistic value
Uncertain random variables
Uncertainty theory
Simulation

AMS Classification 2010:
68T37, 03E72

ABSTRACT

In many situations, uncertainty and randomness concurrently occur in a system. Thus this paper presents a new concept for uncertain random variable. Also, a simulation algorithm based on uncertain random variables is presented to approximate the chance distribution using pessimistic value and optimistic value. An example is also given to illustrate how to use the presented simulation algorithm.

1. Introduction

Liu [1] introduced the uncertain random variable for modeling complex systems. In other words; uncertain random variable is improved to illustrate the phenomenon which mixes uncertainty with randomness. If we receive historical data from the sample, we can estimate the probability distribution. But if we have a new product, we can not achieve the probability distribution of this new product owing to lack of data. In this case, we run across both randomness and human uncertainty. Human uncertainty is investigated by some scholars. As a branch of mathematics based on normality, duality, subadditivity and product axioms, uncertainty theory was introduced by Liu [2] in 2007. Gao [3] presented uncertain bimatrix game. Yang and Gao [4] studied uncertain differential game. Gao and Qin [5] introduced the degree connectivity of uncertain graph. Dalman [6] presented a model for the uncertain multi-item solid transportation problem. Dalman [7] constructed models of uncertain random multi-item solid transportation problem. To model uncertain random event, Liu [8] introduced the chance theory to networks optimization problem. Some scholars derived properties of uncertain random entropy [9, 10].

By employing chance theory, an uncertain random project scheduling problem is presented by Ke et al. [11]. They introduced an uncertain random simulation which randomly produces the sample points. But, this algorithm is ambivalent due to produces different values at different time. A simulation algorithm is presented to solve uncertain random shortest path problem by Sheng and Gao [12].

Therefore this paper presents an algorithm which includes the inverse uncertainty distribution and uniformly produces the sample points. It has powerful performances on the reliability and stability than the algorithms in [11, 12]. The paper is built as follows: some basic knowledge of uncertainty theory and chance theory is presented in section II. Section III shows some formulas for uncertain random variables presents an uncertain random simulation algorithm. A numerical example is presented in Section IV. Finally, this paper closes in Section V.

2. Preliminaries

2.1. Uncertainty theory

Let Γ be a nonempty set, $\mathcal{L}$ be a σ-algebra over Γ and $\mathcal{M}$ be an uncertain measure. Then (Γ, $\mathcal{L}$, $\mathcal{M}$) is a measurable space. A set function $\mathcal{M}$: $\mathcal{L}\rightarrow [0,1]$ is called an uncertain measure if it satisfies the following four axioms:

Axiom 1. *(Normality Axiom)(Liu [2]):* $\mathcal{M}\{\Gamma\}$ $=1$ *for the universal set* Γ.

Axiom 2. *(Duality Axiom)(Liu [2]):* $\mathcal{M}\{\Lambda\}$ + $\mathcal{M}\{\Lambda^c\}=1$ *for any event* Λ.

Axiom 3. *(Subadditivity Axiom)(Liu [2]): For every countable sequence of events* $\Lambda_1, \Lambda_2, \cdots$, *we have*

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty}\Lambda_i\right\} \leq \sum_{i=1}^{\infty}\mathcal{M}\{\Lambda_i\}.$$

Axiom 4. *(Product Axiom)(Liu [13]): Let* $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ *be uncertainty spaces for* $k = 1, 2, \cdots$. *The product uncertain measure* $\mathcal{M}$ *is an uncertain measure satisfying*

$$\mathcal{M}\left\{\prod_{k=1}^{\infty}\Lambda_k\right\} = \bigwedge_{k=1}^{\infty}\mathcal{M}_k\{\Lambda_k\}$$

where Λ_k *are arbitrarily chosen events from* $\mathcal{L}_k$ *for* $k = 1, 2, \cdots$, *respectively.*

Definition 1. *(see [2]): An uncertain variable is a function* ξ *from an uncertainty space* $(\Gamma, \mathcal{L}, \mathcal{M})$ *to the set of real numbers such that* $\{\xi \in B\}$ *is an event for any Borel set* B *of real numbers.*

Remark 1: Note that the event $\{\xi \in B\}$ is a subset of the universal set $\{\xi \in B\} = \{\gamma \in \Gamma | \xi(\gamma) \in B\}$.

Definition 2. *(Liu [14]): An uncertainty distribution* $\Phi(x)$ *is said to be regular if it is a continuous and strictly increasing function with respect to* x *at which* $0 < \Phi(x) < 1$, *and*

$$\lim_{x\rightarrow -\infty}\Phi(x) = 0, \lim_{x\rightarrow +\infty}\Phi(x) = 1.$$

Definition 3. *(Liu [14]): Let* ξ *be an uncertain variable with regular uncertainty distribution* $\Phi(x)$. *Then the inverse function* $\Phi^{-1}(\alpha)$ *is called the inverse uncertainty distribution of* ξ.

Theorem 1. *(Liu [14]): Let* $\xi_1, \xi_2, \cdots, \xi_n$ *be independent uncertain variables with regular uncertainty distributions* $\Phi_1, \Phi_2, \ldots, \Phi_n$, *respectively. If* $f(\xi_1, \xi_2, \cdots, \xi_n)$ *is strictly increasing with respect to* $\xi_1, \xi_2, \cdots, \xi_m$ *and strictly decreasing with respect to* $\xi_{m+1}, \xi_{m+2}, \cdots, \xi_n$, *then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n) \tag{1}$$

has an inverse uncertainty distribution.

$$\begin{aligned}\Psi^{-1}(\alpha) =& f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha),\\ & \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)).\end{aligned} \tag{2}$$

2.2. Chance theory

Definition 4. *(Liu [1]): Let* $(\Gamma, \mathcal{L}, \mathcal{M})$ *be an uncertainty space and let* $(\Omega, \mathcal{A}, \Pr)$ *be a probability space. Then the product* $(\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{A}, \Pr)$ *is called a chance space*

$$(\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{A}, \Pr) = (\Gamma \times \Omega, \mathcal{L} \times \mathcal{A}, \mathcal{M} \times \Pr).$$

Definition 5. *(see [1]): An uncertain random variable is a function* ξ *from a chance space* $(\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{A}, \Pr)$ *to the set of real numbers such that* $\{\xi \in B\}$ *is an event in an event in* $\mathcal{L} \times \mathcal{A}$ *for any Borel set* B *of real numbers.*

Theorem 2. *(see [14]): Let* $\xi_1, \xi_2, \cdots, \xi_n$ *be uncertain random variables on the chance space* $(\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{A}, \Pr)$, *and let* f *be a measurable function. Then*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$$

is an uncertain random variable determined by

$$\xi(\gamma, \omega) = f(\xi_1(\gamma, \omega), \xi_2(\gamma, \omega), \cdots, \xi_n(\gamma, \omega))$$

for all $(\gamma, \omega) \in \Gamma \times \Omega$.

Theorem 3. *(Liu [15]): Let* $\eta_1, \eta_2, \cdots, \eta_m$ *be independent random variables with probability distributions* $\Psi_1, \Psi_2, \cdots, \Psi_m$, *respectively, and let* $\tau_1, \tau_2, \cdots, \tau_n$ *be independent uncertain variables. Assume* f *is a measurable function. Then the uncertain random variable*

$$\xi = f(\eta_1, \eta_2, \cdots, \eta_m, \tau_1, \tau_2, \cdots, \tau_n)$$

has a chance distribution

$$\begin{aligned}\Phi(x) = \int_{\Re^m} & F(x; y_1, y_2, \cdots, y_m)\mathrm{d}\Psi_1(y_1)\\ & \mathrm{d}\Psi_2(y_2)\cdots\mathrm{d}\Psi_m(y_m).\end{aligned} \tag{3}$$

where $F(x; y_1, y_2, \cdots, y_m)$ *is the uncertainty distribution of the variable*

$$f(y_1, y_2, \cdots, y_m, \tau_1, \tau_2, \cdots, \tau_n).$$

Definition 6. *(Liu [1]): Let* ξ *be an uncertain random variable. Then its chance distribution is defined by*

$$\Phi(x) = \mathrm{Ch}\{\xi \leq x\} \tag{4}$$

for any $x \in \Re$.

Theorem 4. *Let* ξ *be an uncertain random variable. Then its expected value is*

$$E[\xi] = \int_0^{+\infty}\mathrm{Ch}\{\xi \geq \mathrm{r}\}\mathrm{dr} - \int_{-\infty}^{0}\mathrm{Ch}\{\xi \leq \mathrm{r}\}\mathrm{dr} \tag{5}$$

provided that at least one of the two integrals is finite.

Theorem 5. *Let ξ be an uncertain random variable with regular chance distribution Φ. If the expected value exists, then*

$$E[\xi] = \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha. \tag{6}$$

In order to obtain the chance distribution, we prove following theorems and develop a simulation algorithm.

3. A simulation algorithm with uncertain random variables

In this section, by employing the above definition and theorems, the chance distribution will be obtained. To do this, the following theorems is proved and besides a simulation algorithm is developed.

In fact, Formula (3) is a theoretical formula, which is not easy to use in most cases due to the complexity of chance distribution function. To overcome the difficulty, an uncertain random simulation is proposed to evaluate the chance distribution. First, we introduce the concepts of $\alpha-$pessimistic value and $\alpha-$optimistic value for an uncertain random variable. Then, we approximate the chance distribution, $\alpha-$pessimistic value and $\alpha-$optimistic value by using a numerical integration method.

Definition 7. *Let ξ be an uncertain random variable on chance space $(\Gamma, \mathcal{L}, \mathcal{M}) \times (\Omega, \mathcal{A}, \mathrm{Pr})$ and $\alpha \in (0, 1]$. Then,*

$$\xi_{\inf}(\alpha) = \inf\{r|\mathrm{Ch}\{\xi \leq r\} \geq \alpha\} \tag{7}$$

and

$$\xi_{\sup}(\alpha) = \sup\{r|\mathrm{Ch}\{\xi \geq r\} \geq \alpha\} \tag{8}$$

are called the $\alpha-$pessimistic value and the $\alpha-$optimistic value of ξ, respectively.

Note that Random variables and uncertain variables are special uncertain random variables. The $\alpha-$pessimistic value and the $\alpha-$optimistic value of linear uncertain variable $\mathcal{L}(a, b)$ are $\xi_{\inf}(\alpha) = (1-\alpha)a + \alpha b$ and $\xi_{\sup}(\alpha) = \alpha a + (1-\alpha)b$.

Theorem 6. *Let ξ be an ordinary uncertain random variable and $\alpha \in (0, 1]$. Then, we have*

$$\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} = \alpha. \tag{9}$$

Proof. Since the chance distribution is continuous, it follows from the definition of the $\alpha-$pessimistic value for each $\alpha \in (0, 1]$, we have $\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} = \lim_{\mathrm{n}\to\infty} \mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha) - 1/\mathrm{n}\} \leq \alpha$, and $\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} = \lim_{n\to\infty} \mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha) + 1/n\} \geq \alpha$, which imply that $\mathrm{Ch}\{\xi \leq \xi_{\inf}(\alpha)\} = \alpha$ holds. The theorem is proved. □

Theorem 7. *Let ξ be an ordinary uncertain random variable and $\alpha \in (0, 1]$. Then, we have*

$$\mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} = \alpha. \tag{10}$$

Proof. Since the continuity of chance distribution, it follows from the definition of the $\alpha-$optimistic value for each $\alpha \in (0, 1]$ that $\mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} = \lim_{\mathrm{n}\to\infty} \mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha) - 1/\mathrm{n}\} \geq \alpha$ and $\mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} = \lim_{n\to\infty} \mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha) + 1/n\} \leq \alpha$, which imply that $\mathrm{Ch}\{\xi \geq \xi_{\sup}(\alpha)\} = \alpha$ holds. The theorem is proved. □

Theorem 8. *Let ξ be an uncertain random variable and $\alpha \in (0, 1]$. Then, we have*

$$\xi_{\inf}(\alpha) = \Phi^{-1}(\alpha). \tag{11}$$

Proof. It follows from Definition 7 immediately. □

Theorem 9. *Let ξ be an uncertain random variable and $\alpha \in (0, 1]$. Then, we have*

$$\xi_{\inf}(\alpha) = \xi_{\sup}(1-\alpha). \tag{12}$$

Proof. It follows from Equation (8) that

$$\mathrm{Ch}\{\xi \geq \xi_{\sup}(1-\alpha)\} = 1 - \alpha.$$

Thus,

$$\begin{aligned}\mathrm{Ch}\{\xi \leq \xi_{\sup}(1-\alpha)\} &= 1 - \mathrm{Ch}\{\xi \geq \xi_{\sup}(1-\alpha)\} \\ &= 1 - (1-\alpha) \\ &= \alpha\end{aligned}$$

Thus, we have $\xi_{\inf}(\alpha) = \xi_{\sup}(1-\alpha)$. The theorem is proved. □

Theorem 10. *Let ξ be an uncertain random variable and $\alpha \in (0, 1]$. Then, we have*

$$\xi_{\sup}(\alpha) = \Phi^{-1}(1-\alpha) \quad \text{and} \quad \xi_{\sup}(1-\alpha) = \Phi^{-1}(\alpha). \tag{13}$$

Proof. It follows from Theorems 8 and 9. □

Theorem 11. *Let ξ be an ordinary uncertain random variable. Then, we have*

$$E[\xi] = \int_0^1 \xi_{\inf}(\alpha)\mathrm{d}\alpha. \tag{14}$$

Proof. Since $\Phi(x)$ is strictly increasing and continuous, we get

$$\begin{aligned}\xi_{\inf}(\alpha) &= \inf\{\mathrm{r}|\mathrm{Ch}\{\xi \leq \mathrm{r}\} \geq \alpha\} \\ &= \inf\{\mathrm{r}|\Phi(\mathrm{r}) \leq \alpha\} \\ &= \Phi^{-1}(\alpha).\end{aligned}$$

According to Definition (), we have $E[\xi] = \int_0^1 \xi_{\inf}(\alpha)\mathrm{d}\alpha$. The Theorem is proved. □

Theorem 12. *Let ξ be an ordinary uncertain random variable. Then, we have*

$$E[\xi] = \int_0^1 \xi_{\sup}(\alpha)\mathrm{d}\alpha. \tag{15}$$

Proof. Since $\Phi(x)$ is strictly increasing and continuous, we get

$$\begin{aligned}\xi_{\sup}(\alpha) &= \sup\{\mathrm{r}|\mathrm{Ch}\{\xi \geq \mathrm{r}\} \geq \alpha\} \\ &= \sup\{\mathrm{r}|\Phi(\mathrm{r}) \leq 1-\alpha\} \\ &= \Phi^{-1}(1-\alpha).\end{aligned}$$

Thus, for all $\alpha \in (0,1)$, we have

$$\begin{aligned}\int_0^1 \xi_{\sup}(\alpha)\mathrm{d}\alpha &= \int_0^1 \Phi^{-1}(1-\alpha)\mathrm{d}\alpha \\ &= \int_0^1 \Phi^{-1}(\alpha)\mathrm{d}\alpha = \mathrm{E}[\xi]\end{aligned}$$

The theorem is proved. □

Theorem 13. *Let ξ be an ordinary uncertain random variable. Then, we have*

$$E[\xi] = \frac{1}{2}\int_0^1 \left[\xi_{\inf}(\alpha) + \xi_{\sup}(\alpha)\right]\mathrm{d}\alpha. \tag{16}$$

Proof. It follows directly from Theorem 11 and Theorem 12. □

Theorem 14. *Let $\eta_1, \eta_2, \cdots, \eta_m$ be independent random variables with probability distributions $\Psi_1, \Psi_2, \cdots, \Psi_m$, and let $\tau_1, \tau_2, \cdots, \tau_n$ be independent uncertain variables with regular uncertainty distributions $\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_n$. Assume $f(\eta_1, \eta_2, \cdots, \eta_m, \tau_1, \tau_2, \cdots, \tau_n)$ is strictly increasing with respect to $\tau_1, \tau_2, \cdots, \tau_k$ and strictly decreasing with respect to $\tau_{k+1}, \tau_{k+2}, \cdots, \tau_n$. Then*

$$\xi = f(\eta_1, \eta_2, \cdots, \eta_m, \tau_1, \tau_2, \cdots, \tau_n)$$

has a chance distribution

$$\begin{aligned}\Phi(x) = \int_{\Re^m} F(x; y_1, y_2, \cdots, y_m)\mathrm{d}\Psi_1(y_1) \\ \mathrm{d}\Psi_2(y_2)\cdots\mathrm{d}\Psi_m(y_m).\end{aligned} \tag{17}$$

where $F(x; y_1, y_2, \cdots, y_m)$ is determined by its inverse uncertainty distribution

$$\begin{aligned}&F^{-1}(y_1, y_2, \ldots, y_m, \Upsilon_1^{-1}(\alpha), \ldots, \\ &\Upsilon_2^{-1}(\alpha), \ldots, \Upsilon_n^{-1}(\alpha)),\end{aligned} \tag{18}$$

$$\begin{aligned}&F^{-1}(y_1, y_2, \cdots, y_m, (\xi_1)_{\sup}(1-\alpha), \cdots, \\ &(\xi_k)_{\sup}(1-\alpha), (\xi_{k+1})_{\sup}(\alpha), \cdots, (\xi_n)_{\sup}(\alpha)).\end{aligned} \tag{19}$$

Proof. It follows from Theorem 1 that inverse uncertainty distribution of $F(x; y_1, y_2, \cdots, y_m)$ is determined by

$$\begin{aligned}&F^{-1}(y_1, y_2, \cdots, y_m, \Upsilon_1^{-1}(\alpha), \cdots, \Upsilon_k^{-1}(\alpha), \\ &\Upsilon_{k+1}^{-1}(1-\alpha), \cdots, \Upsilon_n^{-1}(1-\alpha)).\end{aligned} \tag{20}$$

According to Theorem 10, we substitute $\Upsilon_1^{-1}(\alpha), \cdots, \Upsilon_k^{-1}(\alpha)$ with $(\xi_1)_{\sup}(1-\alpha), \cdots, (\xi_k)_{\sup}(1-\alpha)$ and $\Upsilon_{k+1}^{-1}(1-\alpha), \cdots, \Upsilon_n^{-1}(1-\alpha)$ with $(\xi_{k+1})_{\sup}(\alpha), \cdots, (\xi_n)_{\sup}(\alpha)$. Thus, Formula (19) holds. The theorem is completed. □

Theorem 15. *Let $\eta_1, \eta_2, \cdots, \eta_m$ be independent random variables with probability distributions $\Psi_1, \Psi_2, \cdots, \Psi_m$, and let $\tau_1, \tau_2, \cdots, \tau_n$ be independent uncertain variables with regular uncertainty distributions $\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_n$. Assume $f(\eta_1, \eta_2, \cdots, \eta_m, \tau_1, \tau_2, \cdots, \tau_n)$ is strictly increasing with respect to $\tau_1, \tau_2, \cdots, \tau_k$ and strictly decreasing with respect to $\tau_{k+1}, \tau_{k+2}, \cdots, \tau_n$. Then*

$$\xi = f(\eta_1, \eta_2, \cdots, \eta_m, \tau_1, \tau_2, \cdots, \tau_n)$$

has a chance distribution

$$\begin{aligned}\Phi(x) = \int_{\Re^m} F(x; y_1, y_2, \cdots, y_m)\mathrm{d}\Psi_1(y_1) \\ \mathrm{d}\Psi_2(y_2)\cdots\mathrm{d}\Psi_m(y_m).\end{aligned} \tag{21}$$

where $F(x; y_1, y_2, \cdots, y_m)$ is determined by its inverse uncertainty distribution

$$\begin{aligned}&F^{-1}(y_1, y_2, \ldots, y_m, \Upsilon_1^{-1}(\alpha), \ldots, \\ &\Upsilon_2^{-1}(\alpha), \ldots, \Upsilon_n^{-1}(\alpha))\end{aligned} \tag{22}$$

$$\begin{aligned}&F^{-1}(y_1, y_2, \ldots, y_m, (\xi_1)_{\inf}(\alpha), \cdots, (\xi_k)_{\inf}(\alpha), \\ &(\xi_{k+1})_{\inf}(1-\alpha), \cdots, (\xi_n)_{\inf}(1-\alpha)).\end{aligned} \tag{23}$$

Proof. The proof is similar to that of Theorem 14. □

According to Theorems 1, 17, 8, 9 and 10, we design the following uniform discretization algorithm to simulate $\Phi(x)$, α-pessimistic value and α-optimistic value. The presented algorithm is very flexible. Because it can even simulate the empirical distribution.

Algorithm 1 (Uniform Discretization Algorithm)

Step 1. Discretize the range of the random variable η_i into N_i equally spaced points.

Step 2. Discretize α into K equally spaced points.

Step 3. Calculate $F^{-1}(\alpha; y_1, y_2, \cdots, y_m)$.

Step 4. Calculate $F(x; y_1, y_2, \cdots, y_m) =$

$$\begin{cases} 0 & if \quad x \leq x_1, \\ \alpha_i + (\alpha_{i+1} - \alpha_i)\dfrac{x - x_i}{x_{i+1} - x_i} & if \quad x_i \leq x \leq x_{i+1}, \\ & 1 \leq i \leq K, \\ 1 & if \quad x \geq x_K, \end{cases}$$

Step 5. Apply numerical integration to compute $\Phi(x)$, α-pessimistic value and α-optimistic value.

The uniform discretization algorithm is illustrated by the following example.

4. A numerical example

Example Suppose that η_1 and η_2 are independent random variables with probability distributions $U(1,2)$ and $U(2,4)$, and suppose that τ_1 and τ_2 are independent uncertain variables with uncertainty distributions $\mathcal{L}(1,5)$ and $\mathcal{L}(1,3)$. Then $\xi = \eta_1 + \eta_2 + \tau_1 - \tau_2$ is an uncertain random variable. Assume that ξ has chance distribution $\Phi(x)$.

For the sake of simplicity, we set $N_1 = 10, N_2 = 10, K = 10$. We also can assign a large number to N_1, N_2 and K. The probability distribution function of η_1 and η_2 are

$$\Psi_1(y_1) = \begin{cases} 0 & if \quad y_1 \leq 1, \\ y_1 - 1 & if \quad 1 \leq y_1 \leq 2, \text{and} \\ 1 & if \quad y_1 \geq 2, \end{cases}$$

$$\Psi_2(y_2) = \begin{cases} 0 & if \quad y_2 \leq 2, \\ \dfrac{y_2 - 2}{2} & if \quad 2 \leq y_2 \leq 4, \\ 1 & if \quad y_2 \geq 4. \end{cases}$$

Then, we can have discrete forms of $\Psi_1(y_1)$ and $\Psi_1(y_2)$ in which $y_1 = 1 + 0.1 \cdot i$ and $y_2 = 2 + 0.2 \cdot j$ for $i = 1, 2, \cdots, 10, j = 1, 2, \cdots, 10$. The inverse uncertainty distribution function of τ_1 and τ_2 are $(\tau_1)_{\inf}(\alpha) = (1-\alpha) \cdot 1 + \alpha \cdot 5 = 1 + 4 \cdot \alpha$ and $(\tau_2)_{\inf}(1-\alpha) = \alpha \cdot 1 + (1-\alpha) \cdot 3 = 3 - 2 \cdot \alpha$.

The chance distribution of ξ is

$$\Phi(x) = \frac{1}{2} \int_2^4 \int_1^2 F(x; y_1, y_2) \mathrm{d}y_1 \mathrm{d}y_2 \qquad (24)$$

where $F(x; y_1, y_2)$ is obtained by the inverse uncertainty distribution function $F^{-1}(\alpha; y_1, y_2) = y_1 + y_2 + (1 + 4 \cdot \alpha) - (3 - 2 \cdot \alpha)$ for each $\alpha \in (0, 1]$. This implies that $F^{-1}(\alpha; y_1, y_2) = x_k, 1 \leq k \leq 10$ for each $\alpha \in (0, 1]$. The value of $F^{-1}(\alpha; y_1, y_2)$ is listed on the Table I.

Then, according to Step 4, we obtain $F(x; y_1, y_2)$. Now, we get the chance distribution function of ξ.

$$\begin{aligned} \Phi(x) =& \frac{1}{2} \int_2^4 \int_1^2 F(x; y_1, y_2) \mathrm{d}y_1 \mathrm{d}y_2 \\ =& \frac{1}{2} \sum_{i=1}^{10} \sum_{j=1}^{10} F(x; 1 + 0.1 \cdot i, 2 + 0.2 \cdot j) \cdot 0.1 \cdot 0.2 \end{aligned}$$

Applying the above presented algorithm, the chance distribution of ξ is obtained as (Figure 1): $\xi_{\inf}(0.2)$=3.9 and $\xi_{\sup}(0.2)$=7.4.

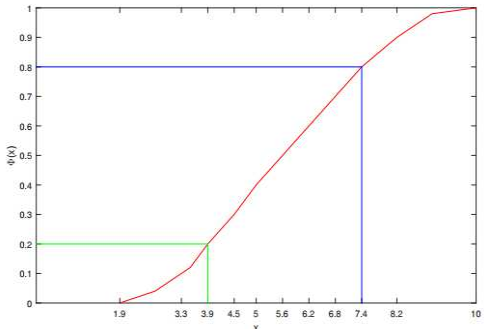


Figure 1. The chance distribution of ξ.

5. Conclusion

Here, an uncertain random simulation algorithm is presented to illustrate chance distribution. The presented simulation method can be generally applied to uncertain random optimization by approximating chance constraints. The results obtained show that the presented algorithm is the successful for the uncertain random simulation.

Acknowledgments

The author would like to thank the anonymous reviewers for their valuable comments.

References

[1] Liu, Y. (2013). Uncertain random variables: A mixture of uncertainty and randomness. *Soft Computing*, 17(4), 625-634.

[2] Liu, B. (2007). Uncertainty theory, 2nd ed., Springer-Verlag, Berlin, Germany.

[3] Gao, J. (2013). Uncertain bimatrix game with applications. *Fuzzy Optimization and Decision Making*, 12(1), 65-78.

[4] Yang, X., and Gao, J. (2016). Linearquadratic uncertain differential game with application to resource extraction problem. *IEEE Transactions on Fuzzy Systems*, 24(4), 819-826.

[5] Gao, Y., and Qin, Z. (2016) On computing the edge-connectivity of an uncertain graph. *IEEE Transactions on Fuzzy Systems*, 24(4), 981-991.

[6] Dalman, H. (2018). Uncertain programming model for multi-item solid transportation problem. *International Journal of Machine Learning and Cybernetics*, 9(4), 559-567.

[7] Dalman, H. (2018). Uncertain random programming models for fixed charge multi-item solid transportation problem, *New Trends in Mathematical Sciences*, 6(1), 37-51.

[8] Liu, B. (2014). Uncertain random graph and uncertain random network. *Journal of Uncertain Systems*, 8(1), 3-12.

[9] Zhou, J., Yang, F., and Wang, K. (2014). Multi-objective optimization in uncertain random environments. *Fuzzy Optimization and Decision Making*, 13(4), 397-413.

[10] Ahmadzade, H., Gao, R., and Zarei, H. (2016). Partial quadratic entropy of uncertain random variables. *Journal of Uncertain Systems*, 10(4), 292-301.

[11] Ke, H., Liu, H., and Tian, G. (2015). An uncertain random programming model for project scheduling problem. *International Journal of Intelligent Systems*, 30(1), 66-79.

[12] Sheng, Y., and Gao, Y. (2016). Shortest path problem of uncertain random network. *Computers and Industrial Engineering*, 99, 97-105.

[13] Liu, B. (2009). Some research problems in uncertainty theory. *Journal of Uncertain Systems*, 3(1), 3-10.

[14] Liu, B. (2010). Uncertainty Theory: A Branch of Mathematics for Modeling Human Uncertainty, Springer-Verlag, Berlin, Germany.

[15] Liu, Y. (2013). Uncertain random programming with applications. *Fuzzy Optimization and Decision Making*, 12(2), 153-169.

Hasan Dalman *received the Ph.D. degree in Mathematical Engineering from Yildiz Technical University in 2015. He is currently assistant professor in the Department of Computer Engineering, Istanbul Gelisim University, Turkey. His current research interests include fuzzy systems, uncertainty theory and mathematical optimization.*

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.2, pp.201-207 (2018)
https://doi.org/10.11121/ijocta.01.2018.00564

RESEARCH ARTICLE

Gain scheduling LQI controller design for LPV descriptor systems and motion control of two-link flexible joint robot manipulator

Yusuf Altun*

Department of Computer Engineering, Faculty of Engineering, Duzce University, Turkey
altunyusf@hotmail.com

ARTICLE INFO

Article history:
Received: 21 November 2017
Accepted: 12 April 2018
Available Online: 27 April 2018

Keywords:
Robotic manipulator
Gain scheduling
LQR controller
Descriptor system

AMS Classification 2010:
70E60, 70Q05, 93B51, 93B50, 93C85, 93C40, 93C95, 93D09, 93D21

ABSTRACT

This paper proposes a gain scheduling linear quadratic integral (LQI) servo controller design, which is derived from linear quadratic regulator (LQR) optimal control, for non-singular linear parameter varying (LPV) descriptor systems. It is assumed that state space matrices are non-singular since many mechanical systems do not have any non-singular matrices such as the natural state space forms of robotic manipulator, pendulum and suspension systems. A controller design is difficult for the systems due to rational LPV case. Therefore, the proposed gain scheduling controller is designed without the difficulty. Accordingly, the motion control design is implemented for two-link flexible joint robotic manipulator. Finally, the control system simulation is performed to prove the applicability and performance.

1. Introduction

Many researchers have recently considered the problem of robust controller design on linear parameter varying or linear time invariant systems in the regular state-space form such as in [1–14], where the systems are common linear parameter varying (LPV) systems. Unlike the regular state-space forms, descriptor systems enable an expression which includes algebraic conditions on physical factors. Therefore, these systems have attracted attention for the past decade because many systems such as mechanical systems have this structure, and the systems have ability to describe many systems such as chemical processes, robotic systems, aircrafts, etc. Thus, many researchers have investigated in the literature, for instance, robust stability analysis and stabilisation [15–17], robust controllability/ observability analysis [18], H_2/H_∞ norm characterization and control [19–27], robust filtering [21], positive real analysis and control [22].

They have two types which are singular and non-singular descriptor systems. Singular systems, which are also called differential algebraic equations (DAE), are generally used to describe some systems including algebraic constraints. On the other hand, many of mechanical systems such as robotic manipulators, pendulum systems and suspension systems have not any singular matrices although they can be represented as singular system form by including algebraic constraints. This is because, they are naturally in the form of non-singular descriptor state space forms due to using Euler-Lagrange method based on energy for the mathematical modelling. On the other hand, non-singular LPV descriptor systems can occur rational LPV form due to uncertain parameters, for which the controller design is difficult. Accordingly, the mentioned papers are mostly for the singular systems.

Flexible robotic manipulators have many advantages on the rigid robots which need less material, lighter in weight, less power, smaller actuators, more manoeuvrable and transportable, and thus their cost are lower. These robots have a wide application area as an industrial robot in industry. In addition, two-link flexible manipulators are preferred because they present more flexibility for applications. Nevertheless, their control is difficult to attain exact positioning. Also, the complexity of problem arises because they are multi-input multi-output (MIMO) systems, which are affected by several factors such as payload changing and vibration effects of between links. Moreover, they have nonlinearity and some uncertain parameters. Therefore, numerous researchers have investigated for the control of robotic manipulators in [28–34]. For instance, H_∞ control and μ-synthesis are in [28, 29, 33].

*Corresponding author

In [30], the control of flexible robotic manipulator is dealt with finite element theory. In [31], a robust control method of a two-link flexible manipulator with neural network. In [32], H_∞ control is performed for LPV descriptor model with affine parametric dependence by using a linear fractional representation (LFR). In [34], LPV control is performed by converting equivalent to a rational LPV system. In [35], general H_∞ LPV control is designed for the non-singular descriptor flexible robotic manipulator.

Thus, the designs in above papers are mostly for the singular LPV systems. In addition, a gain scheduling Linear Quadratic Integral (LQI) controller has not been tackled for the non-singular descriptor systems in the literature. Accordingly, many mechanical systems are naturally in the form of non-singular LPV system. In this paper, we consider the construction of a gain scheduling LQI controller for a non-singular uncertain descriptor system. The controller is applied to two-link robotic manipulator. Natural two-link robotic manipulator model has not any singular matrices, but it and similar mechanical systems can be converted to singular form by LPV conversion methods as in [22, 36, 37], which is not in the focus of paper. Because of avoiding the complexity of rational LPV systems, the proposed controller can be applied to a robotic manipulator system or any non-singular LPV descriptor system without using rational LPV form. Finally, the main aim of the paper is to design a static controller without any extra methods and assumptions for the rational LPV systems. From the references and the other literature studies, there is not a gain-scheduling controller LQI design for the LPV descriptor systems without conversion methods.

2. Two link robotic manipulator model

The robotic manipulator is shown in Figure 1. It is a two-link flexible planar manipulator which is driven by geared two DC motors. θ_1 and θ_2 are the shoulder and elbow joint angles, respectively. τ_1 and τ_2 are the corresponding control torques. The second-order form of the manipulator nonlinear motion equations [28] are as in (1) where $M(\theta_2)$ is the inertia matrix, D is the damping matrix, and K is the stiffness matrix, F is input vector and. $M(\theta_2)$ is as in (2) where $M(\pi)$ and $M(\pi/2)$ are given by (3).

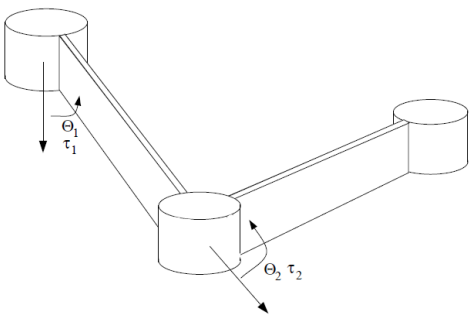


Figure 1. Two link robotic manipulator schema [28].

$$M(\theta_2)\ddot{q}(t) + D\dot{q}(t) + Kq(t) = Fu(t) \quad (1)$$

$$M(\theta_2) = M(\pi/2) + \cos(\theta_2)\left[M(\pi/2) - M(\pi)\right] \quad (2)$$

$$\mathrm{M}\left(\frac{\pi}{2}\right) = \begin{bmatrix} 34.7077 & 9.7246 & 23.6398 & 5.9114 \\ 9.7246 & 9.8783 & 9.7246 & 5.9114 \\ 23.6398 & 9.7246 & 17.5711 & 5.9114 \\ 5.9114 & 5.9114 & 5.9114 & 3.7233 \end{bmatrix},$$
$$\mathrm{M}(\pi) = \begin{bmatrix} 17.0296 & 0.8856 & 9.7776 & 0.8430 \\ 0.8856 & 9.8783 & 4.7016 & 5.9114 \\ 9.7776 & 4.7016 & 7.5249 & 3.0311 \\ 0.8430 & 5.9114 & 3.0311 & 3.7233 \end{bmatrix} \quad (3)$$

Also, D the damping matrix, and K the stiffness matrix and F input vector are

$$\mathrm{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.09 & 0 \\ 0 & 0 & 0 & 0.05 \end{bmatrix},$$

$$\mathrm{K} = = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 89.1473 & 0 \\ 0 & 0 & 0 & 45.6434 \end{bmatrix}, \mathrm{F} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

3. Gain-scheduling LQI controller design

LQR control is an optimal control method, which is commonly used for the state feedback design in the literature. But, it has no output matrix. Therefore, LQI control, which is based on LQR, has the output feedback with integral action. Thus, LQR is regulator while LQI is servo design. That is, LQI controller is used for the reference tracking. In this paper, gain scheduling LQI controller design is proposed for the non-singular LPV descriptor form, and the method is applied to two link robotic manipulator.

$$\dot{x}(t) = Ax(t) + Bu(t),\ x(0) = x_0 \quad (4)$$

Consider the linear time invariant (LTI) system in (4), J the performance index (or cost functional) is defined in (5) for the state feedback optimal controller $u(t) = -Kx(t)$, where $Q^{n\times n}$ is positive semidefinite matrix and $R^{m\times m}$ is positive symmetric matrix. The aim of LQR control is to design a state feedback controller $u(t) = -Kx(t)$ which minimizes performance index J and stabilizes the system. In addition, LQI controller is obtained by including output and reference error. Therefore, if integral action is included in system, LQI controller is designed. Also, its effect is to drop steady state errors.

$$J(t) = \int_0^\infty \left(x^T(t)Qx(t) + u^T(t)Ru(t)\right)dt \quad (5)$$

Robotic manipulators are in the form of (6) where time-varying parameter set is given by (7). It is assumed that $E(\upsilon)$ is non-singular matrix for all υ.

$$\begin{aligned} &E(\upsilon)\dot{x}(t)=Ax(t)+Bu(t),\ \ x(0)=x_0 \\ &y=Cx(t) \end{aligned} \tag{6}$$

$$\begin{aligned} \upsilon_D &= \left\{\dot{\upsilon}(t)\in\Re^n : \underline{\dot{\upsilon}_i} \ \le\ \dot{\upsilon}_i(t)\le \overline{\dot{\upsilon}_i},\ \forall_i=1,...,n\right\} \\ \upsilon_R &= \left\{\upsilon(t)\in\Re^n : \underline{\upsilon_i} \ \le\ \upsilon_i(t)\le \overline{\upsilon_i},\ \forall_i=1,...,n\right\} \end{aligned} \tag{7}$$

Their motion control is a reference tracking problem. Therefore, system error is given by.

$$\begin{aligned} &e(t)=r(t)-y(t), \\ &\dot{e}(t)=\dot{r}(t)-\dot{y}(t)=\dot{r}(t)-C\dot{x}(t) \end{aligned}$$

Accordingly, the extended system is as follows.

$$\underbrace{\begin{bmatrix} E(\upsilon) & I \end{bmatrix}}_{\hat{E}}\begin{bmatrix}\ddot{x}(t)\\ \dot{e}(t)\end{bmatrix}=\underbrace{\begin{bmatrix} A & B\\ C & 0\end{bmatrix}}_{\hat{A}}\underbrace{\begin{bmatrix}\dot{x}(t)\\ e(t)\end{bmatrix}}_{\hat{x}}+\underbrace{\begin{bmatrix}B\\ 0\end{bmatrix}}_{\hat{B}}\dot{u}(t), \tag{8}$$

$$\hat{x}(0)=\hat{x}_0$$

So, LQI controller including output feedback is as follows.

$$K_{LQI}(\upsilon(t))\hat{x}(t)=K_1(\upsilon(t))x(t)+K_2(\upsilon(t))\int_0^t y(t).$$

Accordingly, the following theorem presents the proposed gain scheduling LQI controller design.

Theorem: Consider a non-singular descriptor LPV system in **(6)** and (8), an optimal controller input $K_{LQI}(\upsilon(t))\hat{x}(t)=K_1(\upsilon(t))x(t)+K_2(\upsilon(t))\int_0^t y(t)$ and controller matrix in (10) which minimizes the performance index J in (5), if and only if there exist parameter-dependent symmetric positive-definite matrices $Y\in\Re^{n\times n}$, $Z(\upsilon)\in\Re^{q\times q}$ and block diagonal positive-definite matrix $X\in\Re^{n\times n}$, the inequalities in (9) hold all for all $\upsilon(t)\in\upsilon_R\times\upsilon_D$.

$\min\gamma$, such that

$$\begin{bmatrix} \hat{A}Y\hat{E}^T(\upsilon)+(*)^T-\hat{B}Z(\upsilon)\hat{B}^T & \hat{E}(\upsilon)Y \\ Y\hat{E}^T(\upsilon) & X\end{bmatrix}<0 \tag{9}$$

$$\begin{bmatrix}-\gamma & \hat{x}_0^T\\ \hat{x}_0 & -Y\end{bmatrix}<0$$

$$K_{LQI}(\upsilon)=Z(\upsilon)\hat{B}^TY^{-1} \tag{10}$$

Proof: The non-singular descriptor system in (8) is rewritten as follows.

$$\dot{\hat{x}}(t)=\hat{E}(v)^{-1}\hat{A}\hat{x}(t)+\hat{E}(v)^{-1}\hat{B}\hat{u}(t),\hat{x}(0)=\hat{x}_0 \tag{11}$$

When Lyapunov stability criteria is applied for asymptotic stability for Lyapunov function $V=\hat{x}^TP\hat{x}$, the following condition should be provided [38]:

If the controller $\hat{u}(t)=-K_{LQI}(\upsilon)\hat{x}(t)$ provides the condition $\frac{dV}{dt}\le 0$, the system is stable and (12) is obtained for the performance index. If $V=\hat{x}^TP\hat{x}$ is replaced, the optimal controller in (13) is obtained.

$$\frac{dV}{dt}+\hat{x}^T(t)Q\hat{x}(t)+\hat{u}^T(t)R(\upsilon)\hat{u}(t)=0 \tag{12}$$

$$\hat{u}(t)=-R^{-1}(\upsilon)\hat{B}^TP\hat{x}(t)=-K(\upsilon)\hat{x}(t) \tag{13}$$

In that case, algebraic Riccati matrix equation in (14) is yielded.

$$\begin{aligned} &\hat{A}^T\hat{E}^{-T}(\upsilon)P+P\hat{E}^{-1}(\upsilon)\hat{A} \\ &-P\hat{E}^{-1}(\upsilon)\hat{B}R^{-1}(\upsilon)\hat{B}^T\hat{E}^{-T}(\upsilon)P+Q=0 \end{aligned} \tag{14}$$

The matrix equation is converted to Linear matrix inequality (LMI) because LMI approach is proposed for the design. Because, it is commonly used for the controller design since LMI approach has many advantages.

For minimizing the cost function, the following equality is obtained. The matrix equation in (14) together with initial condition can be expressed

$$\begin{aligned} J_{\min}(t)&=\min\int_0^\infty\left[\hat{x}^T(t)Q\hat{x}(t)+\hat{u}^T(t)R(\upsilon)\hat{u}(t)\right]dt \\ &=\hat{x}_0^T(t)P\hat{x}_0(t) \end{aligned} \tag{15}$$

Accordingly, following Schur formula in [38] is commonly used in linear algebra applications:

Consider a symmetric matrix $T=\begin{bmatrix}T_{11} & T_{12}\\ T_{12}^T & T_{22}\end{bmatrix}$, $T<0$, if and only if $T_{22}<0$ and $T_{11}-T_{12}T_{22}^{-1}T_{12}^T<0$.

Thus, if (14) is multiplied with $P^{-1}=Y$, LMI form in (16) can be written as follows. And then, if Schur formula is applied, the inequality in (17) is obtained.

$$\begin{aligned} &\hat{E}^{-1}(\upsilon)\hat{A}Y+Y\hat{A}^T\hat{E}^{-T}(\upsilon) \\ &-\hat{E}^{-1}(\upsilon)\hat{B}R^{-1}(\upsilon)\hat{B}^T\hat{E}^{-T}(\upsilon)+YQY<0 \end{aligned} \tag{16}$$

$$\begin{bmatrix}\begin{pmatrix}\hat{E}^{-1}(\upsilon)\hat{A}Y+(*)^T\\ -\hat{E}^{-1}(\upsilon)\hat{B}R^{-1}(\upsilon)\hat{B}^T\hat{E}^{-T}(\upsilon)\end{pmatrix} & Y\\ Y & Q^{-1}\end{bmatrix}<0 \tag{17}$$

The matrix $P\left(P^{-1}=Y\right)$ is minimized to get minimum performance index. Therefore, defining γ, $\gamma>\hat{x}_0^TP\hat{x}_0$ can be obtained. Then, if Schur formula is applied to $\hat{x}_0^TP\hat{x}_0-\gamma<0$, the inequality in (18) is obtained.

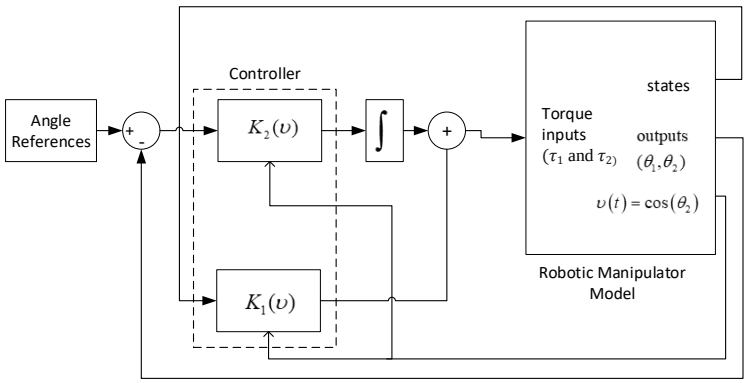


Figure 2. Control system diagram.

$$\begin{bmatrix} -\gamma & \hat{x}_0^T \\ \hat{x}_0 & -Y \end{bmatrix} < 0 \tag{18}$$

The inequality in (17) is multiplied from left and right with the matrix in (19) and its transpose, respectively.

$$\begin{bmatrix} \hat{E}(\upsilon) & 0 \\ 0 & I \end{bmatrix} \tag{19}$$

Thus, we get the following inequality by also defining $X = Q^{-1}$ and $Z(\upsilon) = R^{-1}(\upsilon)$, and the proof completes.

$$\begin{bmatrix} \hat{A}Y\hat{E}^T(\upsilon) + (*)^T - \hat{B}Z(\upsilon)\hat{B}^T & \hat{E}(\upsilon)Y \\ Y\hat{E}^T(\upsilon) & X \end{bmatrix} < 0 \tag{20}$$

Finally, an optimal feedback controller $K_{LQI}(\upsilon(t))\hat{x}(t) = K_1(\upsilon(t))x(t) + K_2(\upsilon(t))\int_0^t y(t)$, which minimizes performance index J and stabilizes the system, can be obtained by solving the optimization problem in (9) for the system in (8) including output feedback.

Remark: The inverse of $E(\upsilon)$ is a drastic case due to rational LPV form and there is no a formula for this, so an extra complexity occurs, and some assumptions or extra methods are needed as in [22, 28, 36, 37]. For example, since $\frac{A(\theta)}{E(\theta)} = E^{-1}(\theta)A(\theta)$ is dependent a parameter theta, the making inverse of the parameter dependent matrix is a difficult problem. But, the proposed solution eliminates to take its inverse. Thus, the proposed design has no the complexity and does not need any extra methods.

4. Simulation results

The simulation tests are performed with MATLAB. Figure 2 shows the simulation block diagram for the two-link flexible manipulator control system. The time-varying parameter is assumed as follows for robotic manipulator system.

$$\upsilon(t) = \cos(\theta_2)$$

Therefore, the bounds of the time-varying parameter is $-1 \le \upsilon(t) \le 1$.

Throughout the paper, affine parametric uncertainties are defined for the controller design, and so all parameter-dependent matrices affinely depend on the uncertain parameter such as the controller matrix in (21).

$$K_{LQI}(\upsilon) = K_0 + \sum_{i=1}^{n} \upsilon_i K_i \tag{21}$$

In addition, the selected matrix X is for the state weights in the optimization problem in (9). The determination of this matrix is not dependent on the certain rule, but it is optimized by some artificial intelligence methods in the literature. In these paper, classical approach (trial and error) is used according to system behaviour.

$$X = \text{blkdiag}\left(10^6 \times (2050, 200, 16, 9, 7, 1.2, 2.3, 3.0, 0.05, 0.03)\right)$$

If the optimization problem is solved, the effects on the system of controller in (21) are presented in Figure 3, Figure 4 and Figure 5. Figure 3 shows joint angles θ_1 and θ_2 of robotic manipulator as degree. Thus, the tracking performance is good. The settling times are nearly 3 and 4 second for θ_1 and c, respectively. Figure 4 shows the time-varying parameter $\upsilon(t)$, and accordingly Figure 5 shows the generated torques by the parameter dependent controller as Nm. the results show that the proposed method has good performance such as low settling time and overshoot. The proposed controller is static feedback which changes with θ_2, and, there is no assumptions while the controller in [29] is dynamic controller, and there are some asumptions such as fixed angles. Nevertheless, the results for the proposed design are somewhat better than the results in [29] especially for θ_2 in point of settling time.

5. Conclusions

In this paper, we have addressed the construction of parameter-dependent LQI controller design for the non-singular LPV descriptor systems, and it has been applied to two-link robotic manipulator. The results have shown that good tracking performances are

obtained for the joint angle references in view of the settling time and overshoot. In addition, the proposed method has not needed any extra requirements such as rational LPV conversion methods. Finally, the gain scheduling LQI controller has been designed for the non-singular descriptor system without any extra methods for the rational LPV form.

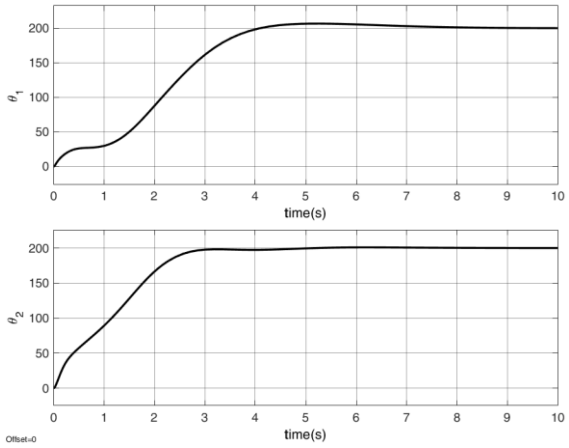


Figure 3. Tracking responses of joint angles.

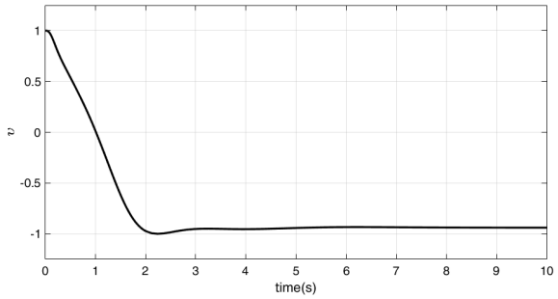


Figure 4. The change of time-varying parameter $\upsilon(t)$.

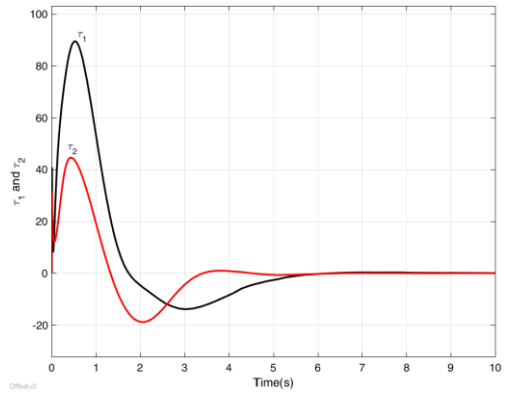


Figure 5. Torque responses.

References

[1] Park BY, Yun SW, Park P (2012) H 2 state-feedback control for LPV systems with input saturation and matched disturbance. *Nonlinear Dynamics* 67(2):1083–1096.

[2] Venkataraman R, Seiler P (2018) Convex LPV synthesis of estimators and feedforwards using duality and integral quadratic constraints. *International Journal of Robust and Nonlinear Control* 28(3):953–975.

[3] He Z, Xie W (2014) Quadratic H2 gain performance analysis of LPV systems with realization of parametric transfer function. *Proceedings of the 33rd Chinese Control Conference, CCC 2014* (IEEE), pp 1788–1792.

[4] Peng C, Zhang Z, Zou J, Li K, Zhang J (2013) Internal model based robust inversion feedforward and feedback 2DOF control for LPV system with disturbance. *Journal of Process Control* 23(10):1415–1425.

[5] Ku C-C, Chen G-W (2015) Gain-scheduled controller design for discrete-time Linear Parameter Varying systems with multiplicative noises. *International Journal of Control, Automation and Systems* 13(6):1382–1390.

[6] Wu K, Zhang Q, Hansen A (2004) Modelling and identification of a hydrostatic transmission hardware-in-the-loop simulator. *Int J Veh Des* 34(1):63–75.

[7] Lu B, Choi H, Buckner GD, Tammi K (2008) Linear parameter-varying techniques for control of a magnetic bearing system. *Control Engineering Practice* 16(10):1161–1172.

[8] Altun Y (2017) A new extended LMI-based robust gain scheduled state feedback H 2 controller design. *International Journal of Control, Automation and Systems* 15(3):967–975.

[9] Masubuchi I, Kurata I (2011) Gain-scheduled control via filtered scheduling parameters. *Automatica* 47(8):1821–1826.

[10] Abdullah AA (2018) Robust model reference control of linear parameter-varying systems with disturbances. *IET Control Theory & Applications* 12(1):45–52.

[11] Apkarian P, Adams RJ (1998) Advanced gain-scheduling techniques for uncertain systems. *IEEE Transactions on Control Systems Technology* 6(1):21–32.

[12] Huang Y, Na J, Wu X, Gao G-B, Guo Y (2016) Robust adaptive control for vehicle active suspension systems with uncertain dynamics. *Transactions of the Institute of Measurement and Control* 40(4):1237–1249.

[13] Altun Y (2017) The road disturbance attenuation for quarter car active suspension system via a new static two-degree-of-freedom design. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)* 7(2):142–148.

[14] Yazici H, Sever M, Sever M (2017) Design of an optimal state derivative feedback LQR controller and its application to an offshore steel jacket platform. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)* 8(1):84–91.

[15] Lin C, Wang QG, Lee TH (2005) Robust normalization and stabilization of uncertain descriptor systems with norm-bounded perturbations. *IEEE Transactions on Automatic Control* 50(4):515–520.

[16] Fang C-HFC-H (2002) Stability robustness analysis of uncertain descriptor systems - an LMI approach. *Proceedings of the 41st IEEE Conference on Decision and Control, 2002* 2(December):1459–1460.

[17] Lin C, Wang J, Wang D, Soh CB (1997) Robustness of uncertain descriptor systems. *Systems & Control Letters* 31(3):129–138.

[18] Chou JH, Chen SH, Zhang QL (2006) Robust controllability for linear uncertain descriptor systems. *Linear Algebra and Its Applications* 414(2–3):632–651.

[19] Osorio-Gordillo GL, Darouach M, Astorga-Zaragoza CM (2015) H∞ dynamical observers design for linear descriptor systems. Application to state and unknown input estimation. *European Journal of Control* 26:35–43.

[20] Bara GI (2011) Dilated LMI conditions for time-varying polytopic descriptor systems: The discrete-time case. *International Journal of Control* 84(6):1010–1023.

[21] Masubuchi I, Suzuki A (2008) Gain-Scheduled Controller Synthesis Based on New LMIs for Dissipativity of Descriptor LPV Systems. *IFAC Proceedings Volumes* 41(2):9993–9998.

[22] Bouali A, Yagoubi M, Chevrel P (2008) H2 H∞ gain scheduling control for rational LPV systems using the descriptor framework. *IEEE Conference on Decision and Control (CDC)* (IEEE), pp 3878–3883.

[23] Rodrigues M, Hamdi H, Benhadj Braiek N, Theilliol D (2014) Observer-based fault tolerant control design for a class of LPV descriptor systems. *Journal of the Franklin Institute* 351(6):3104–3125.

[24] Ali HS, Darouach M, Zasadzinski M, Alma M (2015) An H∞ LPV control for a class of LPV systems using a descriptor approach: Application to a wind turbine mode. *IFAC-PapersOnLine* 48(26):213–217.

[25] Bara GI (2011) Robust analysis and control of parameter-dependent uncertain descriptor systems. *Systems and Control Letters* 60(5):356–364.

[26] Altun Y (2017) Control Design for Nonsingular Descriptor Systems. *Karaelmas Science and Engineering Journal* 7(1):154–159.

[27] Bouali A, Yagoubi M, Chevrel P (2008) Gain scheduled observer state feedback controller for rational LPV systems. *IFAC Proceedings Volumes* 41(2):4922–4927.

[28] Adams R, Apkarian P, Chretien J (1996) Robust Control Approaches for a Two-Link Flexible Manipulator. *3rd International Conference on Dynamics and Control of Structures in Space*, p 101-116.

[29] Karkoub M, Tamma K, Balas G (1999) Robust Control of Two-Link Flexible Manipulators Using the -Synthesis Technique. *Journal of Vibration and Control* 5(4):559–576.

[30] Karagülle H, Malgaca L, Dirilmiş M, Akdağ M, Yavuz Ş (2015) Vibration control of a two-link flexible manipulator. *Journal of Vibration and Control* 23(12):2023–2034.

[31] Li Y, Liu G, Hong T, Liu K (2005) Robust control of a two-link flexible manipulator with quasi-static deflection compensation using neural networks. *Journal of Intelligent and Robotic Systems: Theory and Applications* 44(3):263–276.

[32] Halalchi H, Laroche E, Bara GI (2010) LPV Modeling and Control of a 2-DOF Robotic. *11th Pan-American Congress of Applied Mechanics*.

[33] Yadav PS, Singh N (2015) Robust Control of Two Link Rigid Manipulator. *International Journal of Information and Electronics Engineering* 5(3):3–8.

[34] Halalchi H, Laroche E, Iuliana Bara G (2014) Flexible-Link Robot Control Using a Linear Parameter Varying Systems Methodology. *International Journal of Advanced Robotic Systems* 11(3):46.

[35] Polat I, Eskinat E, Kose IE (2007) Dynamic output feedback control of quasi-LPV mechanical systems. *IET Control Theory and Applications* 1(4):1114–1121.

[36] Bouali A, Yagoubi M, Chevrel P (2008) H2 gain scheduled observer based controllers for rational LPV systems. *2008 10th International Conference on Control, Automation, Robotics and Vision, ICARCV 2008* 41(2):1811–1816.

[37] Galvão RKH, Kienitz KH, Hadjiloucas S (2017) Conversion of descriptor representations to state-space form: an extension of the shuffle algorithm. *International Journal of Control*:1–15.

[38] Boyd S, El Ghaoui L, Feron E, Balakrishnan V (1994) *Linear matrix inequalities in system and control theory*, Society for Industrial and Applied Mathematics, Philadelphia.

Yusuf Altun *received the Doctoral degree in Electrical Engineering from Yildiz Technical University, Istanbul, Turkey in 2012. He has worked as Assistant Professor at the Duzce University, in Department of Computer Engineering, Duzce, in Turkey since 2013. His main research areas are the control of mechatronic systems such as vehicles, robot, and electromechanical systems.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.2, pp.208-215 (2018)
https://doi.org/10.11121/ijocta.01.2018.00557

RESEARCH ARTICLE

Distance restricted maximal covering model for pharmacy duty scheduling problem

Nuşin Uncu[a *], Berna Bulğurcu[b], Fatih Kılıç[c]

[a] *Department of Industrial Engineering, Adana Science and Technology University, Turkey*
[b] *Deparment of Business, Çukurova University, Turkey*
[c] *Department of Computer Engineering, Adana Science and Technology University, Turkey*
nuncu@adanabtu.edu.tr, bkiran@cu.edu.tr, fkilic@adanabtu.edu.tr

ARTICLE INFO

Article history:
Received: 27 October 2017
Accepted: 4 April 2018
Available Online: 1 June 2018

Keywords:
Pharmacy
Duty scheduling
Location models
Maximal covering
Distance restricted

AMS Classification 2010:
90C11, 90C90

ABSTRACT

Pharmacies are considered as an integral part of health care systems for supplying medicine to patients. In order to access medicine with ease, pharmacies locations in the context of distance and demand are important for patients. In the case of a few numbers of pharmacies may be on duty at nights or during holidays, pharmacies duty scheduling problem occur and can be associated with location models. In contrast to widely used *p*-median model which aims to minimize the demand-weighted distance, we maximize the demand covered over the distance between the patients and the pharmacies on duty. Main contribution of the proposed model is the restriction constraint for the distance between pharmacies on duty in order to ensure fairness in an organizational view of point. We propose a distance restricted maximal covering location model (DR-MCLM) in this study. This mathematical model is a mixed integer linear programming model and solved by Lingo optimization software. The distances between the pharmacies and the sites are obtained using Geographic Information Systems (GIS). The model is applied for the case in Adana, one of the biggest cities in Turkey. The results are given on the maps of the city, including the pharmacies on duty and their assignments to sites in each day of the period.

1. Introduction

Pharmacies supply medicine and medical instruments to patients. Location of pharmacies is important for patients and usually depends on population distribution over city and closeness to hospitals. Patients and their relatives need them to access medicine in the daytime and night period. In Turkey, it is forbidden to sell medicine in anywhere other than in pharmacies. Therefore, people can get medicine only from the pharmacies. This results in the need of pharmacies to be open after the daytime period and meet the urgent demands. However, not all pharmacies are allowed to work at nights. The problem of which pharmacies should be kept open and which ones closed at nights is the main concern of the duty scheduling problem of pharmacies. The Chamber of Pharmacists is responsible for pharmacies duty scheduling as considering fairness toward total number of duties given to each pharmacy. This paper aims to develope a mathematical model for the fairest optimal duty scheduling of pharmacies.

Location theory includes such models representing the scheduling and assignment problems. Set covering model is the well known and base one in the location theory. It is used to find the minimum number of facilities to cover all the demand points in a given region. Other important one is the maximal covering location model (MCLM) and it seeks to locate facilities in order to maximize the coverage of all the demand points. The remained covering models can be classified as *p*-center and *p*-median models. *P*-center model seeks to find the location of *p* facilities to minimize the maximum distance between each facility and the demand point in the given region. The quantity of demand covered is not considered in a *p*-center model, whereas in a *p*-median model demand weighted distance minimization is taken into account. In the case of health care problems, while distance has a significant role in saving life, quantity of demand is also the other critical factor for assigning a facility to a site.

*Corresponding author

In this study, MCLM formulated in [1] is revised for the distance restricted pharmacy duty scheduling problem. In the suggested model, in addition to the demand and the distance between pharmacies and sites, distance restricted between pharmacies is also considered.

In this paper, the introduction part is followed by a literature review with four other sections. The review provides an overview of the existing literature on the pharmacy duty scheduling problem and the maximal covering location models formulated mainly for the aforementioned problem. The third section suggests a new mathematical model constructed based on the new constraint that restricts the distance between pharmacies. In the fourth section, an application of the model is provided. The results are mapped by using geographic information system tools in the following section. Finally, in the last section we draw some conclusions and make further suggestions.

2. Literature review

The duty scheduling problem has been studied in a widespread area of operations research. However, it is not the same for pharmacy duty scheduling problem. In the first part of the review, the studies on the duty scheduling of pharmacies is demonstrated. In the second part, we present the related literature about the maximal covering location model which forms the mathematical basis for the suggested pharmacy duty scheduling problem in this study.

Variable neighbourhood search (VNS) is used to solve a multi period *p*-median problem with the real data obtained from pharmacies in İzmir with particular constraints [2]. The goal of basic, restricted and decomposition version of the variable neighbourhood search method is to find better solutions in the least amount of time by decomposing or restricting the search space. According to the findings, the basic and restricted versions of the method give better results significantly than the decomposition version of small and large scale instances. The customer utility maximization is a focal point in their research. The pharmacy duty scheduling problem is examined by adding special side constraints to a multi-period *p*-median model in [3]. Tabu search algorithm is used to solve the problem for the case in Izmir. They provide that the multiple duty pharmacy scheduling problem is *NP*-hard under the assumption of incapacitated pharmacies. The last study is on the location of pharmacies is belongs to [4]. They use set covering, *p*-center and *p*-median models with geographic information systems in order to select the most suitable triad of pharmacies –hospitals-warehouses to minimize the total transportation distance in the province of Gaziantep. They developed a GIS application to visualize the current and potential locations of the pharmacies.

Facility location is known as the most important starting point in the history of location theory [5]. The problems in this field refer to the several models each of which indicates different objectives and has different constraints. The applications of facility location spread to many areas such as emergency services, hospitals, fire stations, airline hubs, schools and warehouses among many others. In health care applications, the set covering model, maximal covering model and *p*-median model are taken attention more than the others [6].

The roots of MCLM were generalized in [1] and various extensions to the original problem have been made since then. A comprehensive literature review about the recent papers in the set covering model, maximal covering model and *p*-median model can be found in [7].

The basic maximal covering model aims to cover maximum demand with a fixed number of facilities. An extension of the model is proposed in [8]. They emphasize on the full and partial coverage of the demand nodes and lower level of service on the constraint of distance between the facility and the demand node. They mentioned on the behaviour of maximal covering model in which some demand points cannot be covered.

The model in [9] is a dynamic maximal covering model. They called the model dynamic because of the changing locations of the facilities in time steps in the whole period. It is the similar approach as considering opening some facilities in time steps in the period, while the others are closed.

Another extension of MCLM is formulated for different capacitated facilities. Several possible capacity levels of the facility at each potential site are used in contrast to only one fixed capacity level which limited the applications of a modular capacitated maximal covering location problem [10]. The article provided a discussion of modular capacitated maximal covering location problem with either facility constraint or non-facility constraint to determine the optimal location of ambulances.

Although the large scale MCLMs are difficult to solve, heuristics are useful to solve such kind of problems. The large-scale maximal covering problem is solved by genetic algorithms with up to 900 nodes in [11]. The results show that the proposed approach has a capacity to solve problems with up to 2500 nodes. A combination of fuzzy simulation and the greedy variable neighbourhood search is presented to solve large scale maximal coverage location problems in [12]. They use greedy VNS for searching solution space and fuzzy simulation to evaluate each solution obtained by greedy VNS. They solve the problem with 1000, 1500, and 2500 nodes and compare the results with the exact results obtained using CPLEX. MCLM maximize the coverage of population within a rapid nuclear response while minimizing modifications of the existing system [13]. They consider three phases including facility location, reallocation and coverage. At these phases, their first aim is to maximize the use

of the existing system. The second is to determine the trade-offs between competing objectives, and the final one is to reduce the set of alternatives. After adding modified conditional covering problem constraints to the problem, they prevent facilities from covering nodes within close proximity in a worst-case scenario. Another variation of the maximal location coverage model is suggested in [14]. The location of facilities and allocation of customers are two main decision criteria of the problem which are considered in bi-level mathematical formulation. They propose greedy randomized adaptive search procedure (GRASP) and hybrid GRASP-Tabu search heuristics. The results of single level and bi-level formulation are compared based on convergence to optimality. It is determined that two heuristics can find good quality solutions.

As a contribution to the existing literature, besides the known objectives and constraints, the model is improved by adding distance restricted constraint for the special case of fairness between facilities in this study.

3. Mathematical model

A graph consists of a set of nodes and edges that refer to the location points and the lines connecting these location points. In this problem, the nodes represent demand points&location of pharmacies and the edges represent distances between pharmacies and demand points. Nodes indexed by *i=1….I*, indicate the demand (site) points that are to be covered by pharmacies. Nodes indexed by *j = 1… J* indicate pharmacies that are candidates for covering the demand points on a given time period, *T*. The proposed model is an extension of MCLM. For the same reason that MCLM is *NP*- hard, our suggested model is also an *NP*-hard problem with three dimensional decision variables indexed respectively, *i, j,* and *t*. The model in this study is a mixed integer linear model and solved by using Lingo software which includes Simplex, branch and bound, and nonlinear solvers. We obtain feasible solution of the model, thanks to its small size.

Distance restricted maximal covering location model (DR-MCLM) for pharmacy duty scheduling (PDS) problem:

In this model, the distance between pharmacies on duty is restricted for each day of the period. The model tends to maximize the demands per distance covered by *p* pharmacies in a given period *T*.

Notations

Sets and Indices

$i \in I$: set of demand (site) points

$j, m \in J$: set of facilities (pharmacies),

$t \in T$: time periods (days),

$D = \{m, j \in J \mid s_{mj} < DR\}$

Parameters

h_i: demand at node *i*.

d_{ij}: distance between demand point *i* and pharmacy *j*.

s_{mj}: distance between pharmacy site *m* and pharmacy *j*.

n: number of pharmacies on duty in one period

DR: maximum distance restricted between pharmacies on duty

M: a big number

$$y_{jt} = \begin{cases} 1, & \text{if pharmacy } j \text{ is on duty on day } t \\ 0, & \text{otherwise} \end{cases}$$

$$x_{ijt} = \begin{cases} 1, & \text{if demand point } i \text{ is covered by pharmacy } j \text{ on day } t \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Maximize } Z = \sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{i=1}^{I} x_{ijt} h_i / d_{ij} \quad (1)$$

$$x_{ijt} \le y_{jt} \qquad \forall i, j, t \quad (2)$$

$$\sum_{t=1}^{T} y_{jt} = 1 \qquad \forall j \quad (3)$$

$$\sum_{j=1}^{J} x_{ijt} \le 1 \qquad \forall i, t \quad (4)$$

$$\sum_{j=1}^{J} y_{jt} \ge n \qquad \forall t \quad (5)$$

$$y_{mt} \le (1 - y_{jt})M \qquad m \ne j, \forall D \quad (6)$$

$$y_{jt} \in \{0,1\} \qquad \forall j, t \quad (7)$$

$$x_{ijt} \in \{0,1\} \qquad \forall i, j, t \quad (8)$$

The objective function of the model represented by Eq. (1) aims to maximize the demand covered over the distance between site and pharmacies on duty for each day of period *T*. Eq. (2) prevents a site to be covered by a pharmacy which is not open on day *t*. Each pharmacy should be on duty once in the whole period is given in Eq. (3). Eq. (4) states that each site should not be covered more than once (no backup coverage) on each day. Eq. (5) ensures that at least *n* number of the pharmacies must be on duty on each day. Eq. (6) is fairness constraint. It states that the distance between pharmacies on duty at the same day must not be less than the maximum distance restricted. As shown in the notations, the set *D* is given including the index numbers of inconvenient pairs of pharmacies which can not have duty on the same day. Eq.(7) and Eq.(8) stand for the binary values of decision variables.

4. Adana case

In Adana case, the requirement of Adana Chamber of Pharmacists is to ensure fairness between pharmacies. In order to meet this fairness, Adana Chamber of Pharmacists emphasized that the distance between pharmacies on duty must not be less than 1.6 km. It is mentioned by Adana Chamber of Pharmacists that the

given pharmacies on the map are eligible to be on duty in terms of capacity and other requirements.

The model is run with data which were recorded in 2017 by Adana Chamber of Pharmacists. These data include the number of pharmacies and their locations over Çukurova. Çukurova is one of the biggest region in Adana and 114 pharmacies are located in this region. The Chamber of Pharmacists requires that the number of duty pharmacies must be at least 4 for this region in each day of the period. Therefore, the period include 28 days for this special situation.

On the map of Adana, 10 demand points with blue border lines are shown by their numbers in blue circles and the locations of pharmacies are given with their numbers in red symbols in Figure 1. One of the pharmacies with number 26 on the left side of the figure is located in the rural area of the city and it is also included in the model.

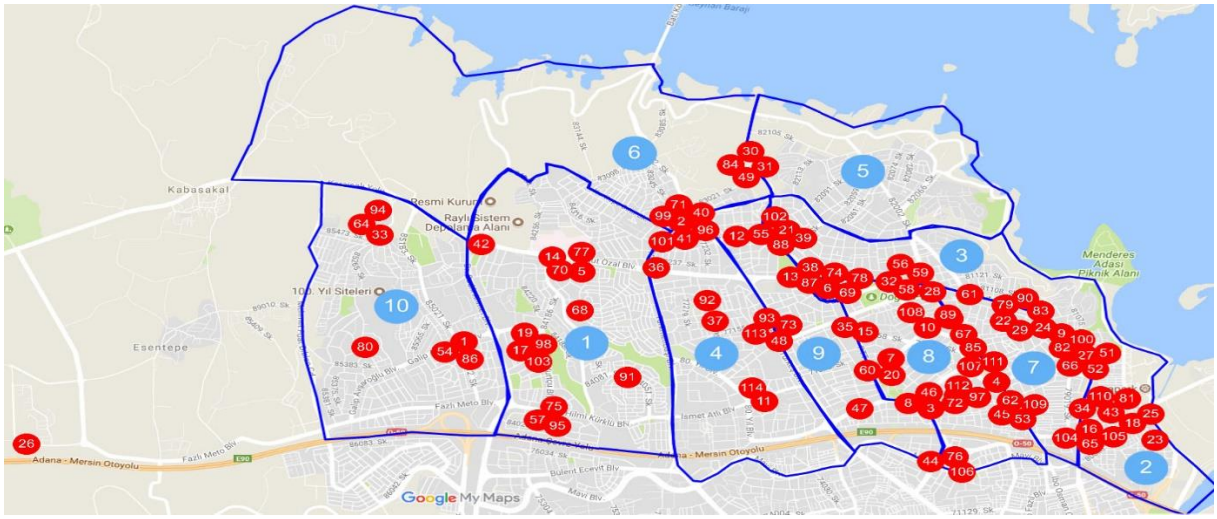
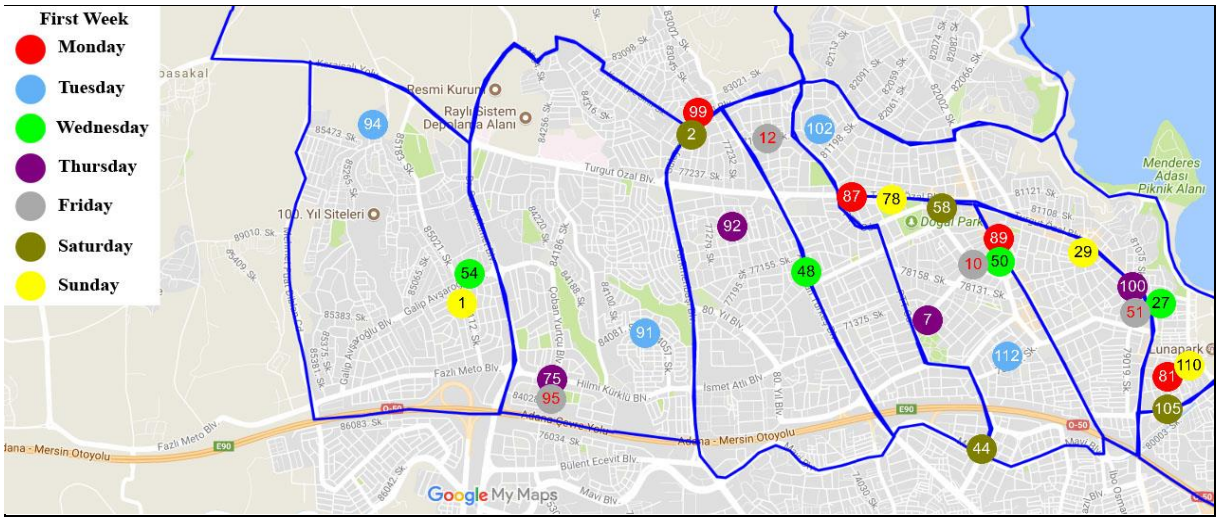


Figure 1. Sites and pharmacies located in the region of Çukurova.

5. Results

The distance restricted maximal covering location model is solved by using Lingo software with an Intel Core i7 2.4 GHz computer. The optimum result is reached after 24 minutes. The solution of DR-MCLM is difficult to be exhibited in one graph, because of its dimensions. Thus, the results are given in timely based graphs and four maps are used for the period of 28 days. Each graph shows a week and each day is shown by one colour. For each day, the points of locations of duty pharmacies are given within the same colour in Figure 2. It can be concluded that the model tends to choose the locations of pharmacies on duty not very close to each other (min 1.6 km) while covering all the sites and maximizing the demand covered over distance.

The results show that the model finds the locations of duty pharmacies on each day of the period with the given constraints successfully. In Figure 2, each group of same coloured balls with the numbers of pharmacies are given for the same day of the period on the region.



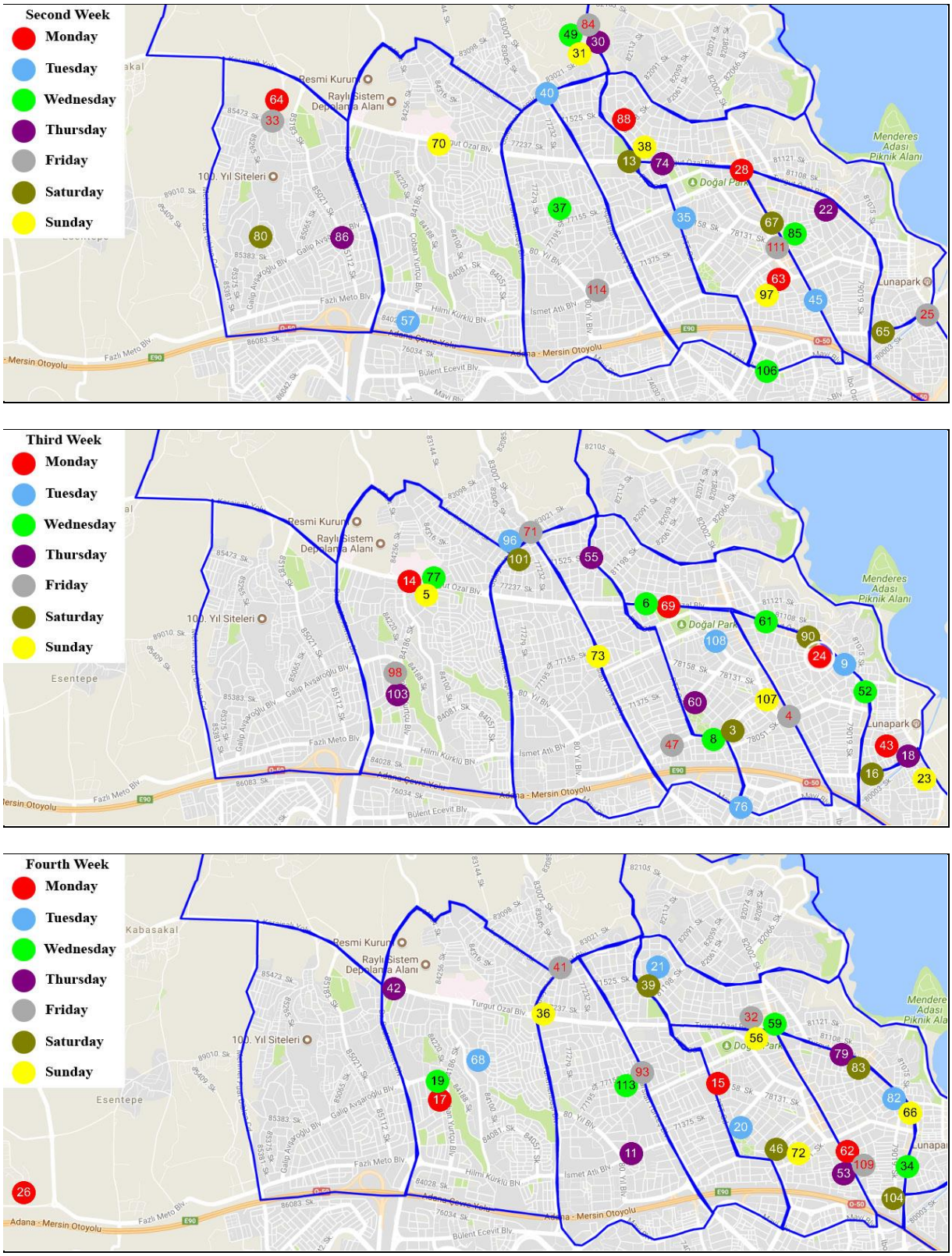


Figure 2. Duty pharmacies on each week in 28 days period.

The objective function value is attained approximately 1.64E+8. This value is the maximum covered population over distances between pharmacies and the sites. This value is not a good indicator alone, however maximizing this value forces the model to assign the pharmacy to the nearest site with higher population. The assignments of each site to duty pharmacy are demonstrated in Table 1. The results can be analysed together with the Figures on the previous pages. All the sites are covered in each day of the period and the results seem fair.

Table 1. Weekly representation of covered sites by duty pharmacies.

1st Week		*2nd Week*		*3th Week*		*4th Week*	
Pharmacy No	**Covered Sites**	**Pharmacy No**	**Covered Sites**	**Pharmacy No**	**Covered Sites**	**Pharmacy No**	**Covered Sites**
81	2,7	28	3,9	14	6,10	15	1,3,4,5,9
87	1,4,5	63	2,4,7,8	24	3,8	17	10
89	3,8,9	64	10	43	2,7	26	6
99	6,10	88	1,5,6	69	1,4,5,9	62	2,7,8
91	1,10	35	1,3,4,5,9	9	7	20	3,4,8,9
94	6	40	6	76	2	21	5,6
102	5	45	2,7,8	96	1,6,10	68	1,10
112	2,3,4,7,8,9	57	10	108	3,4,5,8,9	82	2,7
27	2,7	37	1,4,10	6	1,5	19	6,10
48	1,4,5,6	49	5,6	8	4,8,9	34	2,7,8
50	3,8,9	85	3,7,8,9	52	2,7	59	3,5,9
54	10	106	2	61	3	113	1,4
				77	6,10		
7	3,4,8,9	22	2,3,7,8	18	2,7	11	1,4
75	10	30	6	55	1,5,6	42	6,10
92	1,5,6	74	1,4,5,9	60	3,4,8,9	53	2,7,8
100	2,7	86	10	103	10	79	3,5,9
10	3,4,8,9	25	2	4	2,3,7,8	32	3,5,9
12	1,5,6	33	10	47	1,4,9	41	6,10
51	2,7	84	5,6	71	5,6	93	1,4
95	10	111	3,7,8,9	98	10	109	2,7,8
		114	1,4				
2	1,6,10	13	1,4,5,6	3	4,8,9	39	1,5,6,10
44	8	65	2,7	16	2,7	46	4,8,9
58	3,4,5,9	67	3,8,9	90	3	83	3
105	2,7	80	10	101	1,5,6,10	104	2,7
1	10	31	6	5	6,10	36	1,6,10
29	3,8	38	1,5	23	2,5	56	3,5,9
78	1,4,5,6,9	70	10	73	1,4	66	2,7
110	2,7	97	2,3,4,7,8,9	107	3,7,8,9	72	4,8

6. Conclusion

A distance restricted maximal covering model for the problem of pharmacies duty scheduling is provided in this study. The special case of the problem is the distance restriction that has to be considered between pharmacies on duty. For the classical MCLM, the demand point is assumed to be covered completely if

located within the critical distance of the facility and not covered at all outside of the critical distance. Most of the models in location science consider the distances between facility and the sites. Comparing with classical MCLM, the contribution to the model lies in the critical distance between the facilities taken into account in this study. We emphasized on this critical distance between pharmacies and fortunately this distance does not occur as an infeasibility for our case. The constraints for distances may result an infeasibility, because of the reason that the real system may not been established under this constraint at the beginning. The feasibility may not be reachable case to case depend on the given data. Therefore, the model proposed in this study need to be solved by heuristics or relaxation methods not by exact methods for some data related to large systems, if required. Thanks to the size of the problem and appropriate data in our case, the optimal duty schedules and assignments of demand points are obtained by using Lingo software. The results are found reasonable by the Chamber of Pharmacists based on its applicability, but they mentioned that the patients cannot be restricted to get medicine from an assigned pharmacy obtained by our results, as it is also obvious for us. On the other hand, the assignment of demand points to pharmacies allows to balance the demand over the facilities while scheduling the duties. For further studies, the suggested model could be extended to the whole city in the same case in order to compare the whole schedule in practice. Therefore, the performance of proposed model can be investigated up to a large number of sites and pharmacies.

Acknowledgements

The authors would like to thank Adana Chamber of Pharmacists for providing data used in this manuscript and the editors of this journal and anonomys reviewers for their valuable comments.

References

[1] Church, R., & ReVelle, C. (1974). Maximal covering location problem. *Papers of the Regional Science Association*. 32, 101-118.

[2] Kocatürk, F., & Özpeynirci, Ö. (2014). Variable neighborhood search for the pharmacy duty scheduling problem. *Computers & Operations Research*. 51, 218-226.

[3] Kocatürk, F., & Ağlamaz, E. (2016). Pharmacy duty scheduling problem. *International Transactions in Operational Research.* 23, 459-480.

[4] Özceylan, E., Uslu, A., Erbaş, M., Çetinkaya, C., & İşleyen, S.K. (2017). Optimizing the location-allocation problem of pharmacy warehouses: A case study in Gaziantep. *An International Journal of Optimization and Control: Theories & Applications.* 7(1), 117-129.

[5] Farahani, R.Z., SteadieSeifi, M. & Asgari, N. (2010). Multiple criteria facility location problems: A survey. *Applied Mathematical Modelling*. 34, 1689–1709.

[6] Daskin, M. S., & Dean, L. K. (2004*). Location of Health Care Facilities. Chapter 3 in the Handbook of OR/MS in Health Care: A Handbook of Methods and Applications.* Kluwer.

[7] ReVelle, C.S., Eiselt, H.A., & Daskin, M.S. (2008). A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*, 184, 817–848.

[8] Karasakal, O., & Karasakal, E.K. (2004). A maximal covering location model in the presence of partial coverage. *Computers & Operations Research*, 31, 1515-1526.

[9] Zarandi, M.H.F., Davari, S., & Sisakht, S.A.H. (2013). The large-scale dynamic maximal covering location problem. *Mathematical and Computer Modelling*, 57, 710-719.

[10] Yin, P., & Mu, L., (2012). Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles. *Applied Geography*, 34, 247-254.

[11] Zarandi ,M.H.F., Davari, S., & Sisakht, S.A.H. (2011). The large scale maximal covering location problem. *Scientia Iranica*, 18(6), 1564-1570.

[12] Davari, S., Zarandi ,M.H.F., & Turksen, I.B. (2013). A greedy variable neighborhood search heuristic for the maximal covering location problem with fuzzy coverage radii. *Knowledge-Based Systems*, 41, 68-76.

[13] Paul, N.R., Lunday, B.J., & Nurre, S.G. (2017). A multiobjective, maximal conditional covering location problem applied to the relocation of hierarchical emergency response facilities. *Omega*, 66, 147-158.

[14] Diaz, J.A., Luna, D.E, Vallejo, J-F.C., & Ramirez, M-S.C. (2017). GRASP and hybrid GRASP-Tabu heuristics to solve a maximal covering location problem with customer preference ordering. *Expert Systems with Applications*, 82, 62-76.

Nuşin Uncu *received her PhD from Department of Industrial Engineering in Çukurova University. She is working as an Assistant Professor in Adana Science and Technology University. She received a scholarship granted by TUBITAK for postdoctoral studies on process mining in Eindhoven University of Technology in 2017. Her research area includes modelling, analyzing and mining of complex systems and processes.*

Berna Bulğurcu *received her PhD from Department of Business in Çukurova University. She is working as an Assistant Professor in Çukurova University. Her research interests are multiobjective optimization methods and quantitative decision methods.*

Fatih Kılıç *received his PhD from Department of Electrical&Electronics Engineering in Çukurova University. He is working as an Assistant Professor in Adana Science and Technology University. His research area includes public transit network and search algorithms.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.2, pp.216-227 (2018)
https://doi.org/10.11121/ijocta.01.2018.00516

RESEARCH ARTICLE

A pairwise output coding method for multi-class EEG classification of a self-induced BCI

Nurhan Gursel Ozmen*, Levent Gumusel

Department of Mechanical Engineering, Karadeniz Technical University, Turkey
gnurhan@ktu.edu.tr, gumusel@ktu.edu.tr

ARTICLE INFO

Article history:
Received: 20 July 2017
Accepted: 13 March 2018
Available Online: 14 July 2018

Keywords:
Multi-class EEG classification
Channel reduction
Optimized Output
Brain computer interface (BCI)
Feature extraction

AMS Classification 2010:
62H30

ABSTRACT

In brain computer interface (BCI) research, electroencephalography (EEG) is the most widely used method due to its noninvasiveness, high temporal resolution and portability. Most of the EEG-based BCI studies are aimed at developing methodologies for signal processing, feature extraction and classification. In this study, an experimental EEG study was carried out with six subjects performing imagery mental and motor tasks. We present a multi-class EEG decoding with a novel pairwise output coding method of EEGs to improve the performance of self-induced BCI systems. This method involves an augmented one-versus-one multiclass classification with less time and reduced number of electrodes. Furthermore, a train repetition number is introduced in the training step to optimize the data selection. The difference among right and left hemispheres is also searched. Finally, the difference between experienced and novice subjects is also observed.

The experimental results have demonstrated that, the use of proposed classification algorithm produces high classification accuracies (98%) with nine channels. Reduced numbers of channels (four channels) have 100% accuracies for mental tasks and 87% accuracies for motor tasks with Support Vector Machines (SVM). The classification accuracies are quite high though the proposed one-versus-one technique worked well compared to the classical method. The results would be promising for a real-time study.

1. Introduction

A brain-computer interface (BCI) is a communication system that translates the brain signals into commands for communication or for controlling external devices without requiring any peripheral muscular activity [1-3]. Electroencephalogram (EEG) is the most efficient and widely used recording modality in BCIs due to its non-invasive measurement procedure, portability and reasonable cost [2]. Due to the large numbers of methodologies developed for signal processing, feature extraction and classification of EEG data, there are no gold standards on data processing and machine learning algorithms [4-6]. The output of a BCI contains the decoding of the intended task and then it is transferred to the related device. The discrimination of the tasks is done with a classifier. A number of linear and nonlinear classifiers have been studied for classification of EEG signals under different conditions like Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), Neural Networks (NN) and its special implementations, Bayes Quadratic, Common Spatial Patterns (CSP), Hidden Markov Models and hybrid classifiers [6-10].

It is concluded from the literature [2-4] that, the EEG data has a non-Gaussian nature and differs from person to person so the features and classifiers should adapt that changing character. Therefore, it is difficult to suit a single feature extraction or classification algorithm for the EEG signal. Many BCI methodologies are tried in the literature and the studies that rely on mental task or motor imagery discrimination are regarded as flexible methods [5,6]. The mental task based BCI studies enable independence to the user and to the system developer without the need for an extra monitor or a gaze tracker. Decoding mental tasks can be done with event related potentials (ERP), visually evoked

*Corresponding author

potentials (VEP) and with EEG. ERPs and VEPs require additional interfaces, such as screens of alphanumeric characters or gaze tracking devices.

The principal aim of this study is to introduce a multi-class EEG decoding for BCIs, with a novel pairwise output coding method. For building binary classifiers, one-versus-all (OVA) or one-versus-one (OVO) classification techniques are generally used. The choice between the two methods is based on the computation time and data storage. OVO seems faster and more memory efficient especially with LDA and SVM classifiers [11]. In the classical OVO scheme, a classifier is trained between each pair of classes and the final class of a test sample can be predicted by the max-win voting strategy. Whereas ties can arise in the voting and that could affect the final prediction badly [11]. Different from the literature, a new modification is added to the validation step, in which a train repetition number is introduced, is implemented while determining the training data of the classifiers. The train data are randomly selected from the rest of the test set, and they are obtained with multiple training sessions. The final performance of the classifiers was compared with the classical OVO results showing a higher classification ability.

Moreover, this study introduces a new data set which was recorded during mental and motor task experiments at Mechanical Engineering Department of Karadeniz Technical University. Whereas most of the previous studies were generally conducted from common two data source which are Keirn and Aunon's five class mental task data [12] and and Schlögl and Pfurtscheller's four class motor imagery data [7]. Huan and Palaniappan [13] designed a bi-state BCI for the five class mental task data given in [12], and they used three different feature extraction methods with NN-classifiers. Flores and friends [14] also developed an architecture based on adaptive neuro-fuzzy inference systems through recurrent neural networks. They used five class mental task data in [12]. Solhjoo and friends [15] used the dataset in [7]. They studied the performance of Hidden Markov Models in classification. Tolić and Jović [16] classified the wavelet transformed EEG signals with Neural Network. They studied with the mental task data from [12] and imagined motor tasks [17]. Apart from their work, experimental paradigm enables a user centered flexible environment for performing real time BCI applications in the future.

Furthermore, one more significant contribution of this study is the channel reduction to shorten the data processing time for online BCI applications. Depending on the demand from final-users of BCI applications, portable and easy-to-use devices are encouraged to be developed [4].

This paper is organized as follows. Section 2 gives a brief description of the recording procedure of the experimental design and the data used. Section 3 presents the details of the feature extraction approach and Section 4 contains the improved one-versus-one classification mode for LDA and SVM classifiers. In Section 5, the multiclass classification performance of classifiers for six subjects is reported along with a discussion of the results. Finally, the concluding remarks are given in Section 6.

2. Materials and methods

2.1.Experimental setup and data description

A 64 channel Biosemi ActiveTwo EEG system was used to record the EEG data. All experiments were carried out at the Mechanical Engineering Department of Karadeniz Technical University [18]. Apart from the literature, the recordings were all performed eyes closed which enables participants to concentrate thoroughly. Furthermore, eye blinks and eye movements produce a high amplitude signal called electrooculogram (EOG) that can be many times greater than the EEG signals which are regarded as artifacts [19,20]. For a good and accurate classification, the artifacts added to the EEG signal during the recording session must be removed from the signal itself. One way of solving this problem is to reject the eye blinked segments or the whole trial of that signal. However this can also cause to miss the valuable part of the data and additionally this could not be efficient when there is limited data [21].

1) Participants: The study was performed with 6 healthy participants who were initially naïve to the use of an EEG and the tasks except Subject 1. The 5 men and 1 women, with a mean (standard deviation, SD) age of 30.5 (14.4), had no medical diseases and were all right-handed. Each volunteer participated in several sessions over a period of 2-3 weeks. All of the participants signed the informed consent form before the experiments.

2) Procedure : The subjects were seated on a comfortable chair in a dim lighted, silent room during the recordings. Before each trial, they were informed about the type of the task (resting state, multiplication, right hand, etc.) by auditory cues. The sequence of mental and motor tasks was as follows: resting state, mental arithmetic, imagination of right hand movement, imagination of left hand movement, and imagination of letter 'A' (see Figure 1). Each trial lasted 10 seconds and the interval between consecutive tasks was about 3-4 seconds. The first 2 seconds in trial were the task preparation time for the subject. The experiments comprised of 5 experimental runs of 20 trials each (100 trials per task in total). The details of each task are provided below:

- Resting state (RS): The subjects were asked to sit and relax as much as possible without thinking anything.
- Mental arithmetic task (MA): The subjects were given a two-digit multiplication problem to solve in mind without vocalizing or any movement (e.g. 24×76

=?). The problems were not repeated. After the trial, the subject verified whether he reached the solution or not.

• Right hand imagination task (RH): The subjects were told to imagine right hand movement.

• Left hand imagination task (LH): The subjects were required to imagine left hand movement.

• Letter 'A' imagination task (LA): The subjects were told to imagine the letter 'A' in their mind.

3) Recordings: EEG data were recorded from the subjects during the experiment, using a 64-Channel Biosemi ActiveTwo EEG system with Ag/AgCl electrodes The international 10-20 electrode placement system was used. The grounding electrodes CMS and DRL were mounted on the back of the head. The EEG signals were sampled at 512 Hz.

4) Channel selection: We selected 9 channels from four different brain regions and hemispheres, i.e. frontal (F3/4), central (C3/4), parietal (P3/4, Pz) and occipital (O1/2). Because, each region was constituted with an EEG pair where different EEG rhythms can distinguish patterns of neuronal activity associated with specific motor and cognitive processing functions. Any change in brain patterns could result from different forms of processing or computation in the brain and represent different rhythmic states [7,9-11,13,16,21]. Hence, the alpha wave can be detected primarily from the occipital lobe (O1 and O2), but also from the parietal (P3 and P4) and frontal regions (F3 and F4) of the cerebral cortex. The motor imagery of human right/ left hand is typically reflected in EEG spectra in the beta rhythm obtained from C3 and C4, and mental arithmetic is mostly reflected in frontal cortex at F3 and F4. Feature extraction and classification were performed at each single channel.

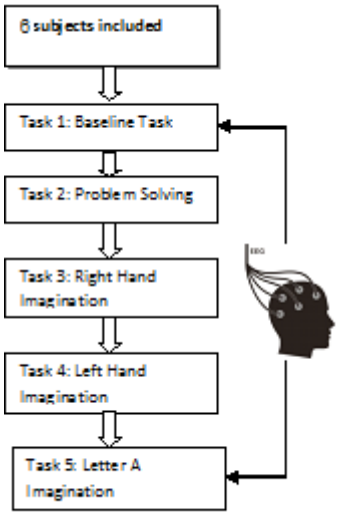


Figure 1. Order of the experiments

2.2. Data preprocessing

The raw data obtained from the Biosemi system is transferred to the Matlab environment. The raw data was normalized by Cz channel's data by subtracting it from the remaining 63 channels. The data were visually inspected, the beginning two seconds part was excluded from the entire 10 seconds signal because of the rough changes at the beginning of the imagination task. Then, the rest 8 seconds signal is divided into two 4 seconds signal by making 200 trials for each task. At the final stage, totally 6000 data samples (200 data samples x 5 tasks x 6 subjects) are collected to be analyzed. For online BCI studies, the user could be trained for some trials before the online application instead of excluding the data. The signal analyses are done on the data samples of 10 channels separately, (F3, F4, C3, C4, Cz, P3, P4, Pz, O1 and O2) which are given in Figure 2. Before the feature extraction step, the EEG signal was filtered using the 10^{th} order 50Hz low-pass digital Butterworth filter.

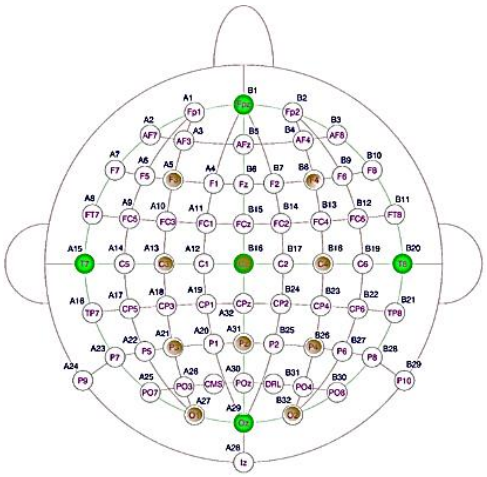


Figure 2. International 10-20 electrode placement system with the selected 10 channels in yellow

2.3. Feature extraction

Many feature extraction methods from basic [22] to highly complex ones [23-25] are proposed in BCI history. As the accuracy, ease of use, efficiency and speed are important parameters to consider [26], the feature extraction approach proposed in [27] is used in this study. This method relies on the band powers of EEG signal which is a common and powerful technique to distinguish different frequencies [28-30]. There, a stable pattern in the PSD was observed with different amplitudes for all subjects and for all tasks. This biologically phenomenon allows a classification between different mental tasks. Based on this biological phenomenon, we extracted three features from the alpha (8-13 Hz) and beta (13-30 Hz) bands of PSD by searching the local peak values in the alpha and beta bands separately. The PSD based on Welch Periodogram: a hamming window of 1024 points was used with a 50% overlap between adjacent windowed sections was computed first. Then the first feature is selected as the highest PSD peak value in the alpha band (8-13 Hz), which is referred to as f1 in the Figure 3. The second and third features are the arbitrary first and second highest PSD peak values in beta band (13-30 Hz) , which are referred to as f2 and f3 in Figure 3.

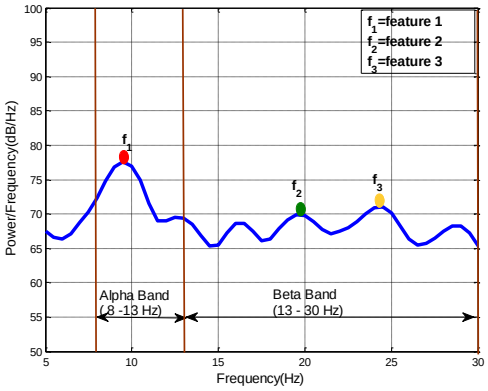


Figure 3. Proposed feature extraction scheme. The curve indicates the PSD of EEG in alpha and beta bands. We select the highest PSD peak value in the alpha band and first two highest PSD peak values in the beta band as the features.

3. Classification taxonomy

Classification is the act of assigning a predefined class to each instance. For this discrimination, we used LDA and SVM classification methods because of their good performances[6,7]. Both of these classification methods were originally designed for binary classification, but in this study they are adapted for multi-class problems. The simplest form is to build independent binary problems and to predict the score according to each binary classifiers result. The structure of binary classifiers are build with one-versus-all (OVA), one-versus-one (OVO) or the error correcting code techniques which are commonly used. The decision between OVO and OVA totally depends on the data type. It is stated that OVO is faster and more memory efficient than the OVA [31]. Regarding with this truth, an extended approach of OVO technique which is tested with LDA and SVM is developed in this study.

3.1. LDA classifier

LDA is an easily implemented, classical classification method. Hence, it is very popular and is often used as the baseline method for comparison with different classification methods. In this method, the data is projected onto a lower-dimensional vector space such that the ratio of the between-class distance to the within-class distance is maximized. Therefore, maximum discrimination is obtained. The optimal projection can be computed by applying the eigendecomposition on the scatter matrices.

According to Fisher's two class LDA, the multivariate observations x are transformed to univariate observations y such that the y's derived from the two classes are separated as much as possible [8]. First of all, the feature vector x is mapped by the following linear transformation in (1):

$$\mathbf{y}=\mathbf{V}^{\mathbf{T}}\mathbf{x} \tag{1}$$

where V represents the projection matrix. It is determined by maximizing the ratio of between-class variance to within-class variance. The within-class variance matrix and between-class variance matrix are defined below with equations (2) and (3):

$$S_W = \sum_{i=1}^{K}\sum_{l=1}^{L_i} (x^l - \mu_i)(x^l - \mu_i)^T \tag{2}$$

$$S_B = \sum_{i=1}^{K} (\mu_i - \mu)(\mu_i - \mu)^T \tag{3}$$

where K is the number of classes, and μ_i is the mean vector of the class i, L_i is the number of samples within-class i and μ is the mean of the entire training sample set. The projection matrix V is calculated by eigenvectors of matrix $S_W^{-1}S_B$. Once the transformation is done, the classification is then performed in the transformed space based on some distance metric, such as Euclidean distance given in (4);

$$d(p,q) = \sqrt{\sum_{\mathrm{i}} (\mathrm{p_i}\text{-}\mathrm{q_i})^2} \tag{4}$$

The final class is attained as $\arg_k \min d(zV, \mu_k V)$, upon the arrival of the new instance z (a row vector). Here, μ_k is the centroid of the k-th class.

The extended OVO approach is built to apply this binary LDA method to multiclass classification. We implemented a computer program for LDA algorithm in MATLAB® for two-class and multi-class classification.

The multi-class case consists of several two-class runs. In a classical OVO approach, K(K-1)/2 binary classifiers are built. A new example is tested according to the **max-win** voting strategy among the classifiers, and the class with the maximum number of votes is assigned. In some cases, ties can arise and the computation for that run is neglected. This is disadvantageous when there is limited data. Apart from the existing OVO approach, the proposed approach in this study uses **K** binary classifiers to classify the K class data. The data and the class labels are introduced as $\{x_i, y_i\}_{i=1}^{N}$ where $x_i \in R$ are training samples with input vectors and $y_i \in \{-1, +1\}$ are class labels. Different from the literature, by means of each electrode channel's separate classification result, the final class label is attained.

The extended OVO approach can be explained in following steps:

We have five classes in this study. Each class contains 200 data samples per subject and per electrode channel, in total 1000 data samples per subject. In theory, we

used 50% of the data for testing and 50% for training which means that the data of a subject is split equally as training and testing. This selection is made randomly by a two-fold cross validation in each run.

Step 1: For each binary classifier, select 50 train data for all task pairs. A task pair contain 100 data. K=5 binary classifiers are built as a module, for instance k1-k1,k1-k2,k1-k3,k1-k4,k1-k5, as depicted in Figure 4. As the proposed method constructs K task pairs for K class data, we use ½ of the train data of each class. Label half of the train data as +1 and -1 in each classifier. As a result of this methodology, 50 train data of the same class will be labeled as +1 and another 50 train data from the same class will be labeled as -1 (k1-k1). The rest of the pairs (k1-k2,k1-k3,k1-k4,k1-k5) are constructed such that 50 train data from one class will be labeled as +1 and another 50 train data from the other class will be labeled as -1.

Step 2: Take another random 50 test data which are totally different from the train data, from any of the classes. As we don't know the test class label, we will compare it with all the K classes. By considering k1-k1 case, if it's true class is k1, the greatest possible majority of a test sample will definitely be 50. The 50 is not the final classification performance value, it is only a mathematical way for us to predict the class labels. Alike pair's comparisons will be different from 50. It is the key point which enables us to predict the test class label. The program is designed to compute all nine channels' percentages separately. The highest classification performance results are observed on all of the nine channels. Then, the predicted class labels are attained by using max-win voting strategy among nine channels. The predicted class labels are displayed as P1, P2, P3, P4 and P5 in Figure 4. In Table 1, the K class module results that give way to decision on final class label is indicated. You will see from Table 1 that, the task pairs Task1- Task1, Task2-Task2, Task3-Task3, Task4-Task4 and Task5-Task5

have values around 50. The task pair results are independent from each other.

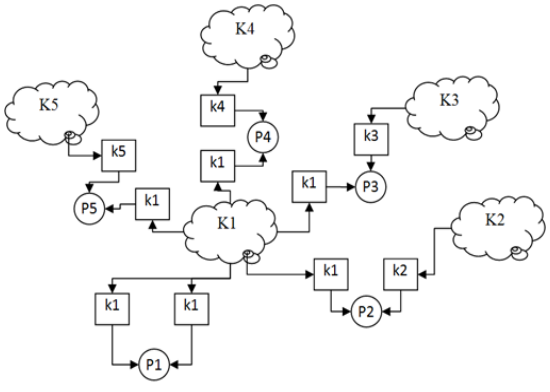


Figure 4. Schematic explanation for extended OVO approach

Step 3: Finally, the overall program is repeated 100 times to obtain a mean value.

3.2. SVM classifier

SVM is a powerful classifier which has demonstrated its excellent generalization properties in various BCI applications [7,8,10]. The basic idea of SVM is finding the optimal separation hyperplane by maximizing the margin.

The general SVM solution is obtained from the following optimization problem [32] given in (5):

$$\min_{w,b,\xi} \frac{1}{2} w^T w + c\sum_{i=1}^{N} \xi_i$$

subject to

$$y_i(w^T \Phi(x_i) + b \geq 1 - \xi_i), \quad \xi_i \geq 0, i = 1, ..., N \qquad (5)$$

Input vector $x_i \in R^n$, and class labels $y_i \in \{+1, -1\}$. N is the number of training samples, *w* is the normal vector and *b* is the bias of the separation hyperplane. If $\Phi(x) = x$, then SVM is called linear, otherwise, if x is mapped to a higher dimensional space, then it is called nonlinear SVM. In that case, the training data might not be separated without error, the slack variable $\xi_i \geq 0$ and c>0 have to be introduced. The output of a binary SVM classifier can be computed by the following expression in (6):

$$y = sign(\sum_{i=1}^{N} \alpha_i y_i k(x_i, x) + b) \qquad (6)$$

where $\alpha_i \geq 0$ are Lagrangian multipliers obtained by solving a quadratic optimization problem, and $k(x_i, x_j)$ is called Kernel function. The most commonly used kernel function is the Gaussian RBF function which is also used in this study as in (7),

$$k(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \qquad (7)$$

σ is the user defined parameter which shows the kernel function's width.

MATLAB's SVM package which is originally designed for binary classification is applied for multiclass classification according to the improved OVO approach that is used in LDA.

Table 1. Proposed final class label decision table

Tasks		**Percentages % with Standard deviations ±**								
Training set	**Test set**	**F3**	**F4**	**C3**	**C4**	**P3**	**P4**	**Pz**	**O1**	**O2**
Task1	Task1	**50.28 ±9.62**	**49.90 ±9.51**	**50.98 ±10.11**	**50.98 ±8.81**	**50.80 ±14.63**	**50.18 ±9.94**	**51.00 ±10.10**	**49.50 ±9.29**	**49.96 ±10.13**
Task2		2.64 ±1.86	1.84 ±2.23	4.76 ±2.89	7.78 ±2.82	4.48 ±2.99	4.74 ±3.03	6.32 ±3.47	5.64 ±3.28	7.70 ±4.33
Task3		9.62 ±4.14	9.98 ±3.79	9.84 ±4.73	8.60 ±3.28	11.66 ±4.79	14.40 ±4.64	14.14 ±4.46	20.40 ±6.72	33.44 ±7.20
Task4		10.96 ±4.41	10.38 ±4.19	13.74 ±4.95	7.54 ±3.13	14.72 ±5.31	14.54 ±4.87	17.14 ±5.49	26.56 ±6.96	34.98 ±8.14
Task5		14.70 ±6.22	15.88 ±6.18	20.04 ±6.23	28.76 ±7.16	13.26 ±5.59	16.22 ±5.21	16.62 ±4.43	15.36 ±5.64	21.44 ±6.50
Task1	Task2	3.82 ±2.77	5.38 ±2.97	6.44 ±3.44	9.42 ±4.06	24.90 ±6.25	10.32 ±3.72	10.72 ±4.11	20.16 ±4.84	22.34 ±5.18
Task2		**49.74 ±9.17**	**49.32 ±9.86**	**49.52 ±9.54**	**50.70 ±10.11**	**48.08 ±15.84**	**51.52 ±9.47**	**51.92 ±9.90**	**50.26 ±9.77**	**51.90 ±9.48**
Task3		23.50 ±5.70	23.24 ±5.68	24.20 ±5.64	48.20 ±8.29	33.60 ±7.72	28.34 ±6.32	26.52 ±5.95	29.38 ±6.48	28.44 ±6.08
Task4		21.82 ±5.15	21.98 ±5.88	22.18 ±6.04	43.96 ±8.28	31.06 ±7.53	26.66 ±5.96	24.32 ±5.38	25.20 ±6.50	27.76 ±6.04
Task5		16.86 ±5.68	16.04 ±5.17	20.98 ±7.36	23.28 ±6.66	33.32 ±6.22	24.30 ±6.48	25.16 ±6.05	33.54 ±7.37	33.24 ±6.65
Task1	Task3	10.44 ±3.97	9.66 ±4.24	8.50 ±3.48	9.60 ±3.89	26.98 ±7.21	13.74 ±5.54	12.70 ±4.39	35.14 ±7.44	31.08 ±7.72
Task2		29.78 ±6.30	30.24 ±6.59	32.84 ±6.23	41.56 ±8.16	27.36 ±6.44	27.96 ±6.36	26.00 ±5.99	26.66 ±6.06	21.76 ±6.85
Task3		**49.76 ±10.57**	**50.32 ±10.49**	**47.86 ±10.28**	**51.50 ±9.41**	**49.86 ±14.44**	**51.70 ±9.67**	**51.78 ±9.02**	**51.00 ±9.70**	**50.40 ±8.98**
Task4		42.74 ±10.04	40.68 ±9.16	46.20 ±9.29	38.72 ±8.65	47.38 ±13.53	46.04 ±10.12	47.16 ±9.67	47.86 ±8.53	47.98 ±9.22
Task5		22.60 ±7.28	24.78 ±7.16	25.94 ±6.88	23.14 ±5.63	34.12 ±8.43	26.12 ±8.35	29.08 ±8.35	39.38 ±8.76	36.86 ±7.37
Task1	Task4	13.86 ±4.79	14.10 ±5.35	13.98 ±4.67	11.08 ±4.56	32.64 ±7.46	16.50 ±5.01	18.56 ±5.61	42.34 ±8.72	33.44 ±9.84
Task2		25.04 ±6.33	28.80 ±6.61	27.14 ±6.65	46.30 ±8.04	22.54 ±6.85	25.48 ±5.48	22.88 ±5.79	21.16 ±5.67	19.94 ±5.68
Task3		50.28 ±8.67	53.34 ±9.00	45.16 ±8.51	47.88 ±7.28	46.86 ±12.34	51.70 ±9.84	51.38 ±10.15	44.34 ±9.16	48.98 ±11.67
Task4		**49.94 ±9.61**	**51.24 ±10.49**	**50.70 ±9.05**	**49.94 ±9.89**	**52.26 ±16.25**	**51.26 ±9.38**	**50.74 ±10.69**	**50.30 ±10.43**	**48.72 ±12.10**
Task5		29.04 ±7.12	29.70 ±7.15	31.62 ±7.84	25.64 ±6.51	34.72 ±7.66	25.88 ±6.15	30.26 ±6.89	33.42 ±7.23	36.94 ±8.34
Task1	Task5	22.12 ±5.39	23.88 ±5.85	23.48 ±6.54	25.12 ±6.81	31.96 ±6.84	31.78 ±6.70	25.32 ±6.16	31.20 ±6.57	36.92 ±7.49
Task2		22.72 ±6.25	23.06 ±5.70	24.76 ±6.20	19.42 ±6.05	33.02 ±6.57	27.02 ±6.05	29.12 ±6.21	31.52 ±6.50	29.04 ±6.10
Task3		38.32 ±7.40	36.56 ±6.81	27.98 ±6.68	22.86 ±5.75	41.78 ±9.80	30.44 ±5.98	30.26 ±7.17	46.78 ±7.17	44.16 ±8.06
Task4		36.46 ±8.27	33.70 ±6.93	28.62 ±6.41	16.92 ±5.28	38.28 ±8.87	27.56 ±6.41	29.14 ±6.99	42.54 ±6.48	39.46 ±8.64
Task5		**50.54 ±9.17**	**50.00 ±10.50**	**49.36 ±8.96**	**49.10 ±8.97**	**47.86 ±14.07**	**50.80 ±8.36**	**48.48 ±9.15**	**51.32 ±9.53**	**51.08 ±10.24**

4. Results and discussion

The classification performance of the proposed approach and the regular OVO with LDA and SVM were given in Table 2 for all six subjects and nine channels. The classification performance per mental and motor task is defined as the number of correct predictions in a run over the total number of data points in that run and expressed in % for 100 times trials. In the tables, a two group representation is performed.

First group represents the results of extended OVO approach and the next group shows the results of regular OVO approach. The leftmost column represents the subjects from one to six, the next column shows the classifiers, the following five columns represent the classification accuracies of five tasks and the last column represents the mean values of all tasks. All computations were done by using the same number of train and test data. It is seen from Table 2 that, the data of Task1 is finely discriminated from the rest in all the subjects. The highest classification rates are 100 with SVM and 99.96 with LDA for Subject1. It is 99.90 for Subject 2 with SVM, 63.90 for Subject 3 with SVM, 75.92 for Subject 4 with SVM, 83.44 for Subject 5 with SVM and finally 79.39 for Subject 6 with LDA. Moreover, the classification performances of Task 2 and Task 5 are fairly good for Subject 1 and Subject 2, and medially for the rest four subjects. The classification results of motor tasks, Task 3 and Task 4 which are Imagination of Right Hand and Imagination of Left Hand, are lower than the mental task results. The highest rates are again observed with Subject 1 which is 75.44 for Task 3 and 75.20 for Task 4 which are obtained with SVM. The highest mean values of subjects for different tasks with LDA are 79.34 and 81.64 with SVM. The overall performances of all subjects show that the results are subject dependent. Hence, more trained subjects (Subject 1, familiar with EEG and BCI studies) produce a higher classification performance. To be able compare the proposed extended OVO approach with the regular OVO approach, the classification accuracies of five tasks and the mean values of all tasks were also included in Table 2. The maximum mean classification performances of subjects were achieved by Subject 2 as 69.36% with LDA and 61.95 % with SVM which are very poor when compared with extended OVO results (81.38 % LDA and 89.98 % SVM). The mean classification results calculated due to six subjects and also for five tasks show that extended OVO approach performs better than the regular method which is the superiority of the proposed approach. The reason for this performance drop may be the occurrence of many ties at the max-win voting strategy during the final class decision of regular OVO approach. It is clear in Table 2 that the classification performance of Task 1 and Task 2 data samples are very high for all six subjects with both methods. An important finding from Table 2 is that Task 3 and Task 4 (the motor imagery tasks) have very low performance results for both OVO approaches (54.83% mean with extended OVO and 58.37% mean with regular OVO). The computation times for both classifiers are given in Table 7. The proposed method has shorter computation times (0.9 seconds for LDA and 2.42 seconds for SVM) than the regular OVO approach (3.44 seconds with LDA and 4.21 seconds with SVM).

Another point which is observed for this subject is reducing the number of electrodes from 9 to four depends on the cortex placements (Table 3). Four electrodes F3, F4, C3 and C4 are selected because they are at the frontal regions and sensorimotor area of the cortex and moreover they are more important for recognition of mental and motor tasks than the rest of the channels. The results of this study also supports this truth that during motor tasks while the classification accuracies with four channels (F3, F4, C3, C4) are higher, increasing the number of channels to nine does not increase the classification accuracy. So, it can be concluded that, the use of four channels (F3, F4, C3 and C4) data is enough also for the motor task classifications.

In Tables 4-5, the classification results of each classifier for left (F3, C3, P3, O1) and right (F4, C4, P4, O2) hemisphere electrodes and the midline electrode (Pz) is given. The classification results of (F3 C3 P3 O1 Pz) are as follows: for Task 1 100 with LDA and 100 with SVM, for Task 2 95.10 with LDA and 100 with SVM, for Task 3 47.54 with LDA and 81.44 with SVM, and for Task 4 57.66 with LDA and 69.06 with SVM and finally for Task 5 90.64 with LDA and 99 with SVM. Whereas, the classification results at Table 4 are very close to the results of Table 5 which means that we cannot judge about hemispheric changes under these circumstances for this subject. However, during the imaginary tasks, the performance of the task results are much more dependent on how well the imagination performed. Both 5-electrode configurations had better classification results for right than for left arm movement imagination (see Tables 4-5), instead for having better results for the corresponding arm movement. A possible explanation for this fact could be that the dominant hand characteristics may affect the classification results.

One important point of this study is to obtain an ideal training data without discarding noisy or bad data during the analysis. While working with an online BCI system, it would be difficult to discard the data. For this reason, a parameter called the train repetition number is introduced to select a fine train data set. The train data set is randomly selected for n times (n=1,…,5). The selection of train data set is different from cross-validation. Here, the selection of a fine train data set is searched. First, the data is split to equal number of train and test sets. Then, by keeping the test set unchanged, train set is formed after n repetitions (n=1,…,5). As the n increases, the computation time increases which is not preferred for online applications. The effect of repetition number to the computation time for LDA and SVM classifiers is given in Table 6. It is obvious from the table that, the response time of LDA is faster than the SVM's and reducing the number of electrodes used from nine to four makes a 0.25 seconds time consuming with LDA, on the other hand, this is about 0.56 seconds with SVM.

Table 2. Multi-class classification results with standard deviations for 9 electrodes with extended OVO and regular OVO approaches. S:Subjects, C:Classifiers, M:Mean

		Extended OVO						**Regular OVO**					
S	**C**	**Task 1**	**Task 2**	**Task 3**	**Task 4**	**Task 5**	**M**	**Task 1**	**Task 2**	**Task 3**	**Task 4**	**Task 5**	
S1	LDA	99.96 ±0.19	97.88 ±1.28	48.90 ±4.14	61.18 ±4.68	90.66 ±3.16	79.71	94.00 ±0.10	26.00 ±5.18	56.00 ±6.16	48.00 ±3.98	74.00 ±3.56	59.20
	SVM	100 ±0.00	100 ±0.00	75.44 ±5.03	75.20 ±4.23	99.28 ±0.75	**89.98**	93.00 ±2.00	50.00 ±5.10	38.00 ±7.03	44.00 ±4.28	72.00 ±5.75	59. 40
S2	LDA	95.88 ±2.19	75.56 ±5.10	55.00 ±4.72	80.50 ±2.22	100.00 ±0.00	81.38	91.08 ±0.19	75.56 ±7.44	45.50 ±3.76	50.50 ±2.22	83.50 ±4.00	**69.36**
	SVM	99.90 ±0.10	53.30 ±3.90	50.40 ±4.50	65.76 ±3.88	95.22 ±1.98	72.91	86.10 ±4.22	73.30 ±4.20	39.40 ±5.50	30.76 ±3.12	80.22 ±2.18	61.95
S3	LDA	60.90 ±6.88	73.20 ±5.16	50.00 ±6.28	45.18 ±7.13	64.20 ±5.98	58.69	59.90 ±3.48	61.20 ±5.00	30.40 ±6.28	45.18 ±7.13	64.20 ±5.98	52.17
	SVM	63.90 ±7.12	91.96 ±1.24	55.16 ±2.78	65.00 ±4.44	52.32 ±5.22	65.66	65.20 ±5.16	81.16 ±4.08	55.26 ±7.12	50.10 ±4.66	49.72 ±4.42	60.28
S4	LDA	65.00 ±2.78	75.04 ±1.56	45.44 ±5.82	38.16 ±7.64	75.48 ±3.32	59.82	55.04 ±8.18	63.84 ±5.56	40.32 ±4.20	29.16 ±6.10	55.14 ±4.03	48.70
	SVM	75.92 ±3.00	78.20 ±4.28	42.00 ±6.34	45.00 ±5.28	77.28 ±4.18	63.68	60.96 ±2.08	68.20 ±2.98	46.00 ±4.34	55.00 ±3.70	70.28 ±2.10	60.08
S5	LDA	71.00 ±3.66	75.66 ±2.27	39.30 ±5.76	42.50 ±5.34	69.72 ±4.12	59.63	64.00 ±3.66	70.66 ±2.27	40.30 ±5.76	41.50 ±5.34	63.72 ±4.12	56.03
	SVM	83.44 ±1.02	71.56 ±2.98	55.79 ±4.34	57.95 ±4.72	77.16 ±2.16	69.18	73.44 ±1.02	70.06 ±2.98	67.70 ±4.34	53.90 ±4.72	66.16 ±2.16	66.25
S6	LDA	79.39 ±2.06	72.98 ±2.74	54.90 ±5.70	49.00 ±5.52	76.00 ±4.54	66.45	55.39 ±4.08	68.08 ±3.24	53.90 ±4.10	47.30 ±5.60	60.50 ±3.98	57.03
	SVM	66.72 ±5.02	78.44 ±2.72	50.22 ±5.12	41.34 ±4.86	73.92 ±3.08	62.12	56.72 ±4.86	70.34 ±4.12	55.60 ±4.70	51.34 ±3.86	68.12 ±5.18	60.42
M	LDA	78.68	78.91	48.92	52.75	79.34	67.72	69.29	60.89	44.40	43.60	66.84	**57.19**
	SVM	**81.64**	78.38	**54.83**	**58.37**	79.19	70.48	72.57	68.84	**50.32**	**47.50**	67.75	**61.39**

It was observed that, the classification performances of motor tasks are so low compared to the other tasks. Using several train repetition numbers also affects the classification accuracy of Right Hand Imagination task results which are displayed in Figure 5 and Figure 6. By considering Figure 6 and Table 6, the train repetition number three is suggested for optimum classification accuracy. As it is noticed in Figures 5-6, any increase in train repetition number also increases accuracy at 15.2 in LDA and at 12.1 in SVM for right hand imagination task.

The accuracy, sensitivity and specificity values for the classifiers are obtained for the two class case and the results of a binary SVM are given in Table 8. Sensitivity and specificity measures are used to measure the statistical performance of a binary classification test. Sensitivity is defined as the proportion of number of true positives to the total number of true positives and number of false negatives. On the other hand, specificity is defined as the ratio of number of true negatives over the total number of true negatives and false positives. The definition of sensitivity and specificity are given in Table 7. All calculations are performed by writing a MATLAB

script. Accuracy is used to see how well the result of a binary classifier correctly identified [33]. Hence, an accuracy of 100 means that the tested values are exactly the same as the true values. The results are obtained for each electrode channel separately. The overall accuracies of Task 1 are around 90. The overall sensitivity values are around 87 and the overall specificity values are around 92. These results also support the good classification ability of the proposed method and the classifiers.

Table 3. Multi-class classification results of Subject 1 for 4 channels, (F3 F4 C3 C4)

Classifier		Test Data					Mean
		Task 1	Task 2	Task 3	Task 4	Task 5	
LDA	Accuracy	100	95.10	47.54	57.66	90.64	78.188
	Standard Deviation	±0.00	±1.76	±4.95	±4.15	±2.72	
SVM	Accuracy	100	100	**81.44**	69.06	99.00	89.900
	Standard Deviation	±0.00	±0.00	±3.53	±4.59	±0.90	

Table 4. Multi-class classification results of Subject 1 for 5 channels, (F3 C3 P3 O1 Pz)

Classifier		Test Data					Mean
		Task 1	Task 2	Task 3	Task 4	Task 5	
LDA	Accuracy	99.96	89.58	46.48	54.80	89.36	76.036
	Standard Deviation	±0.19	±2.79	±5.51	±4.86	±2.59	
SVM	Accuracy	100	100	**87.04**	72.66	99.56	91.852
	Standard Deviation	±0.00	±0.00	±3.55	±4.31	±0.61	

Table 5. Multi-class classification results of Subject 1 for 5 channels, (F4 C4 P4 O2 Pz)

Classifier		Test Data					Mean
		Task 1	Task 2	Task 3	Task 4	Task 5	
LDA	Accuracy	100	94.72	47.88	57.16	89.58	77.868
	Standard Deviation	±0.00	±2.63	±4.73	±5.26	±2.19	
SVM	Accuracy	100	99.96	**79.70**	69.74	98.92	89.664
	Standard Deviation	±0.00	±0.19	±4.02	±3.50	±0.89	

Table 6. Effect of repetition number to computation time. C:Classifiers, NE:Number of electrodes

C	NE	Repetition Number					Regular OVO
		1	2	3	4	5	
		Computation Time (seconds)					
LDA	9	0.90	1.33	1.78	2.24	2.64	**3.44**
	4	0.65	0.85	1.03	1.23	1.43	
SVM	9	2.42	4.18	5.54	7.21	8.85	**4.21**
	4	1.82	2.94	3.70	4.76	6.00	

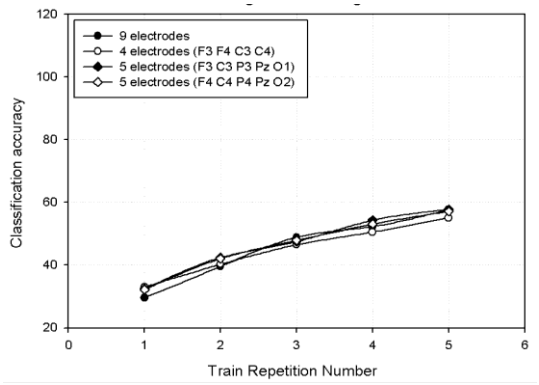


Figure 5. Change in classification accuracy for right hand imagination with LDA

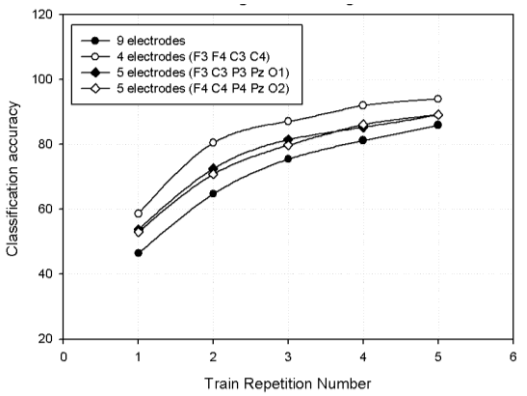


Figure 6. Change in classification accuracy for right hand imagination with SVM

Table 7. Calculation of sensitivity and specificity

	Disease present	Disease absent
Test Positive	a (TP)	b (FP)
Test Negative	c (FN)	d (TN)
	Sensitivity	Specificity
	a/(a+c)	d / (b+d)

TP:True Positive, FP:False Positive, FN:False Negative, TN:True Negative

5. Conclusion

The main research finding of this study is proposing an alternative solution step that brings about an extended approach for one-versus-one classification of data. With this method, the computation time and the data storage are lessened. One other finding is the necessity of certain electrode channels required for BCI systems. For mobile BCI systems, reduced number of all technical equipment including electrode channels are preferred. Therefore, in this study, a four channel system provides results that are on par with more channels which is a success. We obtained better classification performance with SVM; on the other hand less computation time with LDA which is a fact. A thorough comparison between mental and motor tasks and between right and left hemispheres were searched. For homogenous seperation of train data, a repetition number is introduced. Moreover, the difference between the experinced and novice subjects were searched and it is concluded that the a short training period for subjects before the online applications will improve the overall performance.

It is also observed that, selecting the proper electrode channel is an important task. In this study, it is concluded that, the use of frontal and central lobe electrodes would be enough to distinguish some basic tasks especially mental and motor tasks separately with the proposed features and classifiers.

The main contribution of this paper is its original extended OVO output coding methodology which can be used instead of regular OVO algorithm during the multiclass classification scenarios. An extra contribution is the use of less channel data that reduces the processing time and producing a quick response.

Table 8. Accuracy, Sensitivity and Specificity values for a two-class SVM

Tasks			Channels Percentages %								
Training set	Test set		**F3**	**F4**	**C3**	**C4**	**P3**	**P4**	**Pz**	**O1**	**O2**
Task 1 Task 2	**Task 1**	Accuracy	96.80 ±1.66	96.76 ± 1.54	94.17 ± 1.95	92.69 ± 2.42	84.93 ± 3.14	92.44 ± 2.08	91.71 ± 2.17	86.82 ± 2.51	84.31 ± 2.96
		Sensitivity	96.27 ±2.65	95.11 ± 2.73	93.80 ± 2.74	91.39 ± 3.84	78.60 ± 3.98	90.38 ± 3.15	89.75 ± 2.95	82.00 ± 3.41	80.21 ± 3.90
		Specificity	97.47 ±1.94	98.66 ± 1.55	94.71 ± 2.78	94.39 ± 3.08	95.46 ± 3.08	94.95 ± 2.68	94.09 ± 3.04	93.77 ± 3.32	90.31 ± 4.28
Task 1 Task 3		Accuracy	90.40 ±2.42	90.80 ± 2.57	90.28 ± 2.56	91.55 ± 2.24	81.57 ± 2.94	86.37 ± 2.66	87.33 ± 3.03	71.80 ± 3.32	67.54 ± 3.63
		Sensitivity	88.96 ±3.79	89.83 ± 3.73	90.66 ± 4.02	90.02 ± 3.53	77.89 ± 3.96	86.16 ± 4.25	86.99 ± 4.24	69.24 ±4.09	68.52 ± 4.86
		Specificity	92.32 ±3.30	92.09 ± 3.05	90.42 ± 4.23	93.54 ± 3.20	87.14 ± 4.50	87.06 ± 3.59	88.13 ± 4.16	76.25 ± 5.11	67.26 ± 4.02
Task 1 Task 4		Accuracy	89.21 ±2.76	89.20 ± 2.71	86.92 ± 3.16	91.83 ± 2.65	77.52 ±3.73	85.61 ± 3.16	83.84 ± 2.83	65.93 ± 3.87	66.38 ± 3.51
		Sensitivity	87.37 ±3.84	86.77 ±3.89	86.64 ± 4.07	89.21 ± 4.00	73.51 ± 4.29	84.25 ± 4.12	82.90 ± 4.05	63.81 ± 3.90	67.48 ± 4.53
		Specificity	91.71 ±4.14	92.36 ± 3.60	87.60 ± 4.22	65.16 ± 2.95	84.07 ± 5.14	87.53 ± 4.21	85.37 ± 4.00	69.91 ± 6.03	65.91 ± 3.80
Task 1 Task 5		Accuracy	81.43 ±3.33	80.65 ± 3.09	77.74 ± 3.24	73.46 ± 3.84	76.67 ± 3.65	76.02 ± 3.15	79.58 ± 3.69	75.82 ± 3.46	69.55 ± 3.76
		Sensitivity	78.87 ±4.43	77.91 ± 4.35	77.13 ± 4.84	73.84 ± 4.88	72.23 ± 4.37	71.92 ± 3.74	76.20 ± 4.47	72.32 ± 3.78	67.41 ± 4.10
		Specificity	85.24 ±4.62	84.70 ± 4.22	79.23 ± 4.53	73.76 ± 4.76	84.47 ± 4.86	83.08 ± 5.13	84.78 ± 5.07	81.55 ± 5.40	73.28 ± 5.40

Finally, the study imparts that multiclass SVM using extended technique and combined with the proposed feature extraction algorithm can be used for classification of motor task EEG signals for various applications when verified with more subjects.

For further study, the current results obtained from this study would be supported with different BCI data sets.

Acknowledgments

This work is supported by Karadeniz Technical University Scientific Research Projects, No.2007.112.03.3, Trabzon, Turkey.

References

[1] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. & Vaughan, T.M. (2002). Brain-Computer Interfaces for Communication and Control, Clinical Neurophysiology, 113, 767-791.

[2] Hwang, H.J., Kim, S., Choi, S. & Im, C.H. (2013). EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey, International Journal of Human-Computer Interaction, 29(12), 814-826.

[3] Brunner, C., Birbaumer, N., Blankertz, B., Guger, C., Kübler, A., Mattia, D., & Ramsey, N. (2015). BNCI Horizon 2020: towards a roadmap for the BCI community. Brain-computer interfaces, 2(1), 1-10.

[4] Mühl, C., Allison, B., Nijholt, A., & Chanel, G. (2014). A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. Brain-Computer Interfaces, 1(2), 66-84.

[5] Bashashati, A., Fatourechi, M. (2007). A Survey of Signal Processing Algorithms in brain-computer interfaces based on electrical brain signals, Journal of Neural Engineering, 4, 32-57.

[6] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. & Arnaldi, B. (2007). A Review of Classification Algortihms for EEG-based Brain-computer Interfaces, Journal of Neural Engineering, 4, R1-R13.

[7] Schlögl, A., Lee, F., Bischof, H. & Pfurtscheller, G. (2005). Characterization of four-class Motor Imagery EEG Data for the BCI-competition, Journal of Neural Engineering, 2, L14-L22.

[8] Li, T., Zhu, S. & Ogihara, M. (2006). Using Discriminant Analysis for Multi-class Classification: An Experimental Investigation, Knowledge of Information Systems, 10(4), 453-472.

[9] Zhou, S.M., Gan, J.Q., & Sepulveda, F. (2008). Classifying Mental Tasks Based on Features of Higher-order Statistics from EEG Signals in Brain-computer Interface, Information Sciences, 178, 1629-1640.

[10] Liang, N.Y., Saratchandran, P., Huang, G.B. & Sundararajan, N. (2006). Classification of Mental Tasks from EEG Signals Using Extreme Learning Machine, International Journal of Neural Systems, 16(1), 29-38.

[11] Lee, F., Scherer, R., Leeb, R., Neuper, C., Bischof, H., & Pfurtscheller, G. (2005). A comparative analysis of multi-class EEG classification for brain computer interface. In Proceedings of the 10th Computer Vision Winter Workshop (195-204).

[12] Keirn, Z. A., Aunon, J. I. (1990). A New Mode of Communication Between Man and His Surroundings. IEEE Transactions on biomedical engineering, 37,12.

[13] Huan, N., Palaniappan, R. (2004). Brain Computer Interface Design Using Mental Task Classification, Multimedia Cyberscape Journal, 1, 35-43.

[14] Flores, M.E., Cortés, J.M.R., Gil, P.G., Aquino, V.A. (2013). Mental Tasks Temporal Classification Using an Architecture Based on ANFIS and Recurrent Neural Networks, Recent Advances on Hybrid Intelligent Systems, Studies in Computational Intelligence 451, 135-146.

[15] Solhjoo, S., Nasrabadi, A. M. & Golpayegani, M.R.H.(2005). Classification Of Chaotic Signals Using HMM Classifiers:EEG-Based Mental Task Classification, 13th Signal Processing Conference, September 4-8, 2005, Antalya,Turkey.

[16] Tolić, M. & Jović, F. (2013). Classification of Wavelet Transformed EEG Signals With Neural Network For Imagined Mental And Motor Tasks, Kinesiology, 45(1), 130-138.

[17] Schalk, G. (2009). EEG motor movement/imagery dataset. Retrieved November 18, 2011 from http://www.physionet. org/pn4/eegmmidb.

[18] Özmen, N. G. & Gümüşel, L. (2010). Mental and Motor Task Classification by LDA, MEDICON 2010, IFMBE Proceedings, 29, 172-175.

[19] Benbadis, S.R. (2006). Introduction to EEG. In: Lee-Chiong T. Sleep: A Comprehensive Handbook. Hoboken, NJ: Wiley & Sons Inc.

[20] Carriea, J., Gorodnitsky, I.F. & Kutas, M. (2004). Automatic removal of eye movement and blink artifacts from EEG data using blind component separation, Psychophysiology, 41.

[21] Wang, D., Miao, D. & Blohm, G. (2012). Multi-class motor imagery EEG decoding for brain-computer interfaces, Frontiers in Neuroscience, 6,151,1-13.

[22] Sun, S. & Zhang, C.(2006). Adaptive feature extraction for EEG signal classification, Medical Biological Engineering and Computing, 44, 931-935.

[23] Hsu, W.Y. (2010). EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features, Journal of Neuroscience Methods, 189(2), 295–302.

[24] Lin, Y., Wang, C.H., Jung, T.P., Wu, T. L., Jeng,

S.K., Duann, J.R. & Chen, J. H. (2010). EEG-Based Emotion Recognition in Music Listening, IEEE Transactions On Biomedical Engineering, 57(7), 1798-806.

[25] Barachant, A., Bonnet, S., Congedo, M. & Jutten, C. (2012). Multi-class Brain Computer Interface Classification by Riemannian Geometry, IEEE Transactions on Biomedical Engineering, 59(4), 920-928.

[26] Nijboer, F., OudeBos, D. P., Blokland,Y., Wijk, R. & Farquhar, J. (2014). Design requirements and potential target users for brain-computer interfaces–recommendations from rehabilitation professionals, Brain-Computer Interfaces, 1(1), 50-61.

[27] Özmen, N.G. & Gümüşel, L. (2011). Discrimination Between Mental and Motor Tasks of EEG Signals Using Different Classification Methods, 2011 IEEE International Conference on Innovations in Intelligent Systems and Applications, 15-18 June 2011.

[28] Subasi, A., Erçelebi, E., Alkan, A., & Koklukaya, E. (2006). Comparison of Subspace-Based Methods With AR Parametric Methods In Epileptic Seizure Detection, Computers in Biology and Medicine, 36(2), 195-208.

[29] Siuly, Y.L. & Wen, P. (2009). Classification of EEG signals using sampling techniques and least square support vector machines, Fourth International Conference on Rough Sets and Knowledge Technology 2009, LNCS 5589, 375-382.

[30] Mosquera, C. G., Trigueros, A.M., Franco, J.I. & Vaszquez, A.N. (2010). New feature extraction approach for epileptic EEG signal detection using time-frequency distributions, Medical and Biological Engineering and Computing, 48(4), 321-330.

[31] Rifkin, R. & Klautau, A. (2004). Parallel networks that learn to pronounce english text. Journal of Machine Learning Research, 5, 101–141.

[32] Vapnik V. (1995). The Nature of Statistical Learning Theory, New York, Springer.

[33] Anderson, C.W., Stolz, E.A. & Shamsunder, S. (1998). Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks, IEEE Transactions Biomedical Engineering, 45(3), 277-286.

Dr. Özmen *received her Bachelor Degree from Balıkesir University, Department of Mechanical Engineering in 2001. She got her Msc and PhD Degrees from Karadeniz Technical University, Department of Mechanical Engineering in 2005 and 2010 respectively. She studied at University of Gent, Belgium in 2005-2006 during her PhD. She worked as a visiting Professor at the Technology University of Delft, Netherlands in 2014-2015 with TUBITAK's 2219-scholarship. She worked on the" 4D-EEG project" which is an Advanced Grant from the European Research Council, funded under the Seventh Framework Programme (FP7).ERC n. 291339-4D-EEG. Her research topics include robotics, biomedical signal processing and control applications. She is fluent in English with speaking and writing.*

Prof. Gümüşel *is at Department of Mechanical Engineering of Karadeniz Technical University. He received his Msc, and Ph.D degrees from The Catholic University of America, Washington DC., ABD in 1986 and 1990 respectively. His research interests include robotics, mechatronics, and dynamics of machinery, brain computer interfaces, and automatic control of systems. He has several papers on robotics and control systems. He works on different scientific projects. He is fluent in English.*

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.228-236 (2018)
http://doi.org/10.11121/ijocta.01.2018.00610

RESEARCH ARTICLE

Hermite collocation method for fractional order differential equations

Nilay Akgönüllü Pirim and Fatma Ayaz*

Department of Mathematics, Gazi University, Ankara, Turkey
nilayakgonullu@gmail.com, fayaz@gazi.edu.tr

ARTICLE INFO

Article History:
Received 03 May 2018
Accepted 13 July 2018
Available 20 July 2018

Keywords:
Fractional order differential equations
Hermite polynomials
Collocation methods

AMS Classification 2010:
26A33,33C45,35C10

ABSTRACT

This paper focuses on the approximate solutions of the higher order fractional differential equations with multi terms by the help of Hermite Collocation method (HCM). This new method is an adaptation of Taylor's collocation method in terms of truncated Hermite Series. With this method, the differential equation is transformed into an algebraic equation and the unknowns of the equation are the coefficients of the Hermite series solution of the problem. This method appears as an useful tool for solving fractional differential equations with variable coefficients. To show the pertinent feature of the proposed method, we test the accuracy of the method with some illustrative examples and check the error bounds for numerical calculations.

1. Introduction

In many branches of science, mathematical models of physical processes require differential equations and nowadays, it is verified that some of these models can be better defined by fractional order equations due to the material and hereditary properties. Consequently, too many applications of the fractional order differential equations exist (see [1–5]). Unfortunately, model equations are usually in complex nature and involves nonlinear terms, therefore, analytical solutions can not easily be obtained. As a result, we still need more powerful numerical or approximate methods. Nowadays, many researchers are studying on numerical or approximate solutions of the fractional order equations and some new techniques have been introduced or adopted with the existent ones for ordinary case. For instance, finite difference [6–10], fractional linear multistep methods [11–13], Adomian decomposition [14–16], variational iteration method [16–18], differential transform or Taylor collocation method [19, 20] and spectral method [21–24] can be cited here. For some classes of fractional differential equations, Kumar and Agarwal mentioned about polynomial approximation methods and detailed information can be found in [25–27]. There are also some other studies which worth to cite here [28–30]. These are some valuable studies on fractional partial differential equations. In the recent years, many works have also been published on solving fractional differential equations but most of them have been concerned with a single term and the order is less than one. However, here, we adopt the Hermite Collocation method (HCM) for obtaining solutions to higher order multi-term fractional differential equations with variable coefficients. This technique evaluates an analytical solution in the form of a truncated Hermite series with unknown coefficients. In many physical problems, orthogonal functions or polynomials are used as a basis for obtaining solutions to the problems. On the other hand, the orthogonal Hermite polynomials are extensively

*Corresponding Author

used in some problems of hydrodynamics and meteorology [31]. The present method uses Hermite polynomials smilar to the Taylor collocation method and so is called Hermite Collocation Method which was first developed for higher-order linear Fredholm integro differential equation [32]. Using this method has advantages on some partical types of physical processes as we mentioned above.

The second section of this study involves preliminary definitions and related theorems of the fractional calculus. In section 3, we recall the fundamental properties of the Hermite series and making adaptation of the method to the fractional order equation. Section 4 deals with the error bounds for the calculations and the section 5 involves some illustrative examples. Finally, we conclude the research with some highlights.

2. Preliminary information and notations

We start with the definition of Caputo derivative which was first introduced by Caputo ([33]). This derivative is preferred by many researches to make it easier to incorporate the initial and boundary conditions to the problem. Therefore, all the derivatives will be defined as Caputo derivatives throughout this study.

Definition 1. *[1] Let $f \in AC^n[a,b]$ then, the Caputo fractional derivative of a function f of order $\alpha > 0$ is defined by*

$$({}^{C}D_a^{\alpha} f)(x) = \frac{1}{\Gamma(n-\alpha)} \int_a^x (x-t)^{n-\alpha-1} (\frac{d}{dt})^n f(t) dt, \tag{1}$$

where Γ is the gamma function and $n-1 < \alpha < n, n \in N$.

Additionally, we can state the Caputo fractional derivative of a power function as follows. Let $\alpha \geq 0$, and $f(x) = (x-c)^{\beta}$ for some $\beta \geq 0$. Letting c is any number, then

$${}^{C}D_a^{\alpha} f(x) = \begin{cases} 0, & if\ \beta \in \{0,1,2,...,n-1\} \text{ and } \beta < n, \\ \frac{\Gamma(\beta+1)}{\Gamma(\beta+1-\alpha)}(x-c)^{\beta-\alpha}, & \beta \in N \text{ and } \beta \geq n \text{ or}, \\ & \beta \notin N \text{ and } \beta .> n-1 \end{cases} \tag{2}$$

Some properties of the Caputo derivative can be given as follows:

Lemma 1. *[1] Let $\alpha > 0$ and let $y \in L^{\infty}(a,b)$ or $C([a,b])$. Then, $({}^{C}D_a^{\alpha} I^{\alpha} y)(x) = y(x)$, where I^{α} defines the integral operator.*

Lemma 2. *[1] Let $\alpha > 0$ and $n = [\alpha] + 1$ where $[\alpha]$ is the integer part of α. If $y \in AC^n([a,b])$ or $y \in C^n([a,b])$, then*

$(I^{\alpha C} D_a^{\alpha} y)(x) = y(x) - \sum_{k=0}^{n-1} \frac{y^{(k)}(a)}{k!}(x-a)^k$.

Theorem 1. *[34] For every $\alpha, \beta \in \mathbb{R}_+$ the following relation holds,*

$${}^{C}D_a^{\alpha C} D_a^{\beta} f(x) = {}^{C} D_a^{\alpha+\beta} f(x).$$

3. Hermite-collocation method for fractional order differential equations

This section deals with the establishment of the theory of HCM for solving following multi-term fractional differential equations with variable coefficients,

$$\sum_{k=0}^{m} P_k(x) {}^{C}D_a^{k\alpha} y(x) = g(x), \tag{3}$$

where $a \leq x \leq b, n-1 < m\alpha < n$ $(0 < \alpha < 1)$, $n > 1, n \epsilon \mathbf{N}$, $P_k(x)$ and $g(x)$ continuous on $a \leq x \leq b$. Initial conditions are:

$${}^{C}D^{j} y(a) = \lambda_j,\ j = 0,1,2,...,m\alpha - 1. \tag{4}$$

In Eqs.(3)-(4), D_a^{α} or, for convenience D^{α} defines the Caputo derivative of order α and , $m\alpha - 1$ is an integer number. We approximate the solution of the form as the following truncated Hermite series,

$$y(x) = \sum_{k=0}^{N} a_k H_k(x^{\alpha}), \tag{5}$$

where a_k are unknown Hermite coefficients and $N \in \mathbf{N}^+$ which satisfies $N \geq m\alpha$. To obtain the solution of Eq.(3) of the form Eq.(5), we first define the collocation points as $x_i = a + (\frac{b-a}{N})i$ ($i = 0,1,2,...N$, and $x_0 = a, x_N = b$). Now, letting that $H(x^{\alpha}) = [H_0(x^{\alpha})\ H_1(x^{\alpha})\ H_2(x^{\alpha})...H_N(x^{\alpha})]$, $A = [a_0\ a_1\ a_2...a_N]^T$ then, Eq.(5) is written in matrix form as follow:

$$[y(x)] = H(x^{\alpha})A. \tag{6}$$

Eventually, at collocation points, Eq.(6) is shown by $[y(x_i)] = H(x_i^\alpha)A$.

3.1. Fractional hermite collocation method

The Hermite polynomials of degree n are generated by the very well known formula,

$$H_n(x) = \sum_{n=0}^{\left[\left\lfloor \frac{n}{2} \right\rfloor\right]} (-1)^m \frac{n!}{(n-2m)!m!} x^{n-2m} \quad -\infty < x < \infty \ , \ n = 0, 1, 2..., N.$$

Now we can define them in matrix notation (see [32])as below. If N is an odd number, then the matrix notation of Hermite polynomials is written as

$$\underbrace{\begin{bmatrix} H_0(x^\alpha) \\ H_1(x^\alpha) \\ \vdots \\ H_{N-1}(x^\alpha) \\ H_N(x^\alpha) \end{bmatrix}}_{H^T(x^\alpha)} = \underbrace{\begin{bmatrix} 2^0 & 0 & \dots & 0 & 0 \\ 0 & 2^1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (-1)^{\left(\frac{N-5}{2}\right)} \frac{2^0}{0!} \frac{(N-1)!}{\left(\frac{N-1}{2}\right)!} & 0 & \dots & 2^{N-1} & 0 \\ 0 & (-1)^{\left(\frac{N-1}{2}\right)} \frac{2^1}{1!} \frac{N!}{\left(\frac{N-1}{2}\right)!} & \dots & 0 & 2^N \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} 1 \\ x^\alpha \\ \vdots \\ x^{\alpha(N-1)} \\ x^{\alpha N} \end{bmatrix}}_{X^T(x^\alpha)}, \tag{7}$$

if N is even then it follows,

$$\underbrace{\begin{bmatrix} H_0(x^\alpha) \\ H_1(x^\alpha) \\ \vdots \\ H_{N-1}(x^\alpha) \\ H_N(x^\alpha) \end{bmatrix}}_{H^T(x^\alpha)} = \underbrace{\begin{bmatrix} 2^0 & 0 & \dots & 0 & 0 \\ 0 & 2^1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & (-1)^{\left(\frac{N-2}{2}\right)} \frac{2^1}{1!} \frac{(N-1)!}{\left(\frac{N-2}{2}\right)!} & \dots & 2^{N-1} & 0 \\ (-1)^{\left(\frac{N-4}{2}\right)} \frac{2^0}{0!} \frac{N!}{\left(\frac{N}{2}\right)!} & 0 & \dots & 0 & 2^N \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} 1 \\ x^\alpha \\ \vdots \\ x^{\alpha(N-1)} \\ x^{\alpha N} \end{bmatrix}}_{X^T(x^\alpha)}. \tag{8}$$

Consequently, we can write the above matrices shortly,

$$H^T(x^\alpha) = FX^T(x^\alpha),$$

or

$$H(x^\alpha) = X(x^\alpha)F^T. \tag{9}$$

More generally, if we show that

$$X(x^\alpha) = \left[(x-c)^0 (x-c)^{1\alpha} \cdots (x-c)^{(N-1)\alpha} (x-c)^{N\alpha}\right],$$

then, the substitution of Eq.(9) into Eq.(6) yields,

$$y(x) = X(x^\alpha)F^T A. \tag{10}$$

3.2. Caputo derivatives of operational matrix

Now, we need to determine any $k\alpha$ th order Caputo fractional derivatives of Eq.(10) by the following procedure,

$${}^C D^{k\alpha} y(x) = {}^C D^{k\alpha} X(x^\alpha) F^T A. \tag{11}$$

$${}^C D^{k\alpha} X(x^\alpha) = [{}^C D^{k\alpha}(x-c)^0 \ {}^C D^{k\alpha}(x-c)^{1\alpha} \dots {}^C D^{k\alpha}(x-c)^{(N-1)\alpha} \ {}^C D^{k\alpha}(x-c)^{N\alpha}] \tag{12}$$

$$\underbrace{\begin{bmatrix} {}^C D^\alpha (x-c)^0 \\ {}^C D^\alpha (x-c)^{1\alpha} \\ {}^C D^\alpha (x-c)^{2\alpha} \\ \vdots \\ {}^C D^\alpha (x-c)^{(N-1)\alpha} \\ {}^C D^\alpha (x-c)^{N\alpha} \end{bmatrix}}_{{}^C D^{k\alpha} X^T(x^\alpha)} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \Gamma(\alpha+1) & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{\Gamma(2\alpha+1)}{\Gamma(\alpha+1)} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{\Gamma(N\alpha+1)}{\Gamma((N-1)\alpha+1)} & 0 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} (x-c)^0 \\ (x-c)^{1\alpha} \\ (x-c)^{2\alpha} \\ \vdots \\ (x-c)^{(N-1)\alpha} \\ (x-c)^{N\alpha} \end{bmatrix}}_{X^T(x^\alpha)}$$

Consequently, one can write that ${}^{C}D^{\alpha}X(x^{\alpha}) = X(x^{\alpha})B^{T}$ and the following theorem holds.

Theorem 2. *Let $X(x^{\alpha})$ be the Hermite polynomial vector, for any $\alpha > 0$, then we have,*

$$ {}^{C}D^{k\alpha}y(x) = X(x^{\alpha})(B^{T})^{k}F^{T}A, $$

Proof. By the help of Theorem 1, the successive α th order Caputo fractional derivatives of $X(x^{\alpha})$ become,

$$ \begin{aligned} {}^{C}D^{\alpha C}D^{\alpha}X(x^{\alpha}) &= \underbrace{{}^{C}D^{\alpha}X(x^{\alpha})}_{X(x^{\alpha})B^{T}}B^{T}, \\ {}^{C}D^{2\alpha}X(x^{\alpha}) &= X(x^{\alpha})(B^{T})^{2}, \\ &\vdots, \\ {}^{C}D^{k\alpha}X(x^{\alpha}) &= X(x^{\alpha})(B^{T})^{k}. \end{aligned} \tag{13} $$

Hence, substitution of Eq.(13) into Eq.(11) gives,

$$ {}^{C}D^{k\alpha}y(x) = X(x^{\alpha})(B^{T})^{k}F^{T}A. \tag{14} $$

Eq.(14), is also shown by the following formula at collocation points $x = x_i$ as,

$$ {}^{C}D^{k\alpha}y(x_i) = X(x_i^{\alpha})(B^{T})^{k}F^{T}A. \tag{15} $$

Now, let recall the differential equation redefined at collocation points as same as below,

$$ \sum_{k=0}^{m} P_k(x_i)\,{}^{C}D^{k\alpha}y(x_i) = g(x_i), \quad i = 0, 1, 2, \ldots, N, \tag{16} $$

therefore, Eq.(16) is written in the following matrix form:

$$ \underbrace{\begin{bmatrix} P_k(x_0) & 0 & \cdots & 0 \\ 0 & P_k(x_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_k(x_N) \end{bmatrix}}_{P_k} \times \underbrace{\begin{bmatrix} {}^{C}D^{k\alpha}y(x_0) \\ {}^{C}D^{k\alpha}y(x_1) \\ \vdots \\ {}^{C}D^{k\alpha}y(x_N) \end{bmatrix}}_{Y^{k\alpha}} = \underbrace{\begin{bmatrix} g(x_0) \\ g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix}}_{G}. $$

In the compact form, Eq.(16) can be given as,

$$ \sum_{k=0}^{m} P_k Y^{k\alpha} = G. \tag{17} $$

On the other hand, we have by Eq.(15),

$$ \underbrace{\begin{bmatrix} {}^{C}D^{k\alpha}y(x_0) \\ {}^{C}D^{k\alpha}y(x_1) \\ \vdots \\ {}^{C}D^{k\alpha}y(x_N) \end{bmatrix}}_{Y^{k\alpha}} = \underbrace{\begin{bmatrix} X(x_0^{\alpha}) \\ X(x_1^{\alpha}) \\ \vdots \\ X(x_N^{\alpha}) \end{bmatrix}}_{X^{\alpha}} \left[(B^{T})^{k}F^{T}A\right], $$

where the matrix X^{α} is equivalent to

$$ X^{\alpha} = \underbrace{\begin{bmatrix} X(x_0^{\alpha}) \\ X(x_1^{\alpha}) \\ \vdots \\ X(x_N^{\alpha}) \end{bmatrix}}_{X^{\alpha}} = \begin{bmatrix} 1 & (x_0 - c)^{1\alpha} & \cdots & (x_0 - c)^{(N-1)\alpha} & (x_0 - c)^{N\alpha} \\ 1 & (x_1 - c)^{1\alpha} & \cdots & (x_1 - c)^{(N-1)\alpha} & (x_1 - c)^{N\alpha} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & (x_N - c)^{1\alpha} & \cdots & (x_N - c)^{(N-1)\alpha} & (x_N - c)^{N\alpha} \end{bmatrix}. $$

Hence, we can rewrite Eq.(15) as follows,

$$ Y^{k\alpha} = X^{\alpha}(B^{T})^{k}F^{T}A. \tag{18} $$

Finally, the substitution Eq.(18) into Eq. (17) gives the fundamental matrix equation such as

$$ \sum_{k=0}^{m} P_k X^{\alpha}(B^{T})^{k}F^{T}A = G. \tag{19} $$

Moreover, denoting

$$ W = \sum_{k=0}^{m} P_k(x) X^{\alpha}(B^{T})^{k}F^{T}, $$

where $W = [w_{ij}]$ ($i, j = 0, 1, 2, \ldots N$), then, Eq. (19) is shown by,

$$ W.A = G. \tag{20} $$

Now, Eq. (20) generates an algebraic system which consists of (N+1) rows and (N+1) columns. Then, the augmented matrix of the system is written by,

$$[W;G] = \left[\begin{array}{ccccc;c} w_{00} & w_{01} & \cdots & w_{0N} & ; & g(x_0) \\ w_{10} & w_{11} & \cdots & w_{1N} & ; & g(x_1) \\ \vdots & \vdots & \vdots & \vdots & ; & \vdots \\ w_{(N-1)0} & w_{(N-1)1} & \cdots & w_{(N-1)N} & ; & g(x_{N-1}) \\ w_{N0} & w_{N1} & \cdots & w_{NN} & ; & g(x_N) \end{array}\right] \quad (21)$$

This method can be modified to handle the initial conditions defined at particular point a. Therefore, we recall the initial Eq. (4),

$$^{C}D^{j}y(a) = \lambda_j, j = 0, 1, 2, \ldots m\alpha - 1.$$

Hence, substitution of these conditions into Eq.(15) yields,

$$X^{\alpha}(a)(B^T)^j F^T A = \lambda_j. \quad (22)$$

Therefore, defining U_j as,

$$U_j = X^{\alpha}(a)(B^T)^j F^T = \left[\begin{array}{cccc} u_{j0} & u_{j1} & u_{j2} & \cdots \end{array}\right. \left.\begin{array}{c} u_{jN} \end{array}\right]$$

then, Eq.(22) can be shown by

$$U_j A = \lambda_j, \quad (23)$$

and corresponding augmented matrix is written of the form,

$$[U_j; \lambda_j], j = 0, 1, 2, \ldots, m\alpha - 1,$$

and denoted by

$$[U_j;\lambda_j] = \left[\begin{array}{ccccc;c} u_{00} & u_{01} & \cdots & u_{0N} & ; & \lambda_0 \\ u_{10} & u_{01} & \cdots & u_{1N} & ; & \lambda_1 \\ \vdots & \vdots & \vdots & \vdots & ; & \vdots \\ u_{(m\alpha-1)0} & u_{(m\alpha-1)1} & \cdots & u_{(m\alpha-1)N} & ; & \lambda_{m\alpha-1} \end{array}\right]. \quad (24)$$

Now, if the $m\alpha$th row of the the augmented matrix Eq.(21)of the system is replaced by the augmented matrix of initial conditions Eq.(24), then one can write the following matrix form,

$$\left[\widetilde{W};\widetilde{G}\right] = \left[\begin{array}{ccccc;c} w_{00} & w_{01} & \cdots & w_{0N} & ; & g(x_0) \\ w_{10} & w_{01} & \cdots & w_{1N} & ; & g(x_1) \\ \vdots & \vdots & \vdots & \vdots & ; & \vdots \\ w_{(N-1-m\alpha)0} & w_{(N-1-m\alpha)1} & \cdots & w_{(N-1-m\alpha)N} & ; & g(x_{N-1-m\alpha}) \\ w_{(N-m\alpha)0} & w_{(N-m\alpha)1} & \cdots & w_{(N-m\alpha)N} & ; & g(x_{N-m\alpha}) \\ u_{00} & u_{01} & \cdots & u_{0N} & ; & \lambda_0 \\ u_{10} & u_{11} & \cdots & u_{1N} & ; & \lambda_1 \\ \vdots & \vdots & \vdots & \vdots & ; & \vdots \\ u_{(m\alpha-1)0} & u_{(m\alpha-1)1} & \cdots & u_{(m\alpha-1)N} & ; & \lambda_{m\alpha-1} \end{array}\right]$$

Hence, the system of algebraic equations are shown by the following notation,

$$\widetilde{W}A = \widetilde{G}. \quad (25)$$

□

Remark 1. *Now, let us consider the system* $\widetilde{W}A = \widetilde{G}$.

If $rank\widetilde{W} = rank\,[\widetilde{W}, \widetilde{G}] = N+1$, (i.e $\det(\widetilde{W}) \neq 0$) *then we can write*

$$A = (\widetilde{W})^{-1}\widetilde{G}. \quad (26)$$

Consequently, the Hermite coefficients a_k $(k = 0, 1, 2, ..., N)$ can be uniquely determined by Eq.(25). As a result, the truncated Hermite series is written as follows,

$$y(x) = \sum_{k=0}^{N} a_k H_k(x^{\alpha}). \quad (27)$$

4. Error bounds

Eq.(27) is the approximate solution to Eq.(3) with the initial conditions, Eq.(4). Therefore, substitution the truncated Hermite series into the problem, we obtain the residuals;

$$\left|\sum_{k=0}^{m} P_k(x_i)\,^{C}D^{k\alpha}y(x_i) - g(x_i)\right|$$

at $x = x_i$ $(-\infty < a \leq x \leq b < \infty)$, $i = 0, 1, 2, \ldots, N$. Then, we call the error function as $E(x_i)$ and this function should be less than ϵ, which is a positive number and can arbitrarily

be chosen as $10^{-k_i\alpha}$. As a result, error function becomes, $E(x_i) \leq 10^{-k_i\alpha}$ where $k_i > 0$ is any constant. If the $\max(10^{-k_i\alpha}) = 10^{-k\alpha}$ is desired accuracy then, the truncation limit N is incerased untill $E(x_i)$ approaches zero. Besides, the global error function is defined as follows,

$$E_N(x) = \sum_{k=0}^{m} P_k(x)\,{}^C D^{k\alpha} y(x) - g(x).$$

Consequently, the global error, $E_N(x) \to 0$ when N is sufficiently large.

5. Illustrative examples

The method which was mentioned so far has been used to solve multi term fractional order differential equations. To show the accuracy of the method, the following examples have been solved. All the numerical calculations have been performed by using Matlab v7.5. and the results have been given by Figure 1 for different values of α. The comparisons between exact and Hermite polynomial solution approximation have been made and shown by Table 1.

Example 1. *First we consider Bagley-Torvik equation [35];*

$$D^2 y(x) + D^{3/2} y(x) + y(x) = x + 1, \qquad (28)$$

where $\alpha = 1/2, m = 4$ and $g(x) = x + 1$. To find HCM solution of the problem here, for convenience, we choose $N = 2$. Because the analytical solution of the problem for $\alpha = 1$ is easily obtainable. Therefore, we only concentrate on the different values of α. Hence,the approximate solution can be written by the following truncated Hermite series:

$$y(x) = \sum_{n=0}^{2} a_n H_n(x^\alpha). \qquad (29)$$

The coefficients of the differential equation are $P_0(x) = P_3(x) = P_4(x) = 1, P_1(x) = P_2(x) = 0$ Since $N = 2$ then, collocation points are taken as $\{x_0 = 0, x_1 = 1/2, x_2 = 1\}$. From the fundamental matrix equation Eq.(19),one can write that

$$\{P_0 X + P_3 X (B^T)^3 + P_4 X (B^T)^4\} F^T A = G.$$

After evaluating the matrices B and F and substituting them into the above equation, then, the augmented matrix is obtained as follows,

$$W.A = G \Rightarrow [W;G] = \left[\begin{array}{ccc;c} 1 & 0 & -2 & 1 \\ 1 & 1.4 & 0 & 1.5 \\ 1 & 2 & 2 & 2 \end{array}\right]$$

Since $\det(W) \neq 0$, *finding the solution of the system defines the coefficients of the truncated series as*

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0 \\ 0.25 \end{bmatrix}$$

Finally, substituting these coefficients into Eq.(29),

$$y(x) = a_0 H_0(x^\alpha) + a_1 H_1(x^\alpha) + a_2 H_2(x^\alpha)$$

then, we obtain Hermite polynomial solution of the problem as

$$y(x) = \frac{3}{2} + \frac{1}{4}(4x^\alpha - 2).$$

From here, if we substitute $\alpha = 1$ in $y(x)$ then, we obtain $y(x) = x + 1$. This is the exact solution of the problem for integer order case.

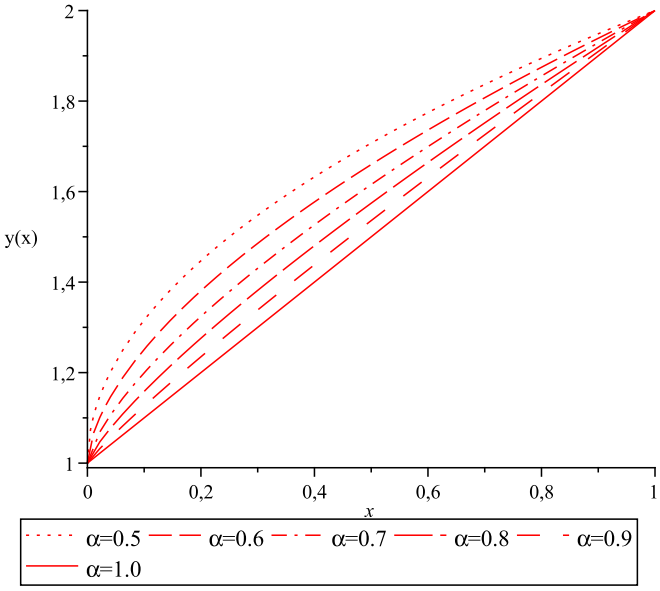


Figure 1. The solution curves of Example 1 for different values of α. The α values change from 0.5 to 1.

Example 2. *Next we consider an initial value problem which is studied in [36] where $0 \leq x \leq 1$ and $\alpha \in (0,1), \beta \leq 0$. Therefore, we can write the equation as:*

$$D^{\alpha}y(x) = \beta y(x) + g(x). \qquad (30)$$

We assume that $\beta = -1$ *and* $g(x) = x^2 + \frac{2x^{2-\alpha}}{\Gamma(3-\alpha)}$ *defined on [0,1] and the initial condition is* $y(0) = 0$. *If we apply the HCM for* $\alpha = \frac{1}{2}$ *and* $N = 4$, *it is obvious that* m *is 2 and* $P_0(x) = P_1(x) = 1, P_2(x) = 0$. *Therefore, we can write the following fundamental matrix of Eq.(30) as,*

$$\left\{P_0X + P_1XB^T + P_2X(B^T)^2\right\}F^TA = G.$$

Therefore, we can easily establish the system $WA = G$ *and constitute* $[W;G]$ *matrix as,*

$$[W;G] = \begin{bmatrix} 1.0000 & 1.7725 & -2.0000 & -10.6347 & 12.0000 & ; & 0.0000 \\ 1.0000 & 2.7725 & 1.2568 & -12.9760 & -23.0721 & ; & 0.2506 \\ 1.0000 & 3.1867 & 3.1915 & -10.9742 & -37.7877 & ; & 0.7818 \\ 1.0000 & 3.5045 & 4.9088 & -7.8548 & -46.2706 & ; & 1.5397 \\ 1.0000 & 3.7725 & 6.5135 & -4.0000 & -50.0901 & ; & 0.0000 \end{bmatrix}.$$

On the other hand, the augmented matrix, which corresponds to initial condition, is obtained by substitution the row,

$$y(0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}F^TA = 0.$$

$$y(0) = X(\mathbf{0})F^TA = \lambda_0 = 0,$$

or

Finally, we find that $[U_0;\lambda_0] = \begin{bmatrix} 1 & 0 & -2 & 0 & 12 & ; & 0 \end{bmatrix}$. Therefore, the augmented matrix of the system, $\widetilde{W}A = \widetilde{G}$, is obtained from Eq.(25) as follows,

$$\left[\widetilde{W};\widetilde{G}\right] = \begin{bmatrix} 1.0000 & 1.7725 & -2.0000 & -10.6347 & 12.0000 & ; & 0.0000 \\ 1.0000 & 2.7725 & 1.2568 & -12.9760 & -23.0721 & ; & 0.2506 \\ 1.0000 & 3.1867 & 3.1915 & -10.9742 & -37.7877 & ; & 0.7818 \\ 1.0000 & 3.5045 & 4.9088 & -7.8548 & -46.2706 & ; & 1.5397 \\ 1.0000 & 0.0000 & -2.0000 & 0.0000 & 12.0000 & ; & 0.0000 \end{bmatrix}$$

Consequently, solution of the above system gives the approximate solution of the problem as,

$$y(x) = 0,986076x10^{-31}x^{1/2} + 0,104468x10^{-13}x^{3/2} + x^2.$$

We note here that our solution is very close to the exact solution, $y(x) = x^2$, since the first two terms vanishes (see Figure 2).

As a result of all these, the HCM solution is very good approximation to the problem even for small N. Table 1 lists both the exact solution and error function at particular x, corresponding to the Example.

Table 1. The exact solution of Example 2 and error function $E(x_i)$ for HCM solution.

x values	Exact Solution	$E(x_i)$ For HCM solution, $N = 4$
0.0	0.0000	9.0206e-17
0.2	4.0000e-2	9.0206e-17
0.4	1.6000e-1	2.7756e-16
0.6	3.6000e-1	4.9960e-16
0.8	6.4000e-1	7.7716e-16
1.0	1.0000	1.1102e-15

In Figure 2, both the analytical and the HCM solutions (for $N = 2$ and $N = 4$) are given. It is clear that the series (for $N = 4$) and the analytical solution are identical.

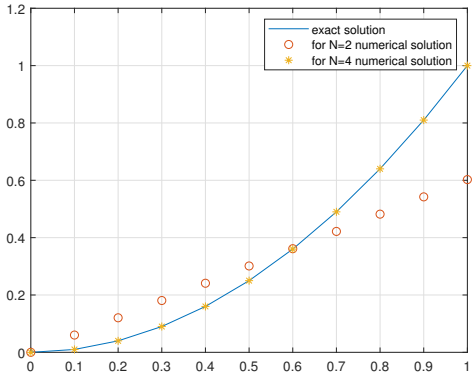


Figure 2. Comparison of Analytical solution and HCM solution for Example 2.

6. Conclusion

The objective of this study is to apply the HCM method for solving higher order multi-term fractional order differential equations. The motivation of this work is that obtaining considerable simplifications in the solutions of the multi-term fractional order differential equations by using HCM, since the analytical solutions of such equations cannot easily be obtained. By using any symbolic toolbox of Matlab programme, the Hermite polynomial coefficients of the solution can be obtained easily. Illustrated examples determine the reliability of the algorithm and give chance to apply the method for wider classes of equations. As a further work, the method will be considered for solving nonlinear fractional order differential equations.

References

[1] Kilbas, A.A., Sirvastava, H.M., Trujillo, J.J. (2006). *Theory and Application of Fractional Differential Equations*. North-Holland Mathematics Studies. Vol 204, Amsterdam.

[2] Mainardi, F., Luchko, Yu., Pagnini, G. (2001). The fundamental Solution of the Space-Time Fractional Diffusion Equation. *Frac. Calc. Appl. Anal.*, 4(2), 153-152.

[3] Ross, B. (1975). *Fractional Calculus and its Applications*. Lecture Notes in Mathematics. Vol. 457, Springer Verlag.

[4] Stanislavsky, A.A. (2004) Fractional Oscillator. *Phys. Rev. E.*, 70(5).

[5] Trujillo, J.J. (1999). On a Riemann-Liouville Generalised Taylor's Formula. *J. Math. Anal. Appl.*, 231, 255-265.

[6] Podlubny, I. (1999). *Fractional Differential Equations*. Academic Press, New York.

[7] Ciesielski, M., Leszcynski, J. (2003). Numerical simulations of anomalous diffusion. Computer Methods Mech. Conference, Gliwice Wisla Poland.

[8] Yuste, S.B. (2006). Weighted average finite difference methods for fractional diffusion equations. *J. Comput. Phys.*, 1, 264-274.

[9] Odibat, Z.M. (2006). Approximations of fractional integrals and Caputo derivatives. *Appl. Math. Comput.*, 527-533.

[10] Odibat, Z.M. (2009). Computational algorithms for computing the fractional derivatives of functions. *Math. Comput. Simul.*, 79(7), 2013-2020.

[11] Ford, N.J., Connolly, A.Joseph. (2009). Systems-based decomposition schemes for the approximate solution of multi-term fractional differential equations. *J. Comput. Appl. Math.*, 229, 382-391.

[12] Ford, N.J. (2001). Simpson, A. Charles, The numerical solution of fractional differential equations:speed versus accuracy. *Numer. Alg.*, 26, 333-346.

[13] Sweilam, N.H., Khader, M.M., Al-Bar, R.F. (2007). Numerical studies for a multi order fractional differential equations. *Phys. Lett. A*, 371, 26-33.

[14] Momani, S., Odibat, Z. (2006). Analytical solution of a time-fractional Navier-Stokes equation by adomian decomposition method. *Appl. Math. Comput.*, 177, 488-494.

[15] Odibat, Z., Momani, S. (2007). Numerical approach to differential equations of fractional order. *J. Comput. Appl. Math.*, 207(1), 96-110.

[16] Odibat, Z., Momani, S. (2008). Numerical methods for nonlinear partial differential equations of fractional order. *Appl. Math. Model.*, 32, 28-39.

[17] Odibat, Z., Momani, S. (2006). Application of variational iteration method to equation of fractional order. *Int. J. Nonlinear Sci. Numer. Simul.*, 7, 271-279.

[18] Wu, J.L. (2009). A wavelet operational method for solving fractional partial differential equations numerically. *Appl. Math. Comput.*, 214, 31-40.

[19] Ertürk,V.S., Momani, S. (2008). Solving systems of fractional differential equations using differential transform method. *J. Comput. Appl. Math.*, 215, 142-151.

[20] Yalçınbaş, S., Konuralp, A., Demir, D.D., Sorkun, H.H. (2010). The solution of the fractional differential equation with the generalized Taylor collocation method. *IJRRAS*, August 2010.

[21] Abd-Elhameed, W.M., Youssri, Y.H. (2016). A Novel Operational Matrix of Caputo Fractional Derivatives of Fibonacci Polynomials: Spectral Solutions of Fractional Differential Equations. *Entropy*, 18(10), 345.

[22] Abd-Elhameed, W.M., Youssri, Y.H. (2016). Spectral Solutions for Fractional Differential Equations via a Novel Lucas Operational Matrix of Fractional Derivatives. *Romanian Journal of Physics*, 61(5-6), 795-813.

[23] Abd-Elhameed, W.M., Youssri, Y.H., (2015). New Spectral Solutions of Multi-Term Fractional-Order Initial Value Problems With Error Analysis. *CMES: Computer Modeling in Engineering and Sciences*, 105(5), 375-398.

[24] Youssri, H., Abd-Elhameed W.M. (2016). Spectral Solutions for Multi-Term Fractional Initial Value Problems Using a New Fibonacci Operational Matrix of Fractional Integration. *Progress in Fractional Differentiation and Applications*, 2(2), 141-151.

[25] Kumar, P., Agrawal, O.P. (2006). An approximate method for numerical solution of fractional differential equations. *Signal Process*, 86, 2602-2610.

[26] Kumar, P. (2006). *New numerical schemes for the solution of fractional differential equations*. Southern Illinois University at Carbondale. Vol 134.
[27] Kumar, P., Agrawal, O.P. (2006). Numerical scheme for the solution of fractional differential equations of order greater than one. *J. Comput. Nonlinear Dyn.*, 1, 178-185.
[28] Yavuz M., Özdemir, N. (2018). A Different Approach to the European Option Pricing Model with New Fractional Operator. *Mathematical Modelling of Natural Phenomena*, 13(1),1-12.
[29] Avci, D., Eroglu İskender, B.B, Özdemir, N. (2017). Conformable heat equation on a radial symmetric plate. *Thermal Science*, 21(2), 819-826.
[30] Yavuz, M., Özdemir, N. (2018). Numerical inverse Laplace homotopy technique for fractional heat equations. *Thermal Science*, 22(1), 185-194.
[31] Dattoli, G. (2004). Laguerra and generalized Hermite polynomials: the point of view of the operation method. *Integral Transforms Special Func.*, 15, 93-99.
[32] Akgönüllü, N.,Şahin, N., Sezer, M. (2010). A Hermite Collocation Method for the Approximate Solutions of Higher-Order Linear Fredholm Integro-Differential Equations. *Numerical Methods for Partial Differential Eqautions*, 27(6), 1708-1721.
[33] Caputo, M. (1967). Linear models of dissipation whose Q is almost frequency independent II. Geophys. *J. Royal Astron. Soc.*, 13, 529-539.
[34] Samko,S., Kilbas, A., Marichev, O. (1993). *Fractional integral and derivatives theory and applications*. Gordon and Breach, New York.
[35] Diethelm, K., Ford, N.J. (2002). Numerical Solution of the Bagley-Torvik Equation. *BIT*, 42, 490-507.
[36] Diethelm K. (1997). An algorithm for the numerical solution of differential equations of fractional order. *Electronic Transactions on Numerical Analysis*,5, 1-6.

Fatma Ayaz *is a Professor in the Department of Mathematics at Gazi University, Ankara, Turkey. She received her Ph.D. of Mathematics in Leeds University (UK), 1995. She has many research papers about numerical solutions of differential equations and fractional ordinary differential equations, computational methods, computational fluid dynamics, mathematical modelling. She has supervised 14 graduate students and has been supervising 6 graduate students.*

Nilay Akgönüllü Pirim *received her PhD. degree in Mathematics from Gazi University in 2014. Her research interests focuses mainly on the approximate solutions of fractional differential equations and Hermite collocation Method.*

An International Journal of Optimization and Control: Theories & Applications
(http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.237-249 (2018)
http://doi.org/10.11121/ijocta.01.2018.00519

RESEARCH ARTICLE

Active Unmatched Disturbance Cancellation and Estimation by State–Derivative Feedback for Plants Modeled as an LTI System

Halil Ibrahim Basturk

Department of Mechanical Engineering, Bogazici University, Turkey
halil.basturk@boun.edu.tr

ARTICLE INFO

Article History:
Received 27 July 2017
Accepted 09 March 2018
Available 20 July 2018

Keywords:
Adaptive Control
Estimation
Backstepping
Disturbance Cancellation

AMS Classification 2010:
93C40, 93E10, 93D21

ABSTRACT

We design adaptive algorithms for both cancellation and estimation of unknown periodic disturbance, by feedback of state–derivatives (*i.e.,* without position information for mechanical systems) for the plants which are modeled as a linear time invariant system. We consider a series of unmatched unknown sinusoidal signals as the disturbance. The first step of the design consists of the parametrization of the disturbance model and the development of observer filters. The result obtained in this step allows us to use adaptive control techniques for the solution of the problem. In order to handle the unmatched condition, a backstepping technique is employed. Since the partial measurement of the virtual inputs is not available, we design a state observer and the estimates of these signals are used in the backstepping design. Finally, the stability of the equilibrium of the adaptive closed loop system with the convergence of states is proven. As a numerical example, a two-degree of freedom system is considered and the effectiveness of the algorithms are shown.

1. Introduction

Active disturbance cancellation arises in many applications such as automotive [1] and marine systems [2–4]. The internal model principle which relies on modeling the disturbance as the output of an exosystem and duplicating its effect to the input channel, is the one of the most popular technique for the solution [5, 6].

The internal model principle differs according to the type of the considered system, uncertainties and the measured signals. In [7], the harmonic control arrays is employed for the parametrization of the sinusoidal signals. The solutions for linear systems are given in [8–11]. The plants that contain nonlinear dynamics are proposed in [12–15]. Most of disturbance cancellation problems can also be formulated as an output regulation problem. In these references, it is assumed that the measurements of the state are available for feedback. Since the accelerometer is the major sensor that are used in mechanical applications, the direct usage of the acceleration measurement in the feedback draws the attention of the researchers.

In [16], acceleration measurement is used together with velocity and position to suppress the vibration. The term state–derivative feedback is widely used in the literature for the case where only the derivative of the considered states are available for measurement. That corresponds to acceleration and velocity feedback without using position information in mechanical systems. The pole placement technique and stability results for state–derivative is given in [17] and [18]. The adaptive cancellation algorithms by state derivative feedback for matched disturbance in a known linear plant is given in [19]. The solution for the case where the unknown disturbance is unmatched and the state of the actuator dynamics is available for the measurement is given in [20].

In this note, the problem where the actuator states are not measured but its derivative is available for a control design, is considered. The design by the derivative of the actuator state is given in [21]. The main difference between [21] is the number of the state of the actuator dynamics. In the present study, an arbitrary m– dimensional general linear system is considered for the actuator dynamics whereas it is restricted by one, in [21]. We consider two cases; first case is state and disturbance estimation in the open loop, and the second case is disturbance cancellation in the closed loop case. Firstly, a parametrization of the unknown disturbance signal is performed. Observers are developed for the unknown disturbance and the unmeasured actuator state for open loop case. We show that perfect state and observer estimations are achieved. The results that are obtained in the open loop case, is used while designing an active cancellation algorithm for the control input. An adaptive control technique is employed for the design of observer based control and update laws. Finally, we prove the stability of the equilibrium of the closed-loop adaptive system and the convergence of the state with perfect disturbance estimation.

The outline of the note is given as follows. Section 2 is dedicated to state the problems in details. The observer designs, theorem that states the convergence and its proof are given in Section 3. The backstepping adaptive controller design with the statement of the stability theorem is presented in Section 4. The theorem is proven in Section 5. As a numerical example, a two–degree of freedom mechanical system is considered. The simulation results of the example are discussed in Section 6. Lastly, concluding remarks are given in Section 7.

2. Problem statement

We consider a general representation of a linear time invariant system which is given as follows

$$\dot{x}(t) = A_n x(t) + B_n \left(b_x^T p(t) + a_x^T x + \nu(t) \right), \quad (1)$$

$$\dot{p}(t) = A_m p(t) + B_m \left(b_p^T p(t) + a_p^T x + u(t) \right), \quad (2)$$

where

$$A_n = \begin{bmatrix} 0_{n-1} & I_{n-1} \\ 0 & 0_{n-1}^T \end{bmatrix}, A_m = \begin{bmatrix} 0_{m-1} & I_{m-1} \\ 0 & 0_{m-1}^T \end{bmatrix}, \quad (3)$$

$$B_n = \begin{bmatrix} 0_{n-1} \\ 1 \end{bmatrix}, \quad B_m = \begin{bmatrix} 0_{m-1} \\ 1 \end{bmatrix}, \quad (4)$$

$$a_x = [a_{11}, \cdots, a_{1n}]^T, \quad a_p = [a_{21}, \cdots, a_{2n}]^T, \quad (5)$$

$$b_x = [b_{11}, \cdots, b_{1n}]^T, \quad b_p = [b_{21}, \cdots, b_{2n}]^T, \quad (6)$$

with $0_i = [0, \ldots, 0]^T \in \mathbb{R}^i$, the states $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$, $p = [p_1, \ldots, p_m]^T \in \mathbb{R}^m$, input $u \in \mathbb{R}$ and disturbance signal $\nu \in \mathbb{R}$ given by

$$\nu(t) = \sum_{i=1}^{q} g_i \sin(\omega_i t + \phi_i), \quad (7)$$

where $i \neq j \Rightarrow \omega_i \neq \omega_j$, $\omega_i \in \mathbb{Q}_+$, $g_i, \phi_i \in \mathbb{R}$. In the given system, the disturbance signal and the control input are separated by m integrators. This situation makes the input $u(t)$ and $\nu(t)$ unmatched. The input of system (1) is $b_x^T p(t)$ that is called virtual input. Because $b_x^T p(t)$ is not the main control input, and it has its own dynamics given by (2). The main aim is to design a control law for $u(t)$ such that the signal $b_x^T p(t)$ cancels the effect of the disturbance while maintaining the boundedness of all signals. Since state $p(t)$ and disturbance $\nu(t)$ are not measured, observers are designed to estimate these signals.

The disturbance $\nu(t)$ that is given by (7) can be also written as the output of a linear exosystem which is given by,

$$\dot{w}(t) = S w(t) \quad (8)$$

$$\nu(t) = h^T w(t) \quad (9)$$

where $w \in \mathbb{R}^{2q}$. The eigenvalues of S depend on the unknown frequencies of disturbance signal $\nu(t)$. Since the initial condition of the exosystem, ($w(0)$) represents the uncertainty of amplitude and phase, the output vector h^T can be considered as known. Therefore, it is possible to obtain an observable (h^T, S) pair.

In order to guarantee the invertibility and to avoid any discontinuity in the design, it is assumed that the terms $a_{11} \neq 0$ and $b_{21}a_{11} - a_{21}b_{11} \neq 0$ **(Assumption 1)**. We also assume that the number of the distinct frequencies, q is known **(Assumption 2)**. The given assumptions are sufficient for observer based control design and show the boundedness of signals. However, the all modes of the exosystem assumed to be excited (*i.e.,* $w(0) \neq 0$) to show the perfect state and disturbance estimations, as well as, the convergence of the state $x(t)$ **(Assumption 3)**.

The main difference between the problem statement given in this note and [21] is the dimension of $p(t)$. In [21] $p(t)$ is restricted as a scalar signal.

However, in this note there is not any restriction on the dimension of $p(t)$ and it can be an m dimensional vector where m is an arbitrary positive integer. Thus this note provides a solution for more general systems.

3. State and disturbance observer design

This section consists of two parts; the first part explains the details of parametrization and observer filters of disturbance signal, in the second part the state observer design for the virtual control input $p(t)$ is given.

3.1. Parametrization of disturbance signal

The disturbance representation given by (8) is not suitable for an observer design since the eigenvalues of system matrix S are on the imaginary axis. The first step of observer design for disturbance is to represent it as the output of a stable LTI system. To this end, we employ the approach given in [22]. Consider the Sylvester equation

$$MS - GM = lh^T. \tag{10}$$

where $M \in \mathbb{R}^{2q\times 2q}$ is the solution of the equation. As it is discussed in [23], the solution M is unique and invertible if and only if (h^T, S) is an observable pair, (G, l) is a controllable pair and the set of eigenvalues of S and G are discrete. Therefore, by choosing $G \in \mathbb{R}^{2q\times 2q}$ as a Hurwitz matrix and vector $l \in \mathbb{R}^{2q}$ such that the pair (G, l) is controllable, we guarantee that M is unique and invertible. We transform state of the exosystem $w(t)$ to a new state $z(t)$ as follows $z = Mw$. By using (10) with the state transformation, exosystem (8)–(9) is transformed into the form

$$\dot{z}(t) = Gz(t) + l\nu(t), \tag{11}$$
$$\nu(t) = \theta^T \dot{z}(t), \tag{12}$$

where

$$\theta^T = h^T(MS)^{-1}. \tag{13}$$

It is also possible to obtain the derivative of disturbance $\nu(t)$ by differentiating (12) and substituting $\ddot{z} = G\dot{z} + l\dot{\nu}$. The parametrized form of disturbance derivative is given by

$$\dot{\nu}(t) = \beta^T \dot{z}(t), \tag{14}$$

where

$$\beta^T = \frac{1}{1-\theta^T l}\theta^T G. \tag{15}$$

It is necessary to show that the term $1-\theta^T l$, that appears in the denominator is not equal to zero. Post-multiplying (10) by $(MS)^{-1}$ and substituting (13), we obtain the following equation

$$I - l\theta^T = GMS^{-1}M^{-1}. \tag{16}$$

Calculating the determinant of both sides of (16) and employing the Sylvester's determinant theorem [24] yield

$$1 - \theta^T l = \det(GMS^{-1}M^{-1}) \tag{17}$$

Recalling that $\det(GMS^{-1}M^{-1}) = \det(G)\det(M)\det(S^{-1})\det(M^{-1})$ and $\det(M^{-1}) = \frac{1}{(\det M)}$, we obtain

$$1 - \theta^T l = \det(G)\det(S^{-1}). \tag{18}$$

Recalling the fact that the determinant of a matrix is the multiplication of its eigenvalues and noting that G and S have no eigenvalues on the origin, we show that $1 - \theta^T l \neq 0$.

The disturbance signal and its derivative are represented as the output of a known and stable LTI system whose input is itself, ans thus, unknown. It is now possible to design an observer to estimate state $z(t)$, since it is the state of exponentially stable linear time invariant system. We present the following lemma that explains the benefits of developed observer filters.

Lemma 1. *The unknown disturbance $\nu(t)$ and $\dot{\nu}(t)$ are represented in the form*

$$\nu(t) = \theta^T\xi(t) + \theta^T\delta(t), \tag{19}$$
$$\dot{\nu}(t) = \beta^T\xi(t) + \beta^T\delta(t), \tag{20}$$

where

$$\xi(t) = \eta(t) + N\dot{x}(t), \tag{21}$$
$$\dot{\eta}(t) = G\left(\eta(t) + N\dot{x}(t)\right) - N\Big(A_n\dot{x}(t) + B_n\big(b_x^T\dot{p}(t) + a_x^T\dot{x}(t)\big)\Big), \tag{22}$$

with

$$N = \frac{1}{B_n^T B_n} l B_n^T, \tag{23}$$

where

$$NB_n = l \tag{24}$$

and the estimation error

$$\delta(t) = \dot{z}(t) - \xi(t), \tag{25}$$

obeys the equation

$$\dot{\delta}(t) = G\delta(t). \tag{26}$$

Proof. Using the fact that $NA_n = 0$ from (3), (4) and (24), the equation (26) is obtained by differentiating (25) and using derivative of (1), (11) and (21). Substitution of (25) into (12) and (14) yields (19) and (20), respectively.

□

The new representation of the disturbance given in (19) and (20) provides benefits to use adaptive control method in the design. In the next section, the design for the state observer together with update law for unknown constant θ are discussed.

3.2. Reciprocal state-space representation and observer design

A standard state space form of a linear system is useful for a state or output feedback design. On the other hand, the reciprocal state–space (RSS) representation that relies on switching state and state–derivatives in the equation, provides advantages for state–derivative feedback design [16]. In this form, it is possible to employ similar techniques that are developed for state and output feedback problems.

Substituting (19) into (1), the RSS form of system (1) is written as

$$\begin{aligned} x(t) =& A_n^T\dot{x}(t) + \frac{1}{a_{11}}\bar{B}_n\Big(\dot{x}_n(t) - a_x^T A_n^T\dot{x}(t) \\ & - b_x^T p(t) - \big(\theta^T\xi(t) + \theta^T\delta(t)\big)\Big). \end{aligned} \tag{27}$$

where $\bar{B}_n = \begin{bmatrix} 1, & 0_{n-1} \end{bmatrix}^T$. Substituting (27) into (2), we obtain

$$\begin{aligned} \dot{p}(t) =& A_m p(t) + B_m\Big(\bar{b}_p^T p(t) + \bar{a}_p^T A_n^T\dot{x}(t) + \frac{a_{21}}{a_{11}} \\ & \times \dot{x}_n(t) - \frac{a_{21}}{a_{11}}\big(\theta^T\xi(t) + \theta^T\delta(t)\big) + u(t)\Big), \end{aligned} \tag{28}$$

where

$$\bar{a}_p = a_p - \frac{a_{21}}{a_{11}}a_x = [\bar{a}_1, \cdots, \bar{a}_n]^T, \tag{29}$$

$$\bar{b}_p = b_p - \frac{a_{21}}{a_{11}}b_x = \big[\bar{b}_1, \cdots, \bar{b}_n\big]^T. \tag{30}$$

Using the fact $\bar{b}_1 = \frac{a_{11}b_{21} - b_{11}a_{21}}{a_{11}}$, under Assumption 1 and 2, the RSS form of (28) is obtained as

$$\begin{aligned} p(t) =& A_m^T\dot{p}(t) + \frac{1}{\bar{b}_1}\bar{B}_m\Big(\dot{p}_m(t) - \bar{b}_p^T A_m^T\dot{p}(t) \\ & - \bar{a}_p^T A_n^T\dot{x}(t) - \frac{a_{21}}{a_{11}}\dot{x}_n(t) \\ & + \frac{a_{21}}{a_{11}}\big(\theta^T\xi(t) + \theta^T\delta(t)\big) - u(t)\Big), \end{aligned} \tag{31}$$

where $\bar{B}_m = \begin{bmatrix} 1, & 0_{m-1} \end{bmatrix}^T$. The state observer in the RSS form of (31) is designed as follows

$$\begin{aligned} \hat{p}(t) =& A_m^T\dot{p}(t) + \frac{1}{\bar{b}_1}\bar{B}_m\Big(\dot{p}_m(t) - \bar{b}_p^T A_m^T\dot{p}(t) \\ & - \bar{a}_p^T A_n^T\dot{x}(t) - \frac{a_{21}}{a_{11}}\dot{x}_n(t) + \frac{a_{21}}{a_{11}}\hat{\theta}^T\xi(t) \\ & - u(t)\Big) + c_{e_p}\Big(\dot{p}(t) - \dot{\hat{p}}(t)\Big), \end{aligned} \tag{32}$$

where $c_{e_p} > 0$ and in state space form it is written as

$$\begin{aligned} \dot{\hat{p}}(t) =& -\frac{1}{c_{e_p}}\hat{p}(t) + \frac{1}{c_{e_p}}\Big(A_m^T + c_{e_p}I_m - \frac{1}{\bar{b}_1}\bar{B}_m \\ & \times \bar{b}_p^T A_m^T\Big)\dot{p}(t) + \frac{1}{c_{e_p}\bar{b}_1}\bar{B}_m\Big(\dot{p}_m(t) - \bar{a}_p^T \\ & \times A_n^T\dot{x}(t) - \frac{a_{21}}{a_{11}}\dot{x}_n(t) + \frac{a_{21}}{a_{11}}\hat{\theta}^T\xi(t) - u(t)\Big). \end{aligned} \tag{33}$$

The signals $\hat{p}(t)$ and $\hat{\theta}(t)$ are the estimates of unmeasured signal $p(t)$ and unknown constant θ, respectively. Although the signal $\dot{p}(t)$ is measured, the signal $\dot{\hat{p}}(t)$ is also needed for representing the dynamics of the state observer and the observer error systems. Defining

$$e_p(t) = p(t) - \hat{p}(t), \tag{34}$$

$$\widetilde{\theta}(t) = \theta - \hat{\theta}(t), \tag{35}$$

and subtracting (31) from (32), we obtain the error dynamics of the state observer in the RSS form as follows,

$$e_p(t) = -c_{e_p}\dot{e}_p(t) + \frac{a_{21}}{\bar{b}_1 a_{11}}\bar{B}_m\Big(\widetilde{\theta}^T(t)\xi(t) + \theta^T\delta(t)\Big). \tag{36}$$

The update law for $\hat{\theta}$ is designed as follows

$$\dot{\hat{\theta}}(t) = \kappa\frac{a_{21}}{\bar{b}_1 a_{11}}\xi(t)\bar{B}_m^T\dot{e}_p(t) \tag{37}$$

where $\kappa > 0$. It should be noted that update law, $\dot{\hat{\theta}}(t)$, is implementable since all signals given in (37) are measured including $\dot{e}_p(t)$.

Comparing the state observer given in [21] and given in (32), it is seen that they are not similar. Since the dimension of $p(t)$ is restricted as one in [21], the observer designed can not be used for the given problem in this note.

We state the following theorem for the designed observers.

Theorem 1. *Consider the system dynamics* (2)*, filter* (21), (22)*, state observer* (33) *and the update law* (37)*, under Assumptions 1–3, for initial conditions* $p(0) \in \mathbb{R}^m, \hat{\theta}(0) \in \mathbb{R}^{2q}, e_p(0) \in \mathbb{R}^m$*, the signals* $e_p(t), \widetilde{\theta}(t), \delta(t), \nu(t) - \hat{\theta}^T(t)\xi(t)$ *converge to zero as* $t \to \infty$*.*

Before proving the theorem, we state the following lemma since it is used in the proof of theorems.

Lemma 2. *There exists* $\rho > 0$ *such that for all* $t_0 \geq 0$*, the following holds*

$$Q_p(\rho, t_0) = \int_{t_0}^{t_0+\rho}\xi(t)\xi^T(t)\,dt - \frac{1}{\rho}\int_{t_0}^{t_0+\rho}\xi(t)\,dt\int_{t_0}^{t_0+\rho}\xi^T(t)\,dt > 0. \tag{38}$$

Proof.

$$\dot{\xi}(t) = G\xi(t) + l\dot{\nu}(t), \tag{39}$$

where $\dot{\nu}(t) = \sum_{i=1}^{q}\omega_i g_i\cos(\omega_i t + \phi_i)$. The proof for the signal given by (40) is given in [19]. Using (26), time derivative of (25) yields

$$\dot{\xi}(t) = G\xi(t) + l\dot{\nu}(t), \tag{40}$$

where $\dot{\nu}(t) = \sum_{i=1}^{q}\omega_i g_i\cos(\omega_i t + \phi_i)$. The proof for the signal given by (40) is given in [19]. □

Proof of Theorem 1: We show the stability together with the convergence. The closed-loop system of $(e_p(t), \widetilde{\theta}(t))$ is written as

$$\dot{\zeta}(t) = E(t)\zeta(t) + F(t)\delta(t), \tag{41}$$

where

$$E(t) = \begin{bmatrix} -\frac{1}{c_{e_p}}I_m & \frac{a_{21}}{c_{e_p}\bar{b}_1 a_{11}}\bar{B}_m\xi(t)^T \\ \kappa\frac{a_{21}}{c_{e_p}\bar{b}_1 a_{11}}\xi(t)\bar{B}_m^T & -\frac{\kappa}{c_{e_p}}\left(\frac{a_{21}}{\bar{b}_1 a_{11}}\right)^2\xi(t)\xi(t)^T \end{bmatrix}, \tag{42}$$

$$F(t) = \begin{bmatrix} \frac{a_{21}}{c_{e_p}\bar{b}_1 a_{11}}\bar{B}_m\theta^T \\ -\frac{\kappa}{c_{e_p}}\left(\frac{a_{21}}{\bar{b}_1 a_{11}}\right)^2\xi(t)\theta^T \end{bmatrix}, \tag{43}$$

$$\zeta(t) = \begin{bmatrix} e_p(t), & \widetilde{\theta}^T(t) \end{bmatrix}^T. \tag{44}$$

Since system matrix, $E(t)$ and input vector, $F(t)$ are time dependent, system (41) is a linear time varying system. The proof consists of two steps; first the homogeneous part is considered and the exponential stability of the equilibrium $\zeta(t) = 0$ is shown, as the second step the signal convergence is proven by employing the results both given in Lemma 2 and the first step of the proof.

For the first step, we choose the Lyapunov function as follows

$$V_{LTV} = \frac{1}{2}\zeta(t)^T P_{LTV}\zeta, \tag{45}$$

where

$$P_{LTV} = \begin{bmatrix} I_m & 0_{m\times 2q} \\ 0_{2q\times m} & \frac{1}{\kappa}I_{2q} \end{bmatrix}. \tag{46}$$

The time derivative of V_{LTV} is given by

$$\dot{V}_{LTV} = \zeta(t)^T C(t)C^T(t)\zeta(t). \tag{47}$$

where

$$C(t)^T = \frac{1}{\sqrt{c_{e_p}}}\begin{bmatrix} -I_m & \frac{a_{21}}{\bar{b}_1 a_{11}}\bar{B}_m\xi(t)^T) \end{bmatrix}, \tag{48}$$

we get

$$\dot{V}_{LTV} = \frac{1}{2}\zeta(t)^T\left(E^T(t)P_{LTV} + P_{LTV}E(t)\right)\zeta(t) = -\zeta(t)^T C(t)C^T(t)\zeta(t) \leq 0. \tag{49}$$

This proves the stability of equilibrium. However, we can not make any conclusion on the exponential stability. Therefore, we need to employ the

series of theorems and lemmas given in [25]. Using (49), we obtain

$$E^T(t)P_{LTV} + P_{LTV}E(t) + \alpha C^T(t)C(t) \leq 0, \tag{50}$$

for some $\alpha > 0$. The exponential stability of the equilibrium of the homogeneous part of (41) is proven by showing the uniform complete observable (UCO) property of the pair $(C(t), E(t))$ [25]. In order to simplify the problem, we employ the following property which is given as follows; for a bounded $H(t)$, the pairs $(C(t), E(t))$ and $(C(t), E(t) + H(t)C(t)^T)$ have the same (UCO) property [25]. Choosing $H(t) = C(t)$, the system for the pair $(C(t), E(t)+H(t)C(t)^T)$ is written as

$$\dot{Y}(t) = 0, \tag{51}$$

$$y(t) = C(t)^T Y(t) \tag{52}$$

The state transition matrix of (51) is $\Phi(t) = I_{(m+2q)}$. Therefore, $(C(t), E(t) + H(t)C(t)^T)$ is a UCO pair if there exist $\alpha_2, \alpha_3, \rho > 0$, such that the observability gramian satisfies

$$\alpha_2 I \geq \int_{t_0}^{t_0+\rho} C(t)C^T(t)\,dt \geq \alpha_3 I, \tag{53}$$

for all $t_0 \geq 0$ [25]. The boundedness of $\xi(t)$ can be shown by (40) then, recalling (48), the upper bound of (53) is satisfied. We show the lower bound in (53). Calculating the integral in (53) yields

$$X = \int_{t_0}^{t_0+\rho} C(t)C^T(t)\,dt$$

$$= \frac{1}{c_{e_p}} \begin{bmatrix} \rho I_m & -\frac{a_{21}}{\bar{b}_1 a_{11}} \int_{t_0}^{t_0+\rho} \xi(t)\,dt \bar{B}_m^T \\ -\frac{a_{21}}{\bar{b}_1 a_{11}} \bar{B}_m \int_{t_0}^{t_0+\rho} \xi^T(t)\,dt & \left(\frac{a_{21}}{\bar{b}_1 a_{11}}\right)^2 \int_{t_0}^{t_0+\rho} \xi(t)\xi^T(t)\,dt \end{bmatrix}. \tag{54}$$

(55)

The lower bound argument is shown by concluding that X is positive definite. To this end, we use Schur complement. The Schur complement of ρI_m of block matrix X is given by,

$$S_h = \frac{1}{c_{e_p}} \left(\frac{a_{21}}{\bar{b}_1 a_{11}}\right)^2 \left(\int_{t_0}^{t_0+\rho} \xi\xi^T\,dt - \frac{1}{\rho} \int_{t_0}^{t_0+\rho} \xi\,dt \int_{t_0}^{t_0+\rho} \xi^T\,dt \right). \tag{56}$$

According to Schur compliment, it is concluded that X is positive definite if and only if S_h is positive definite by denoting $I_m\rho$ is a positive definite matrix. We need to show that S_h is positive definite. Noting that $\frac{1}{c_{e_p}} \left(\frac{a_{21}}{\bar{b}_1 a_{11}}\right)^2$ is a positive scalar, according to Lemma 2 there exists a positive ρ such that for all $t_0 > 0$, then $S_h > 0$. Hence, $(C, E + HC^T)$ is UCO, which implies that (C, E) is UCO. Therefore, the state transition matrix $\Phi(t, t_0)$ corresponding to $E(t)$ in (41) satisfies

$$\| \Phi(t, t_0) \| \leq \kappa_0 e^{-\gamma_0(t-t_0)} \tag{57}$$

for some positive constants κ_0, γ_0. Recalling that G is Hurwitz, the solution of (26) yields

$$|\delta(t)| = |e^{G(t-t_0)}\delta(0)| \leq \kappa_1 e^{-\gamma_1(t-t_0)}|\delta(0)| \tag{58}$$

for some positive constants κ_1, γ_1. Solving (41) yields

$$\zeta(t) = \Phi(t, 0)\zeta(0) + \int_0^t \Phi(t, \tau)F(\tau)\delta(\tau)d\tau. \tag{59}$$

Since $\xi(t)$ is bounded, from (40), it is concluded that $F(t)$ is bounded. Using (57)–(59), we get

$$|\zeta(t)| \leq \kappa_0 e^{-\gamma_0 t}|\zeta(0)| + \frac{\kappa_1\kappa_0 \sup\limits_{0\leq\tau\leq t} |F(\tau)|}{\left(\gamma_0 - \frac{1}{2}\min\{\gamma_0, \gamma_1\}\right)}|\delta(0)| \times e^{-\frac{1}{2}\min\{\gamma_0,\gamma_1\}t}. \tag{60}$$

Finally, by considering (60), we prove that $\zeta(t) = \left[e_p(t), \quad \widetilde{\theta}^T(t)\right]^T$ converge to zero as $t \to \infty$. Since $\widetilde{\theta}(t)$ tends to zero, in other words, the perfect estimation of unknown constant parameter θ is achieved, from Lemma 1, and noting that $\delta(t)$ converges to zero, we also prove that $\nu(t) - \hat{\theta}^T(t)\xi(t)$. □

The result of Theorem 1 enables us to estimate unknown disturbance signal $\nu(t)$ with unmeasured state $p(t)$ for an open loop case. The perfect estimation relies on the excitation of all modes of the disturbance signal. This provides the persistence of excitation that is given in Lemma 2.

In the next section, the algorithm for the input $u(t)$ to cancel the disturbance in the system is given.

4. Active disturbance cancellation

We design an active disturbance cancellation algorithm for the actual input $u(t)$ by employing a backstepping technique. This technique is firstly

proposed by Petar Kokotovic for the control design of nonlinear system which are in strict feedback form [26]. However, the usage of this method is not restricted by only a special class of nonlinear systems. In [27], adaptive controller design methodology without any over parametrization for systems that contain parametric uncertainties which are unmatched with the actual input is given based on aforementioned technique. The idea behind the backstepping is to find a control law that transform the closed loop system to a desired form step by step, this can also be seen as recursive method. Particularly, in the considered problem, firstly we design a law for the virtual control input $p(t)$ then by backstepping, we reach the actual control input. It is called virtual control, since it is not the actual input but it is seen as the input of the subsystem. Therefore, the main role of actual input to drive the virtual input so that the state of the considered system is driven to a desired set. Since $p(t)$, is not measured, this technique can not be applied directly. This problem is handled by using the result obtained in the previous section. Using (36) and (34), we represent the unmeasured $p(t)$ as follows,

$$\begin{aligned} p(t) =& \hat{p}(t) - c_{e_p}\dot{e}_p(t) + \frac{a_{21}}{\bar{b}_1 a_{11}}\bar{B}_m\Big(\widetilde{\theta}^T(t)\xi(t) \\ &+ \theta^T\delta(t)\Big). \end{aligned} \tag{61}$$

4.1. Backstepping design

We can now consider the problem as an adaptive control design. Substituting (19), (61) into (27) and recalling $\theta = \widetilde{\theta}(t) + \hat{\theta}(t)$ from (35) , we obtain

$$\begin{aligned} x(t) =& A_n^T\dot{x}(t) + \frac{1}{a_{11}}\bar{B}_n\Bigg(- b_x^T\hat{p}(t) + c_{e_p}b_x^T\dot{e}_p(t) \\ &+ \dot{x}_n(t) - a_x^T A_n^T\dot{x}(t) - \hat{\theta}^T\xi(t) \\ &- \left(\frac{a_{21}b_{11}}{\bar{b}_1 a_{11}} + 1\right)\left(\widetilde{\theta}^T\xi(t) + \theta^T\delta(t)\right)\Bigg). \end{aligned} \tag{62}$$

It is possible to consider $b_x^T\hat{p}(t)$ as the virtual controller for a backstepping design since the dynamic of $\hat{p}(t)$ depends on only measured signals.

The backstepping procedure consist of the following steps; (1) designing a controller $b_x^T\hat{p}(t)$ that both cancels the disturbance and stabilize the system, (2) defining the error between the desired and actual value $b_x^T\hat{p}(t)$, this step can be considered as a state transformation, (3) by taking the time derivative of the defined error term, we find the dynamics of error. In the last step, we reach the actual control input.

First Step: The desired value of $b_x^T\hat{p}(t)$ is designed by

$$\hat{p}_d(t) = -a_{11}K\dot{x}(t) - \hat{\theta}^T(t)\xi(t) + \dot{x}_n(t) - a_x^T A_n^T\dot{x}(t), \tag{63}$$

where the control gain $K = [k_1, \cdots, k_n] \in \mathbb{R}^{1\times n}$ is chosen so that $\left(A_n^T + \bar{B}_n K\right)$ is Hurwitz with $P = P^T > 0$ that is the solution of

$${A_{cl}}^{-T}P + P{A_{cl}}^{-1} = -2\left(2 + \frac{c_{e_p}}{2}\right)I. \tag{64}$$

where

$$A_{cl} = \left(A_n^T + \bar{B}_n K\right)^{-1}. \tag{65}$$

Second Step: we define the following error term,

$$e_d(t) = b_x^T\hat{p}(t) - \hat{p}_d(t). \tag{66}$$

Third Step: Taking the derivative of (66), we obtain

$$\begin{aligned} \dot{e}_d(t) =& -\frac{1}{c_{e_p}}b_x^T\hat{p}(t) + \Bigg(\frac{1}{c_{e_p}}b_x^T\left(A_m^T + c_{e_p}I_m\right. \\ &\left.-\frac{1}{\bar{b}_1}\bar{B}_m\bar{b}_p^T A_m^T\right) + \left(a_{11}k_n - 1\right)b_x^T\Bigg)\dot{p}(t) \\ &+ \Bigg(\frac{-b_{11}}{c_{e_p}\bar{b}_1}\bar{a}_p^T A_n^T + a_{11}KA_n \\ &+ a_x^T A_n^T A_n + \left(a_{11}k_n - 1\right)a_x^T\Bigg)\dot{x}(t) \\ &+ \frac{b_{11}a_{21}}{c_{e_p}\bar{b}_1 a_{11}}\hat{\theta}^T(t)\xi(t) + \hat{\theta}^T(t)G\xi(t) \\ &+ \dot{\hat{\theta}}^T(t)\xi(t) + \Big(\hat{\theta}^T(t)l + \left(a_{11}k_n - 1\right)\Big) \\ &\times \left(\beta^T\xi(t) + \beta^T\delta(t)\right) + \frac{b_{11}}{c_{e_p}\bar{b}_1}\big(\dot{p}_m(t) \\ &- \frac{a_{21}}{a_{11}}\dot{x}_n(t) - u(t)\big). \end{aligned} \tag{67}$$

The representation (20) is used for $\dot{\nu}(t)$. In this way, we transform $(x(t), p(t))$ system into $(x(t), e_p(t), e_d(t))$ system where the uncertainties are now matched with the real actuator.

As it is realized from (66), if $e_d(t)$ converges to zero, the virtual controller tends to its desired value which cancels the disturbance effect and stabilize system (1).

In the next section, the adaptive controller and the main theorem are given.

4.2. Adaptive controller

Considering the transformed system (1), (36) and (67) and their RSS forms, we design the following adaptive controller

$$\begin{aligned} u =&\dot{p}_m(t) - \frac{a_{21}}{a_{11}}\dot{x}_n(t) + \frac{c_{e_p}\bar{b}_1}{b_{11}}\Bigg(- \frac{1}{c_{e_p}} b_x^T \hat{p}(t) \\ &+ \left(\frac{1}{c_{e_p}} b_x^T \left(A_m^T + c_{e_p} I_m - \frac{1}{\bar{b}_1}\bar{B}_m \bar{b}_p^T A_m^T\right)\right. \\ &\left. + \left(a_{11}k_n - 1\right) b_x^T\right)\dot{p}(t) + \left(\frac{-b_{11}}{c_{e_p}\bar{b}_1}\bar{a}_p^T A_n^T\right. \\ &\left. + a_{11} K A_n + a_x^T A_n^T A_n + \left(a_{11}k_n - 1\right) a_x^T\right)\dot{x}(t) \\ &+ \frac{b_{11}a_{21}}{c_{e_p}\bar{b}_1 a_{11}}\hat{\theta}^T(t)\xi(t) + \hat{\theta}^T(t) G \xi(t) + \dot{\hat{\theta}}^T(t)\xi(t) \\ &+ \left(\hat{\theta}^T(t) l + \left(a_{11}k_n - 1\right)\right)\hat{\beta}^T(t)\xi(t) \\ &- \frac{1}{a_{11}}\dot{x}^T(t) P \bar{B}_n \\ &+ \left(\frac{1}{2}\left(\hat{\theta}^T(t) l + \left(a_{11}k_n - 1\right)\right)^2 + c_{e_d}\right) e_d(t)\Bigg), \end{aligned} \tag{68}$$

where $c_{e_d} > 0$. The dynamics to update parameter estimation is given by

$$\begin{aligned} \dot{\hat{\theta}} =&\kappa_\theta \frac{1}{\bar{b}_1 a_{11}}\xi(t)\Bigg(a_{21}\bar{B}_m^T \dot{e}_p(t) \\ &- \frac{a_{21}b_{11} - \bar{b}_1 a_{11}}{\gamma a_{11}}\dot{x}(t)^T P \bar{B}_n\Bigg), \end{aligned} \tag{69}$$

$$\dot{\hat{\beta}} =\kappa_\beta \left(\hat{\theta}^T(t) l + \left(a_{11}k_n - 1\right)\right)\xi(t) e_d(t). \tag{70}$$

with

$$\gamma = \frac{1}{a_{11}^2}\lambda_{\max}\left(\bar{B}_n^T P b_x^T b_x P \bar{B}_n\right). \tag{71}$$

Both control and update laws consist of only measured signals.

As it is seen by comparing the controller (68) and the one designed in [21], they are different than each other. It is not possible to employ the controller in [21] to stabilize the system given in the problem statement of this note.

Remark 1. Although parameter θ represents the same property, the update laws given for $\hat{\theta}(t)$ in the open loop estimation and the closed loop cancellation are slightly different. The update law (69) employed in closed loop cancellation contains an extra term, which is given by $-\kappa_\theta \xi(t)\frac{a_{21}b_{11} - \bar{b}_1 a_{11}}{b_1 a_{11}^2}\dot{x}(t)^T P \bar{B}_n$. The backstepping technique causes the extra term in the update law.

Remark 2. We design the control and update laws for the case where all state-derivative measurements are accurate. That might not be the case in the implementation of algorithms to actual systems, since the measurements providing by sensors in the applications may contain noise unless an appropriate filtering is applied. In adaptive control, measurement noise and unmodeled dynamics may harm the stability of the equilibrium due to drift in the estimation of the parameters. However, there exist simple robustification tools for mostly update laws to eliminate the effect of noise and unmodeled dynamics [25]. One of the appropriate robustification tools can be employed together with the proposed algorithm to provide robustness with respect to measurement noise and unmodeled dynamics. The main drawback of the these tools is the trade off between robustness and the convergence.

We define the following signal

$$\widetilde{\xi} = \xi - \overline{\xi}, \tag{72}$$

where $\overline{\xi} = \int_0^t e^{G(t-\tau)} G l \dot{\nu}(\tau) d\tau$ that is used in the stability statement.

Theorem 2. *Consider closed-loop system that is composed of the plant* (1),(2) *driven by the unknown disturbance signal* (8), (9)*, the disturbance observer dynamics* (21), (22)*, the state observer* (33) *and the adaptive controller* (68)–(70)*. Under Assumptions 1 and 2, the followings hold:*

- *For all initial conditions, all signals are bounded and* $\dot{x}(t), \dot{e}_p(t), e_d(t), \widetilde{\xi}, \delta$ *converge to zero as* $t \to \infty$*,*
- *In addition, for all* $w(0) \in \mathbb{R}^{2q}$ *such that Assumption 3 holds, the signals* $x(t), e_p(t), \widetilde{\theta}, \nu(t) - \hat{\theta}^T\xi$ *converge to zero as* $t \to \infty$*.*

5. Stability proof

The proof of Theorem 2 is given as follows.

Proof of Theorem 2: The transformed closed system with the state observer is given as follows

$$\begin{aligned} \dot{x}(t) =&A_{cl}x(t) - \frac{1}{a_{11}}A_{cl}\bar{B}_n\Big(- e_d(t) + c_{e_p} b_x^T \dot{e}_p(t) \\ &- \left(\frac{a_{21}b_{11}}{\bar{b}_1 a_{11}} + 1\right)\left(\widetilde{\theta}^T\xi(t) + \theta^T\delta(t)\right)\Big), \end{aligned} \tag{73}$$

$$\dot{e}_p(t) = -\frac{1}{c_{e_p}} e_p(t) + \frac{a_{21}}{c_{e_p}\bar{b}_1 a_{11}} \bar{B}_m \left(\widetilde{\theta}^T(t)\xi(t) + \theta^T \delta(t)\right), \tag{74}$$

$$\dot{e}_d = -\left(\frac{1}{2}\left(\hat{\theta}^T(t)l + (a_{11}k_n - 1)\right)^2 + c_{e_d}\right) e_d(t) + \left(\hat{\theta}^T(t)l + (a_{11}k_n - 1)\right)\left(\widetilde{\beta}^T(t)\xi(t) + \beta^T\delta(t)\right) + \frac{1}{a_{11}}\dot{x}^T(t)P\bar{B}_n \tag{75}$$

where

$$\widetilde{\beta}(t) = \hat{\beta}(t) - \beta. \tag{76}$$

We consider the following Lyapunov function,

$$V = \frac{1}{2}\left(x^T P x + \gamma e_p^2 + e_d^2 + \frac{\gamma}{\kappa_\theta}\widetilde{\theta}^T\widetilde{\theta} + \frac{1}{\kappa_\beta}\widetilde{\beta}^T\widetilde{\beta} + \epsilon_\delta \delta^T P_\delta \delta\right) \tag{77}$$

where

$$G^T P_\delta + P_\delta G = -2I, \tag{78}$$

$$\epsilon_\delta = \left(\frac{a_{21}b_{11}}{\bar{b}_1 a_{11}^2} + \frac{1}{a_{11}}\right)^2 \lambda_{\max}\left(\theta \bar{B}_n^T P P \bar{B}_n \theta^T\right) + \left(\gamma \frac{a_{21}b_{11}}{\bar{b}_1 a_{11}}\right)^2 \lambda_{\max}\left(\theta \bar{B}_m \bar{B}_m^T \theta^T\right) + \lambda_{\max}\left(\beta\beta^T\right). \tag{79}$$

The derivative of V with respect to time, in view of (20),(69),(70) and (73)–(75), is given by

$$\dot{V} = -\left(2 + \frac{c_{e_p}}{2}\right)\dot{x}^T(t)\dot{x}(t) - \gamma c_{e_p}\dot{e}_p^T(t)\dot{e}_p(t) - \left(\frac{1}{2}\left(\hat{\theta}^T(t)l + (a_{11}k_n - 1)\right)^2 + c_{e_d}\right) e_d^2(t) + \frac{1}{a_{11}}\dot{x}^T P\bar{B}_n\left(c_{e_p} b_x^T \dot{e}_p(t) - \left(\frac{a_{21}b_{11}}{\bar{b}_1 a_{11}} + 1\right)\theta^T\delta(t)\right) + \frac{a_{21}}{\bar{b}_1 a_{11}}\dot{e}_p^T \bar{B}_m \theta^T \delta + \left(\hat{\theta}^T(t)l + (a_{11}k_n - 1)\right) \times \left(\widetilde{\beta}^T(t)\xi(t) + \beta^T\delta(t)\right)\beta^T\delta e_d - \epsilon_\delta^2 \delta^T\delta. \tag{80}$$

Using Young's inequality, we get

$$\dot{V} \leq -\dot{x}^T(t)\dot{x}(t) - \frac{\gamma c_{e_p}}{2}\dot{e}_p^T(t)\dot{e}_p(t) - c_{e_d} e_d^2(t) - \frac{\epsilon_\delta}{2}\delta^T(t)\delta(t). \tag{81}$$

From (81), we conclude

$$V(t) \leq V(0). \tag{82}$$

Defining

$$\Theta(t) = \left[x^T(t), e_p(t), e_d(t), \widetilde{\theta}^T(t), \widetilde{\beta}^T(t), \delta^T(t)\right]^T, \tag{83}$$

and using (77) and (82), we get

$$|\Theta(t)|^2 \leq M_1 |\Theta(0)|^2, \tag{84}$$

for some $M_1 > 0$. Taking derivative of (72) and using (40) yield

$$\dot{\widetilde{\xi}}(t) = G\widetilde{\xi}(t). \tag{85}$$

Since G is Hurwitz, using (85), we get

$$\left|\widetilde{\xi}(t)\right| \leq M_2 \mathrm{e}^{-\alpha_1 t}\left|\widetilde{\xi}(0)\right|, \tag{86}$$

for some $M_2, \alpha_1 > 0$. By using (84) and (86), we obtain

$$|\Xi(t)| \leq M_4 |\Xi(0)| \tag{87}$$

where

$$\Xi(t) = \left[\Theta^T(t), \widetilde{\xi}^T(t)\right]^T, \tag{88}$$

for some $M_4 > 0$. From (87), it is concluded that all signals are bounded for all initial conditions. Since the closed loop dynamics given by (73)–(75) are continuous in Ξ and t, (81) is also continuous in Ξ and t. Moreover, (81) is zero at $\Xi = 0$. Therefore, we conclude that $\dot{x}(t), \dot{e}_p(t), e_d(t)$ and δ converge to zero as $t \to \infty$, by the LaSalle-Yoshizawa theorem. This proves part 2 of Theorem 2.

We now prove the convergence of $x(t)$. To this end we first need to show the convergence of $e_p(t), \widetilde{\theta}(t))$. This part is similar to the proof of Theorem 1. The system of $(e_p(t), \widetilde{\theta}(t))$ is written as

$$\dot{\zeta}(t) = E(t)\zeta(t) + F_d(t)d(t), \tag{89}$$

where $E(t)$ and $\zeta(t)$ are given by (42) and (44), respectively and

$$F_d(t) = \begin{bmatrix} 0_{m\times n} \\ \kappa_\theta \frac{a_2 1 b_1 1 - \bar{b}_1 a_1 1}{\gamma a_1^2 1 \bar{b}_1} \xi(t) \bar{B}_n^T P \\ \frac{a_{21}}{c_{e_p} \bar{b}_1 a_{11}} \bar{B}_m \theta^T \\ -\frac{\kappa_\theta}{c_{e_p}} \left(\frac{a_{21}}{\bar{b}_1 a_{11}} \right)^2 \xi(t) \theta^T \end{bmatrix}, \tag{90}$$

$$d(t) = \begin{bmatrix} \dot{x}(t), & \delta^T(t) \end{bmatrix}^T. \tag{91}$$

As it is seen the the homogeneous part of systems (41) and (89) are same. The only difference is the input matrices. We have already shown the exponential stability of the homogeneous part in the proof of Theorem 1. The boundedness of $d(t)$ and $F_d(t)$ is concluded from (91) and (90), respectively. Noting that $d(t)$ goes to zero, from (89) and (57), it follows that $\zeta(t)$ is bounded and $\zeta(t) = \begin{bmatrix} e_p(t), & \widetilde{\theta}^T(t) \end{bmatrix}^T \to 0$ as $t \to \infty$. Since $\delta(t), \widetilde{\theta}(t), \dot{e}_p(t), e_d(t) \to 0$ and A_{cl} is Hurwitz, from (73), it is shown that $x(t)$ converges to zero as $t \to \infty$. In addition, from (19), we conclude that $\hat{\theta}^T(t)\xi(t) - \nu(t) \to 0$ as $t \to \infty$. This proves part 2 of Theorem 2. ■

6. Example

We consider a two–degree of freedom system which is illustrated in Figure 1, as a simulation example.

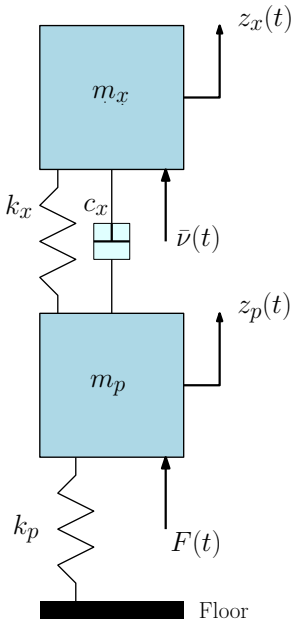


Figure 1. The dynamical system considered for numerical example.

The mass denoted by m_x is forced by an unknown disturbance $\bar{\nu}(t)$ and connected to another mass, m_p which is actuated by a force input. The system dynamics are given by

$$\begin{aligned} m_x \ddot{z}_x(t) = &- k_x \left(z_x(t) - z_p(t) \right) \\ &- c_x \left(\dot{z}_x(t) - \dot{z}_p(t) \right) + \nu(t), \qquad (92) \\ m_p \ddot{z}_p(t) = &- k_p z_p + k_x \left(z_x(t) - z_p(t) \right) \\ &+ c_x \left(\dot{z}_x(t) - \dot{z}_p(t) \right) + F(t). \qquad (93) \end{aligned}$$

Choosing the state as $x(t) = \begin{bmatrix} z_x(t), & \dot{z}_x(t) \end{bmatrix}^T$ and $p(t) = \begin{bmatrix} z_p(t), & \dot{z}_p(t) \end{bmatrix}^T$, we obtain $a_x^T = \begin{bmatrix} -\frac{k_x}{m_x}, & -\frac{c_x}{m_x} \end{bmatrix}$, $b_x^T = \begin{bmatrix} \frac{k_x}{m_x}, & \frac{c_x}{m_x} \end{bmatrix}$, $a_p^T = \begin{bmatrix} \frac{k_x}{m_p}, & \frac{c_x}{m_p} \end{bmatrix}$ and $b_p^T = \begin{bmatrix} -\frac{k_x + k_p}{m_p}, & -\frac{c_x}{m_p} \end{bmatrix}$. The disturbance is $\nu(t) = \frac{\bar{\nu}(t)}{m_x}$ and the input is $u(t) = \frac{F(t)}{m_p}$. The system parameters are chosen as $m_x = 10, m_p = 5, k_x = 10, k_p = 5, c_x = 10$. The mass denoted by m_x is driven by the unknown periodic disturbance $\bar{\nu}(t)$ and the states of mass m_p through the connected spring and damper. The mass m_x is the subject system to stabilize. The effect of the input $F(t)$ reaches the main system through the dynamics of m_p whose order is more than one. Therefore, it should be noted that the design given in [21] can not be used for this example. By employing the developed algorithms, we present the results for estimation and active disturbance cancellation.

6.1. Observer based disturbance and state estimation

In this case, the only aim is estimation. The unknown disturbance signal $\nu(t) = \frac{\bar{\nu}(t)}{m_x} = \sin(\frac{2\pi}{3.5}t) + 0.5\sin(\frac{2\pi}{4}t + \frac{\pi}{4}) + 0.5\sin(\frac{2\pi}{3}t + \pi)$ and initial condition $x(0) = p(0) = \begin{bmatrix} -0.5 & -0.5 \end{bmatrix}^T$. The update and estimations gains is chosen as $\kappa = 15000$ and $c_{e_p} = 20$. The pair (G, l) is $l = \begin{bmatrix} 0_5^T, & 1 \end{bmatrix}^T$, $G = \begin{bmatrix} 0_5 & I_5 \\ 0 & 0_5^T \end{bmatrix} + l \begin{bmatrix} -527.8 & -1146.1 & -1024.2 & -482.5 & -126.4 & -17.5 \end{bmatrix}^T$ which is controllable. Since the critical signal is the position of m_p, we have plotted its estimate in Figure 2. The estimation of the disturbance together with the actual signal are plotted in Figure 3. From Figures 2 and 3, it is seen that $|z_p(t) - \hat{z}_p(t)|$ converge to zero and the estimate of unknown disturbance, $\hat{\theta}^T(t)\xi(t)$ converges to the actual disturbance signal, thus perfect estimation is achieved, as Theorem 1 states.

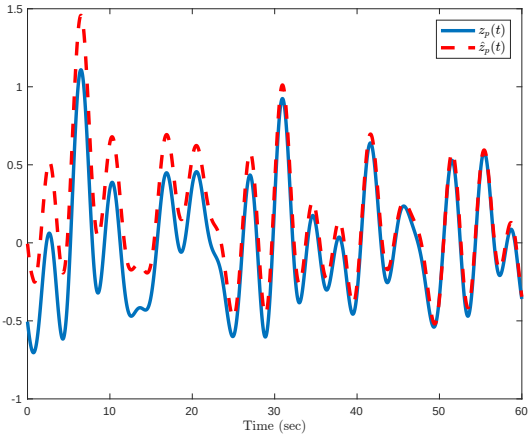


Figure 2. Estimation of the actuator state $\hat{z}_p(t)$ for the observer based estimation.

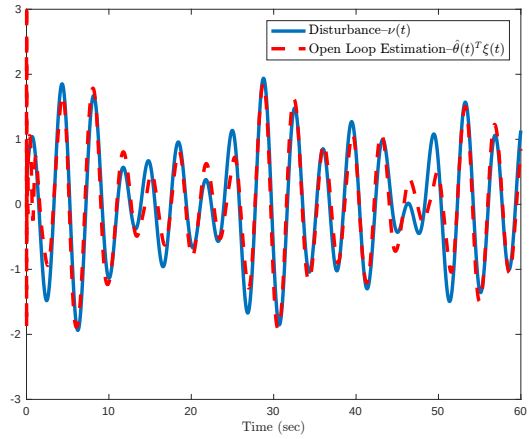


Figure 3. The actual disturbance signal $\nu(t)$ and its estimate, $\hat{\theta}^T(t)\xi(t)$, for the observer based estimation.

6.2. Active cancellation

In this case, the aim is to actuate mass m_p by $F(t)$ so that the $z_p(t), \dot{z}(p)$ cancel the disturbance on mass m_x. The unknown disturbance $\nu(t) = \frac{\bar{\nu}(t)}{m_x} = 2\sin(\frac{2\pi}{3}t) + \sin(\frac{2\pi}{t} + \frac{\pi}{4})$ and the initial conditions of the states are chosen same as the first case. The control gain K is chosen such that the eigenvalues of A_{cl}^{-1} are -3 and -4. The gains are $c_{e_d} = 350, c_{e_p} = 0.2, \kappa_\theta = 50$ and $\kappa_\beta = 10$. The pair (G, l) is $l = \begin{bmatrix} 0_3^T, & 1 \end{bmatrix}^T$, $G = \begin{bmatrix} 0_3 & I_3 \\ 0 & 0_3^T \end{bmatrix} + l \begin{bmatrix} -48.51 & -75.91 & -43.63 & -10.90 \end{bmatrix}^T$ which is controllable. The state of mass m_x is given in Figure 4. The result of the disturbance estimation is presented in Figure 5. From Figures 4 and 5 it is observed that $z_x(t), \dot{z}_x(t)$ converge to zero and the estimate of unknown disturbance, $\hat{\theta}^T(t)\xi(t)$ converges to the actual disturbance signal, thus perfect estimation is achieved, as Theorem 2 states.

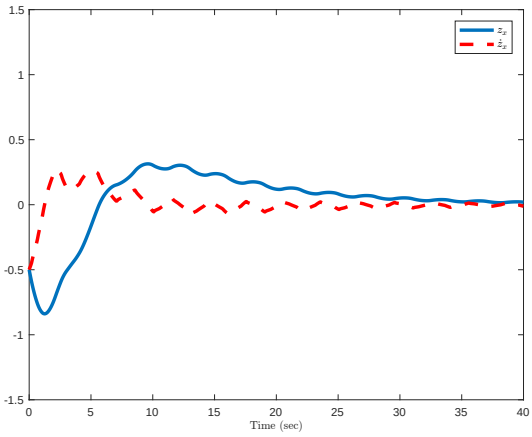


Figure 4. Position, $z_x(t)$, and velocity, $\dot{z}_x(t)$, of mass m_x when the cancellation algorithm is on.

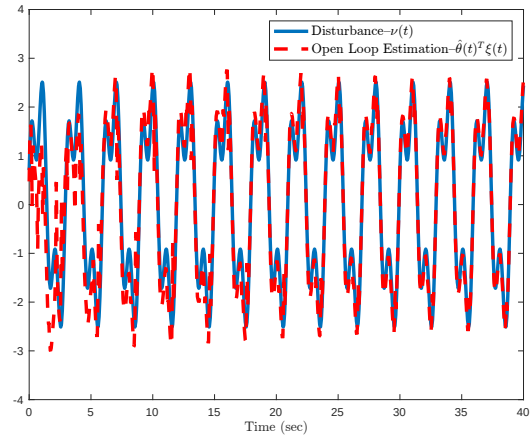


Figure 5. The actual disturbance signal $\nu(t)$ and its estimate, $\hat{\theta}^T(t)\xi(t)$, for the active cancellation.

7. Conclusions

In this note, we consider an active disturbance cancellation problem for the case where the disturbance signal which is modeled as the sum of q sinusoidal functions, and the real control input of the system are unmatched. Moreover, the only available measurements for the feedback are the state derivatives. Two algorithms are developed; the first one is to estimate the unmeasured disturbance and states in the open loop, the second one is to cancel the effect of the disturbance and maintain the state convergence in the closed loop case. To this end, observers are designed for both the unknown disturbance and the unmeasured state that appears as the virtual control input in the system. We perform a backstepping procedure to handle the unmatched condition and obtain the control and update laws by using the estimate of the unmeasured state as the virtual control input. The stability of the equilibrium of the closed-loop adaptive system and the the convergence of state derivatives, $\dot{x}$, to zero in time are shown. By assuming all modes of the disturbance signal are excited, we also achieve to prove that the state $x(t)$ converges to zero as $t \to \infty$ and the estimate of disturbance, $\hat{\theta}^T(t)\xi(t)$ converges to actual disturbance signal, $\nu(t)$. We illustrate the effectiveness of developed algorithms by performing numerical simulations for a two–degree of freedom spring–damper–mass system. The simulation results also verify that the developed algorithms are working correctly as the given theorems state.

References

[1] Arslan Y.Z., Sezgin A.,& Yagiz N.(2015). Improving the ride comfort of vehicle passenger using fuzzy sliding mode controller. *Journal of Vibration and Control*, 21(9), 1667–1679.

[2] Marconi, L., Isidori A., & Serrani, A. (2002). Autonomous vertical landing on an oscillating platform: an internal-model based approach. *Automatica*, 38(1), 21–32.

[3] Basturk, H. I., & Krstic,M. (2013). Adaptive wave cancellation by acceleration feedback for ramp-connected air cushion-actuated surface effect ships. *Automatica*, 49(9), 2591–2602.

[4] Basturk, H. I., Rosenthal, B., & Krstic, M. (2103). Pitch control design for tandem lifting body catamaran by aft lifting body actuation. *IEEE Transactions on Control Systems Technology*, 23(2), 700–707.

[5] Johnson, C. D. (1971). Accommodation of external disturbances in linear regulator and servomechanism problems. *IEEE Transactions on Automatic Control*, 16(6), pp. 635–644.

[6] Francis, D. A., & Wonham, W. N. (1975). The internal model principle for linear multivariable regulators. *Applied Mathematics and Optimization*, 2(2), 170–194.

[7] Dogruel, M., & Celik, H. H. (2011). Harmonic control arrays method with a real time application to periodic position control. *IEEE Transactions on Control Systems Technology*, 19(3), 521–530.

[8] Kim, H., & Shim, H. (2015) "Linear systems with hyperbolic zero dynamics admit output regulator rejecting unknown number of unknown sinusoids," *IET Control Theory and Applications*, **9** , pp. 1472 - 1480, 2015.

[9] Marino, R., & Santosuosso, G. L. (2007). Regulation of linear systems with unknown exosystems of uncertain model. *IEEE Transactions on Automatic Control*, 52(2), 353–359.

[10] Bodson. M., & Douglas, S. (1997) Adaptive algorithms for the rejection of sinusoidal disturbances with unknown frequency. *Automatica*, 33(12), 2213–2221.

[11] Bobtsov A. A., & Pyrkin, A. A. (2009). Compensation of unknown sinusoidal disturbances in linear plants of arbitrary relative degree. *Automation and Remote Control*, 70(3), 449–456.

[12] Marino, R., & Tomei, P. (2005). Adaptive tracking and disturbance rejection for uncertain nonlinear systems. *IEEE Transactions on Automatic Control*, 50(1), 90–95.

[13] Marino, R., & Santosuosso, G. L. (2005). Global compensation of unknown sinusoidal disturbances for a class of nonlinear nonminimum phase systems. *IEEE Transactions on Automatic Control*, 50(11), 1816–1822.

[14] Nikiforov, V. O. (2001). Nonlinear servocompensation of unknown external disturbances. *Automatica*, 37(10), 1617–1653.

[15] Ding, Z. (2003). Universal disturbance rejection for nonlinear systems in output feedback form. *IEEE Transactions on Automatic Control*, 48(7), 1222–1227.

[16] Kwak, S. K., Washington G., & Yedavalli, R. K. (2002) Acceleration-based vibration control of distributed parameter systems using the 'reciprocal state-space framework. *Journal od Sound and Vibration*, 251(3), 543–557.

[17] Abdelaziz, T. H. S. (2008). Robust pole assignment for linear time-invariant systems using state-derivative feedback. *Journal of Systems and Control Engineering*, 223(2), 187–199.

[18] Michiels, W., Vyhlídal, T., Huijberts, H., & Nijmeijer, H. (2009) Stabilizability and stability robustness of state derivative feedback controllers. *SIAM Journal on Control and Optimization*, 47(6), 3100–3117.

[19] Basturk, H. I., & Krstic, M. (2013). Adaptive cancelation of matched unknown sinusoidal disturbances for LTI systems by state derivative feedback. *Journal of Dynamic Systems, Measurement, and Control*, 135(1), 014501–014507.

[20] Basturk, H. I., & Krstic, M. (2012) Adaptive backstepping cancelation of unmatched unknown sinusoidal disturbances for LTI systems by state derivative feedback. *ASME Dynamic Systems and Control Conference*, 1-9.

[21] Basturk, H. I. (2015). Observer based adaptive estimation/cancellation of unmatched sinusoidal disturbances in known LTI systems by state derivative measurement. *ASME Dynamic Systems and Control Conference*, 2015.

[22] Nikiforov, V. O. (2004). Observers of external deterministic disturbances. I. objects with known parameters. *Automation and Remote Control*, 65(10), 1531–1541.

[23] Chen, C. T. (1984) *Linear System Theory and Design.* Rinehart, Winston. New York, NY: Holt.

[24] Salem, A., & Said, K. (2008). A simple proof of Sylvester's (Determinants) Identity. *Applied Mathematical Sciences*, 32(2), 1571–1580.

[25] Ioannou, P., & Sun, J. (1996) *Robust Adaptive Control.* Prentice-Hall.

[26] Kokotovic, P.V. (1990). The joy of feedback: nonlinear and adaptive. *IEEE Control Systems Magazine*, 12(3): 717.

[27] Krstic M., Kanellakopoulos I., & Kokotovic, P. (1995) *Nonlinear and Adaptive Control Design,* Wiley.

Halil Ibrahim Basturk *received his B.S. and M.S. degrees in Mechanical Engineering from Bogazici University in 2006 and 2008, and the Ph.D. degree in Mechanical and Aerospace Engineering from the University of California at San Diego in 2013. Since 2014 he has been an Assistant Professor of Mechanical Engineering at Bogazici University, Istanbul,Turkey. His research interests include disturbance estimation/cancellation, adaptive control, boundary control of infinite dimensional systems and extremum seeking.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.250-258 (2018)
http://doi.org/10.11121/ijocta.01.2018.00608

RESEARCH ARTICLE

Sinc-Galerkin method for solving hyperbolic partial differential equations

Aydin Secer

Department of Mathematical Engineering, Yildiz Technical University, Istanbul, Turkey
asecer@yildiz.edu.tr

ARTICLE INFO

Article History:
Received 30 April 2018
Accepted 14 June 2018
Available 24 July 2018

Keywords:
Sinc basis function
Linear matrix system
LU-decomposition method

AMS Classification 2010:
35L51, 65D32

ABSTRACT

In this work, we consider the hyperbolic equations to determine the approximate solutions via Sinc-Galerkin Method (SGM). Without any numerical integration, the partial differential equation transformed to an algebraic equation system. For the numerical calculations, *Maple* is used. Several numerical examples are investigated and the results determined from the method are compared with the exact solutions. The results are illustrated both in table and graphically.

1. Introduction

Most numerical methods are established on the basis of polynomials. One of these methods is Sinc methods. Frank Stenger firstly introduced the Sinc methods in his works [1, 22] to determine the solutions of some differential equations. The actual detailed analysis on Sinc functions were firstly made by Whittaker in the papers [20-21]. Lund has some works on two-point boundary-value problems [7, 9]. Lewis, Lund and Bowers investigated the parabolic and hyperbolic problems in [6, 12]. In [4] Bowers and Lund worked on singular Poisson and elliptic problems. Numerical solutions of the problems are found by means of SGM. Lund, Bowers and McArthur introduced the Symmetric SGM in [8]. A kind of Sinc methods which is called Sinc Domain Decomposition Method is illustrated in [10, 11, 14, and 15]. Moreover, iterative methods for symmetric Sinc-Galerkin systems are considered in [3, 16, and 17]. Some applications in the various areas of the science and engineering can be seen in [2, 5, 13, 18 and 19]. In the work of Morlet, Lybeck and Bowers in [15], a Volterra integro-differential equation is investigated via the Sinc-collocation method. In paper [1], Stenger made some applications with SGM for the approximate solutions of ODEs and some elliptic and parabolic PDEs. Koonprasert developed a fully SGM for some complex-valued PDEs with time-dependent boundary conditions in [5]. In the work of Stenger [23], some problems related to medical problems are taken and numerical results are found using Sinc methods. In [24], a new algorithm based on Sinc method is applied for the solution of a nonlinear set of PDEs. A new SGM is illustrated for the numerical solutions of convection diffusion equations on half-infinite intervals in [25]. The work in [26] Gamel, Behiry and Hashish dealed with the SGM for solving nonlinear ODEs with various boundary conditions. In [27], sinc-Galerkin method is also applied to a class of the second-order nonlinear ODEs. In the work [28], Zamani focused on some differential operators in one dimension and a Helmholtz eigenvalue problem in two dimensions. In [29], authors use sinc-Galerkin method to obtain approximate solution of fractional partial differential equations. In [30,32] Sinc methods is also applied for the solution of the second-order ODEs with homogeneous Dirichlet-Type boundary conditions. In [31], sinc-Galerkin

method is used for solving fractional boundary value problems approximately. In this paper, we use sinc-Galerkin method to obtain approximate solution of a class of hyperbolic partial differential equations. The rest of this paper is organized as follows. In section 2, we give some definitions and theorems for sinc methods. In section 3, some test problems are given to compare the ability of present methods by using tables and graphics. Finally, in section 4, the paper is completed with a conclusion.

2. Sinc-Galerkin method

2.1. Sinc-Approximation formula for hyperbolic

We consider the following hyperbolic partial differential equation.

$$\begin{array}{l} u_{tt} - u_{xx} = F(x,t), \\ u(0,t) = u(1,t) = 0, 0 < x < 1, \\ u(x,0) = 0, t > 0, \\ u_t(x,0) = 0, t > 0. \end{array} \tag{1}$$

To determine the approximate solution of this equation, Sinc-Galerkin method is used. For the equation given above, the sinc-Galerkin method can be developed in both space and time direction as following:

In general, approximations can be constructed for infinite, semi-infinite, and infinite intervals and both spatial and time spaces will be introduced.

Let us define the function ϕ as

$$\phi(z) = \ln\left(\frac{z}{1-z}\right). \tag{2}$$

Here ϕ is a conformal mapping from D_E, the eye-shaped domain in the z plane, onto the infinite strip, D_S where

$$D_E = \left\{ z = x + iy : \left| \arg\left(\frac{z}{1-z}\right) \right| < d \leq \frac{\pi}{2} \right\}. \tag{3}$$

This is shown in Figure 1.

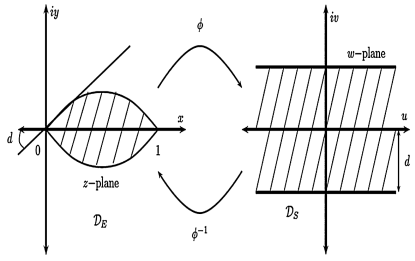


Figure 1. The connection between eye shape domain and the infinite strip [32].

A more general form of the sinc basis according to intervals can be given as following way

$$\begin{array}{c} S(m,h_x) \circ \phi(x) = Sinc\left(\dfrac{\phi(x) - mh_x}{h_x}\right), \\ m = -N_x, \ldots, N_x, \\ S(k,h_t) \circ \gamma(t) = Sinc\left(\dfrac{\gamma(t) - kh_t}{h_t}\right), \\ k = -N_t, \ldots, N_t, \end{array} \tag{4}$$

where

$$Sinc(z) = \begin{cases} \frac{\sin(\pi z)}{\pi z} & , \ z \neq 0 \\ 1 & , \ z = 0 \end{cases} \tag{5}$$

and

$$\begin{aligned} Sinc(k,h)(z) &= Sinc\left(\frac{z-kh}{h}\right) \\ &= \begin{cases} \frac{\sin\left(\pi \frac{z-kh}{h}\right)}{\pi \frac{z-kh}{h}}, & z \neq kh \\ 1, & z = kh \end{cases}, \end{aligned} \tag{6}$$

and the conformal maps for both direction as follows

$$\begin{cases} \phi(x) = \ln\left(\frac{x}{l-x}\right) & , \ x \in (0,l) \\ \gamma(t) = \ln(t) & , \ t \in (0,\infty), \end{cases} \tag{7}$$

are used to define the basis functions on the intervals $(0,l)$ and $(0,\infty)$ respectively. $h_x, h_t > 0$ represents the mesh sizes in the space direction and the time direction respectively. The sinc nodes x_i and t_j are chosen so that $x_i = \phi^{-1}(ih_x)$,$t_j = \gamma^{-1}(jh_t)$. Here the function $x = \phi^{-1}(x) = \frac{e^x}{1+e^x}$ is an inverse mapping of $\phi = \phi(x)$.

We may define the range of ϕ^{-1}on the real line as

$$\Gamma_1 = \left\{ \phi^{-1}(u) \in D_E : -\infty < u < \infty \right\}. \tag{8}$$

For the evenly spaced nodes $\{kh\}_{k=-\infty}^{\infty}$ on the real line, the image which corresponds to these nodes is denoted by

$$x_k = \phi^{-1}(kh) = \frac{e^{kh}}{1+e^{kh}} \tag{9}$$

where $0 < x_k < 1$, for all k.

The sinc basis functions in (4) do not have a derivative when x tends to 0 or 1. We modify the sinc basis functions as

$$\frac{S(m,h_x)\circ\phi(x)}{\phi'(x)}=\frac{Sinc\left(\frac{\phi(x)-mh_x}{h_x}\right)}{\phi'(x)} \tag{10}$$

where

$$\frac{1}{\phi'(x)}=x(1-x). \tag{11}$$

The modified sinc basis functions is shown in Figure 2.

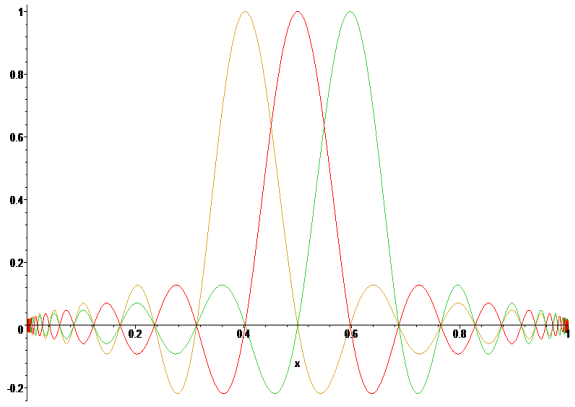


Figure 2. The modified sinc basis on (0, 1) [32].

For the transient space, we generate an approximation via defining the function

$$w=\gamma(r)=\ln(r). \tag{12}$$

Here, w is a conformal mapping from D_W, the wedge-shaped domain in the r-plane onto the infinite strip, D_S, where

$$D_W=\left\{r=t+is:|\arg(r)|<d<\frac{\pi}{2}\right\}. \tag{13}$$

For the SGM, the basis functions are determined from composite translated functions,

$$S(k,h_t)\circ\gamma(t)=Sinc\left(\frac{\gamma(t)-kh_t}{h_t}\right),\quad k=-N_t,.....,N_t. \tag{14}$$

The functions are given in Figure 3 for real values of t.

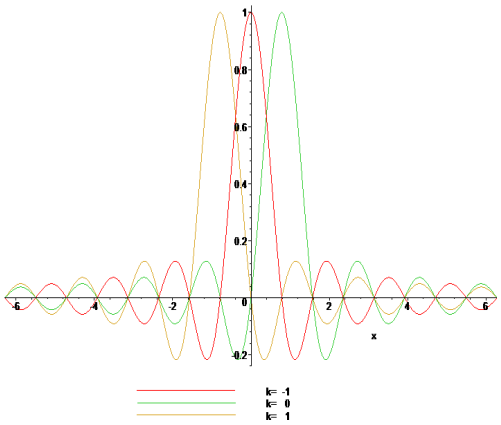


Figure 3. Adjacent members of $S(k,h)\circ\gamma(t)$ when $k=-1,0,1$ and $h=\frac{\pi}{8}$on $(0,\infty)$ [7].

In (14), $w=\gamma(r)$ and $\gamma^{-1}(w)=r=e^w$. We may define γ^{-1}on the real line as

$$\Gamma_2=\left\{\gamma^{-1}(u)\in D_w:-\infty<u<\infty\right\}. \tag{15}$$

For the evenly spaced nodes $\{kh\}_{k=-\infty}^{\infty}$ on the real line, the image which corresponds to these nodes is denoted by

$$t_k=\gamma^{-1}(kh)=e^{kh}, \tag{16}$$

where $0<t_k<\infty$, for all k. A list of conformal mappings may be found in Table 1 below, [9].

Table 1. Conformal mappings and nodes for several subintervals of R.

(a,b)		$\phi(z)$	z_k
a	b	$\ln\left(\frac{z-a}{b-z}\right)$	$\frac{a+be^{kh}}{1+e^{kh}}$
0	1	$\ln\left(\frac{z}{1-z}\right)$	$\frac{e^{kh}}{1+e^{kh}}$
0	∞	$\ln(z)$	e^{kh}
0	∞	$\ln(\sinh(z))$	$\ln\left(e^{kh}+\sqrt{e^{2kh}+1}\right)$
$-\infty$	∞	z	kh
$-\infty$	∞	$\sinh^{-1}(z)$	kh

Definition 1. *Let $B(D_E)$ be the class of functions F that are analytic in D_E and satisfy*

$$\int_{\psi(L+u)}|F(z)|\,dz\to 0,\quad \text{as } u=\mp\infty, \tag{17}$$

where

$$L=\left\{iy:|y|<d\le\frac{\pi}{2}\right\}, \tag{18}$$

and on the boundary of D_E satisfy

$$T(F) = \int_{\partial D_E} |F(z)dz| < \infty. \tag{19}$$

The proof of following theorems can be found in [1].

Theorem 1. *Let* Γ *be* $(0,1)$, $F \in B(D_E)$, *then for* $h > 0$ *sufficiently small*

$$\int_\Gamma F(z)dz - h\sum_{j=-\infty}^{\infty} \frac{F(z_j)}{\phi'(z_j)} = \frac{i}{2}\int_{\partial D} \frac{F(z)k(\phi,h)(z)}{\sin(\pi\phi(z)/h)}dz, \equiv I_F, \tag{20}$$

where

$$|k(\phi,h)|_{z\in\partial D} = \left| e^{\left[\frac{i\pi\phi(z)}{h} sgn(Im\phi(z))\right]} \right|_{z\in\partial D} = e^{\frac{-\pi d}{h}}. \tag{21}$$

For the SGM, the infinite quadrature rule must be truncated to a finite sum; the following theorem demonstrates the conditions under which exponential convergence results.

Theorem 2. *If there exist positive constants* α, β *and* C *such that*

$$\left|\frac{F(x)}{\phi'(x)}\right| \le C \begin{cases} e^{-\alpha|\phi(x)|}, x \in \psi((-\infty,\infty)) \\ e^{-\beta|\phi(x)|}, x \in \psi((0,\infty)), \end{cases} \tag{22}$$

then the error bound for the quadrature rule (20) *is*

$$\left|\int_\Gamma F(x)dx - h\sum_{j=-N}^{N} \frac{F(x_j)}{\phi'(x_j)}\right| \le C\left(\frac{e^{-\alpha Nh}}{\alpha} + \frac{e^{-\beta Nh}}{\beta}\right) + |I_F|. \tag{23}$$

The infinite sum in (20) is truncated with the use of (21) to arrive at this inequality (23). Making the selections

$$h = \sqrt{\frac{\pi d}{\alpha N}}, \tag{24}$$

$$N \equiv \left\| \frac{\alpha N}{\beta} + 1 \right\|, \tag{25}$$

where $\|.\|$ is integer part of statement, then

$$\int_\Gamma F(x)dx = h\sum_{j=-N}^{N} \frac{F(x_j)}{\phi'(x_j)} + O\left(e^{-(\pi\alpha dN)^{1/2}}\right). \tag{26}$$

Theorems 1 and 2 can be used to approximate the integrals that arise in the formulation of the discrete systems.

2.2. Discrete solutions scheme for hyperbolic PDEs

In ordinary differential equations

$$Lu = f, \tag{27}$$

on Γ_1, sinc solution is assumed as an approximate solution u_m in the form of series which $m = 2N+1$ terms

$$u_m(z) = \sum_{j=-N}^{N} c_j S(j,h)\circ\phi(z). \tag{28}$$

The coefficients $\{c_j\}_{j=-N}^{N}$ are determined by orthogonalizing the residual $Lu - f$ with respect to the sinc basis functions $\{S_j\}_{j=-N}^{N}$ where $S_j(z) = S(j,h)\circ\phi(z)$. An inner product for two continuous function such as f_1 and f_2 can be given by the following formula

$$\langle f_1, f_2\rangle = \int_\Gamma f_1 f_2 w dz, \tag{29}$$

where w is the weight function and chosen depending on boundary conditions. If we implement above inner product rule in orthogonalization this yields the discrete sinc-Galerkin system

$$\int_\Gamma (Lu_m - f)(z) S(k,h)\circ\phi(z)\cdot w(z)dz = 0, \quad -N \le k \le N. \tag{30}$$

Now, we are going to derive discrete sinc-Galerkin system for PDEs. Let we assume u_{m_z,m_t} is the approximate solution of equation (1). Then, the discrete system takes the following form

$$u_{m_z,m_t}(z,t) = \sum_{j=-N}^{N}\sum_{k=-N}^{N} c_{jk} S(j,h)\circ\phi(z) \cdot S(k,h)\circ\gamma(t). \tag{31}$$

The coefficients $\{c_{jk}\}_{j,k=-N}^{N}$ are determined by orthogonalizing the residual $Lu_{m_z,m_t} - f$ with respect to the sinc basis functions $\{S_jS_k\}_{j,k=-N}^{N}$ where $S_jS_k(z,t) = S(j,h)\circ\phi(z)S(k,h)\circ\gamma(t)$ for $-N \le j,k \le N$. In this case the inner product takes the following form

$$\langle f_1, f_2\rangle = \int_{\Gamma_t}\int_{\Gamma_z} f_1(z,t) f_2(z,t) w(z,t) dz dt. \tag{32}$$

The choice of the weight function $w(z,t)$ in the double integrand depends on the boundary conditions, the domain, and the partial differential equation. Therefore the discrete Galerkin system is

$$\int_{\Gamma_t}\int_{\Gamma_z}\left(Lu_{m_z m_t}-f\right)(z,t)\cdot S(j,h)\circ\phi(z) \cdot S(k,h)\circ\gamma(t)\cdot w(z,t)dzdt=0. \quad (33)$$

2.3. Matrix representation of the derivatives of sinc basis functions at nodal points

The sinc-Galerkin method actually requires the evaluated derivatives of sinc basis functions at the sinc nodes$z = z_j$. The *rth* derivative of $S_k(z) = S(k,h)\circ\phi(z)$with respect to$\phi$, evaluated at the nodal point $z_j$is denoted by

$$\frac{1}{h^r}\delta_{jk}^{(r)} = \frac{d^r}{d\phi^r}\left(S(k,h)\circ\phi(z)\right)\Big|_{z=z_j.} \quad (34)$$

Here, for each k and j can be stored in a matrix$I^{(r)} = \left[\delta_{jk}^{(r)}\right]$. For $r = 0, 1, 2, ...$

$$I^{(0)} = \delta_{jk}^{(0)} = \left[S(j,h)\circ\phi(x)\right]|_{x=x_k} = \begin{cases} 1, k=j \\ 0, k\neq j, \end{cases} \quad (35)$$

$$I^{(1)} = \delta_{jk}^{(1)} = h\frac{d}{d\phi}\left[S(j,h)\circ\phi(x)\right]\Big|_{x=x_k} = \begin{cases} 0, k=j \\ \frac{(-1)^{k-j}}{(k-j)}, k\neq j, \end{cases} \quad (36)$$

$$I^{(2)} = \delta_{jk}^{(2)} = h\frac{d^2}{d\phi^2}\left[S(j,h)\circ\phi(x)\right]\Big|_{x=x_k} = \begin{cases} \frac{-\pi^2}{3}, k=j \\ \frac{-2(-1)^{k-j}}{(k-j)^2}, k\neq j, \end{cases} \quad (37)$$

where

$$I_m^{(0)} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ : & : & : & \ddots & : \\ 0 & 0 & 0 & ... & 1 \end{bmatrix} = \left[\delta_{jk}^{(0)}\right], \quad (38)$$

$$I_m^{(1)} = \begin{bmatrix} 0 & -1 & \frac{1}{2} & ... & \frac{1}{2N} \\ 1 & 0 & -1 & ... & -\frac{1}{2N-1} \\ -\frac{1}{2} & 1 & 0 & ... & \frac{1}{2N-2} \\ : & : & : & \ddots & : \\ -\frac{1}{2N} & \frac{1}{2N-1} & \frac{1}{2N-2} & ... & 0 \end{bmatrix} = \left[\delta_{jk}^{(1)}\right], \quad (39)$$

$$I_m^{(2)} = \begin{bmatrix} -\frac{\pi^2}{3} & \frac{2}{1^2} & -\frac{2}{2^2} & ... & -\frac{2}{(2N)^2} \\ \frac{2}{1^2} & -\frac{\pi^2}{3} & \frac{2}{1^2} & ... & \frac{2}{(2N-1)^2} \\ -\frac{2}{2^2} & \frac{2}{1^2} & -\frac{\pi^2}{3} & ... & -\frac{2}{(2N-2)^2} \\ : & : & : & \ddots & : \\ -\frac{2}{(2N)^2} & \frac{2}{(2N-1)^2} & -\frac{2}{(2N-2)^2} & ... & -\frac{\pi^2}{3} \end{bmatrix} = \left[\delta_{jk}^{(2)}\right]. \quad (40)$$

The chain rule has been used for the z-derivative of product sinc functions. For example, when $S_j(z) = S(j,h)\circ\phi(z)$;

$$\frac{d\left(S_j(z)w(z)\right)}{dz} = \left(\frac{dS_j(z)}{d\phi(z)}\cdot\frac{d\phi(z)}{dz}\right)w(z) + S_j(z)\frac{dw(z)}{dz} = \frac{dS_j(z)}{d\phi}\phi'(z)w(z) + S_j(z)w'(z), \quad (41)$$

and

$$\frac{d^2\left(S_j(z)w(z)\right)}{dz^2} = \frac{d}{dz}\left(\frac{dS_j(z)}{d\phi}\phi'(z)w(z) + S_j(z)w'(z)\right) = \frac{d^2S_j(z)}{d\phi^2}\left(\phi'(z)\right)^2 w(z) + \frac{dS_j(z)}{d\phi}\phi''(z)w(z) + 2\cdot\frac{dS_j(z)}{d\phi}\phi'(z)w'(z) + S_j(z)w''(z). \quad (42)$$

Now, we are going to develop discrete form for the equation (1). We choose for special case the parameters as follows for the spatial dimension:

$$\left.\begin{array}{l} \phi(z) = \ln\left(\frac{z}{1-z}\right), \\ w_X(z) = \frac{1}{\phi'(z)}, \\ \frac{1}{\phi'(z)} = z(1-z), \end{array}\right\} \quad (43)$$

and for the temporal space as;

$$\left.\begin{array}{l} \gamma(t) = \ln(t), \\ w_T(t) = \frac{1}{\gamma'(t)}, \\ \frac{1}{\gamma'(t)} = t. \end{array}\right\} \quad (44)$$

The discrete form of equation (1) can be given the following form

$$\begin{aligned}\langle Lu-F, S_kS_l\rangle &= \int_{\Gamma_t}\int_{\Gamma_z}(Lu-F)\,S(k,h)\circ\phi(z)\\ &\cdot w_X(x)S(l,s)\circ\gamma(t)\cdot w_T(t)dzdt\\ &= \int_{\Gamma_t}\int_{\Gamma_z}(u_{tt}-u_{xx}-F)\,S(k,h)\circ\phi(z)\\ &\cdot w_X(x)S(l,s)\circ\gamma(t)\cdot w_T(t)dzdt. \end{aligned}\tag{45}$$

We solve this by taking our approximating basis functions to be

$$\begin{cases} S_k(x) = w_X S(k,h)\circ\phi(x),\\ w_X = \frac{1}{\phi'(x)} = x(1-x),\\ \phi(x) = \ln(\frac{x}{1-x}),\\ S_l(t) = w_T S(l,s)\circ\gamma(t),\\ w_T = \frac{1}{\gamma'(t)} = t,\\ \gamma(t) = \ln(t). \end{cases}\tag{46}$$

If we apply sinc-quadrature rules on the definite integral given (45) by using (46) we can get a matrix system. For this purpose, let $A_m(u)$ be a diagonal matrix, whose diagonal elements are $u(x_{-N}), u(x_{-N+1}), ..., u(x_N)$ and non-diagonal elements are zero. Then (45) reproduces following matrixes accordingly:

$$C = \begin{pmatrix} A_1\\ A_2\\ A_3\\ A_4\\ A_5 \end{pmatrix}\tag{47}$$

where

$$\begin{aligned} A_1 &= c_{-N,-N}\ c_{-N,-N+1}\ c_{-N,-N+2}\ \ \dots\ c_{-N,N},\\ A_2 &= c_{-N+1,-N}\ c_{-N+1,-N+1}\ c_{-N+1,-N+2}\ \dots\ c_{-N+1,N},\\ A_3 &= c_{-N+2,-N}\ c_{-N+2,-N+1}\ c_{-N+2,-N+2}\ \dots\ c_{-N+2,N},\\ A_4 &= \vdots\quad \vdots\quad \vdots\quad \ddots\quad \vdots,\\ A_5 &= c_{N,-N}\ c_{N,-N+1}\ \ c_{N,-N+2}\ \dots\ c_{N,N}, \end{aligned}$$

and

$$\begin{cases} B = -2hI_m^{(0)}\left(A_m(w_X)\right) + I_m^{(1)}\left(A_m(w_X')\right) + \frac{I_m^{(2)}}{h},\\ G = A_m(w_T)\left[sI_m^{(0)} - I_m^1\right],\\ D = hA_m(\frac{w_X}{\phi'}),\\ E = sA_m(\frac{w_T}{\gamma'}). \end{cases}\tag{48}$$

Also, for the right side function F given equation (1) can be written as following matrix form;

$$F = \begin{pmatrix} B_1\\ B_2\\ B_3\\ B_4\\ B_5 \end{pmatrix}\tag{49}$$

where

$$\begin{aligned} B_1 &= F_{-N,-N}\ F_{-N,-N+1}\ F_{-N,-N+2}\ \ \dots\ F_{-N,N},\\ B_2 &= F_{-N+1,-N}\ F_{-N+1,-N+1}\ F_{-N+1,-N+2}\ \dots\ F_{-N+1,N},\\ B_3 &= F_{-N+2,-N}\ F_{-N+2,-N+1}\ F_{-N+2,-N+2}\ \dots\ F_{-N+2,N},\\ B_4 &= \vdots\quad \vdots\quad \vdots\quad \ddots\quad \vdots,\\ B_5 &= F_{N,-N}\ F_{N,-N+1}\ \ F_{N,-N+2}\ \dots\ F_{N,N}, \end{aligned}$$

Therefore, we arrive at a matrix system for equation (1) as follows:

$$D^{-1}BC + CGE^{-1} = F\tag{50}$$

Finally, by using Maple Computer Algebra Software, the matrix system (50) can be solved by using LU or QR decomposition method and can be found unknown coefficients. After calculation of C we get approximate solution as follows:

$$u_{x,t} = \sum_{j=-N}^{N}\sum_{k=-N}^{N} c_{jk}S(j,h)\circ\phi(x)\cdot S(k,h)\circ\gamma(t).\tag{51}$$

3. Numerical Examples

In this section, the presented method will be tested on two different problems.

Example 1. *The following hyperbolic equation given*

$$\frac{\partial^2}{\partial t^2}u\,(x,t) - \frac{\partial^2}{\partial x^2}u\,(x,t) = f(x,t),\tag{52}$$

where

$$f\,(x,t) = \frac{e^{-t}\left(A + t^2\left(12 - 11x - 8x^2 + 4x^3\right)\right)}{4\sqrt{t}\sqrt{1-x}},$$

and $A = 3(-1+x)^2x - 12t(-1+x)^2x.$

The exact solution of equation (52) given as follows

$$u\,(x,t) = t^{3/2}e^{-t}x(1-x)^{3/2}.\tag{53}$$

For the equation (52), we choose sinc components here in below:

$$h = s = \frac{0.75}{\sqrt{N}}, x_k = \frac{e^{kh}}{1+e^{kh}}, t_l = e^{sl}, \phi(x) = \ln\left(\frac{x}{1-x}\right),$$
$$\gamma(t) = \ln(t), w_X = \frac{1}{\phi'(x)}, w_T = \frac{1}{\gamma'(t)}. \tag{54}$$

In the light of the above parameters, the numerical results obtained by SGM for equation (52) are indicated in Table 2 and Table 3. Moreover, the graphs of the exact and the approximate solutions for different values are showed in Figure 4, 5 and Figure 6.

Table 2. Numerical results for $N = 5$.

t	x	Exact Sol.	Num. Sol.	Error
0.03	0.3	0.000885	0.002565	0.001679
	0.6	0.000765	0.001809	0.001044
	0.9	0.000143	0.001596	0.001453
0.06	0.3	0.002431	0.014390	0.011958
	0.6	0.002100	0.014650	0.012549
	0.9	0.000393	0.005685	0.005291
0.09	0.3	0.004335	0.018336	0.014000
	0.6	0.003745	0.017395	0.013649
	0.9	0.000702	0.007839	0.007137

Table 3. Numerical results for $N = 20$.

t	x	Exact Sol.	Num. Sol.	Error
0.03	0.3	0.000885	0.001020	0.000134
	0.6	0.000765	0.000474	0.000291
	0.9	0.000143	−0.000060	0.000203
0.06	0.3	0.002431	0.002705	0.000273
	0.6	0.002100	0.001608	0.000492
	0.9	0.000393	0.000033	0.000360
0.09	0.3	0.004335	0.004734	0.000399
	0.6	0.003745	0.003049	0.000695
	0.9	0.000702	0.000186	0.000515

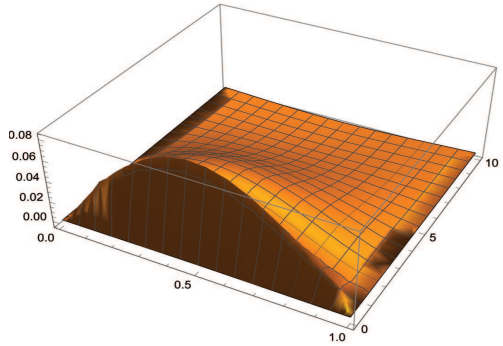


Figure 4. Numerical Simulation of equation (52) according to $N = 5$.

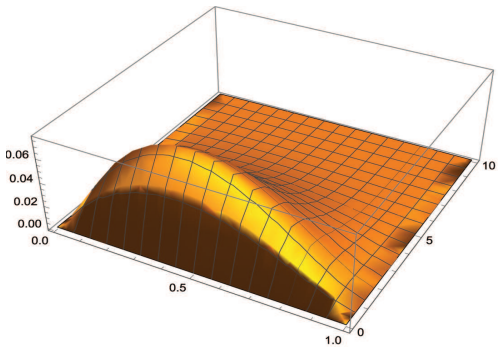


Figure 5. Numerical Simulation of equation (52) according to $N = 20$.

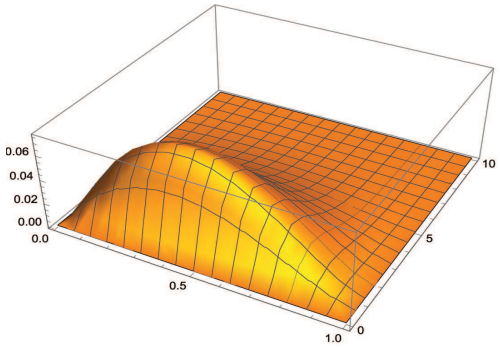


Figure 6. Graph of exact solution of equation (52).

Example 2. *The following hyperbolic equation given*

$$\frac{\partial^2}{\partial t^2}u(x,t) - \frac{\partial^2}{\partial x^2}u(x,t) = f(x,t), \tag{55}$$

where

$$f(x,t) = e^{-t}\left(B + t^2\left(2 + x - x^2\right)\right),$$
and $B = -2(-1+x)x + 4t(-1+x)x.$

The exact solution of equation (55) given as follows

$$u(x,t) = e^{-t}t^2(1-x)x. \tag{56}$$

For the equation (55), we choose sinc components here in below:

$$h = s = \frac{0.75}{\sqrt{N}}, x_k = \frac{e^{kh}}{1+e^{kh}}, t_l = e^{sl},$$
$$\phi(x) = \ln\left(\frac{x}{1-x}\right), \gamma(t) = \ln(t), w_X = \frac{1}{\phi'(x)},$$
$$w_T = \frac{1}{\gamma'(t)}. \tag{57}$$

According to the above parameters, the numerical solutions which are obtained by using the sinc-Galerkin method for equation (55) are presented

in Table 4 and Table 5 for different values. Also the graphs of exact and approximate solutions for different values are presented in Figure 7, 8 and Figure 9.

Table 4. Numerical results for $N = 5$.

t	x	Exact Sol.	Num. Sol.	Error
0.03	0.3	0.000183	0.000078	0.000104
	0.6	0.000209	−0.0023475	0.002556
	0.9	0.000078	0.001439	0.001360
0.06	0.3	0.000711	0.003170	0.002458
	0.6	0.000813	0.000211	0.000601
	0.9	0.000305	0.003259	0.002954
0.09	0.3	0.001554	0.003741	0.002187
	0.6	0.001776	−0.002408	0.004185
	0.9	0.000666	0.005409	0.004743

Table 5. Numerical results for $N = 20$.

t	x	Exact Sol.	Num. Sol.	Error
0.03	0.3	0.000183	0.000180	2.86×10^{-6}
	0.6	0.000209	0.000203	5.75×10^{-6}
	0.9	0.000078	0.000077	1.52×10^{-6}
0.06	0.3	0.000711	0.000711	5.46×10^{-7}
	0.6	0.000813	0.000811	2.43×10^{-6}
	0.9	0.000305	0.000304	4.24×10^{-7}
0.09	0.3	0.001554	0.001554	5.03×10^{-7}
	0.6	0.001776	0.001777	1.12×10^{-6}
	0.9	0.000666	0.000666	2.11×10^{-7}

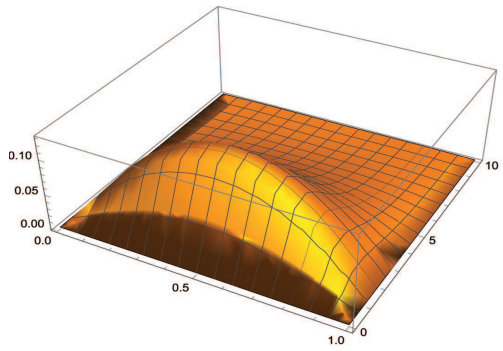


Figure 7. Numerical Simulation of equation (55) according to $N = 5$.

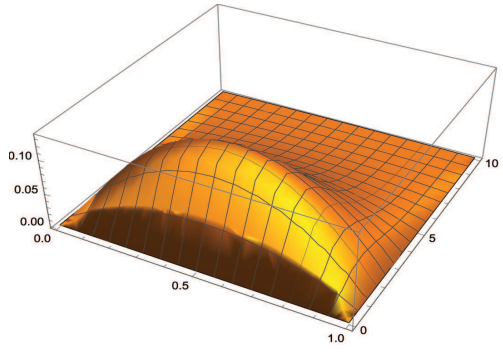


Figure 8. Numerical Simulation of equation (55) according to $N = 20$.

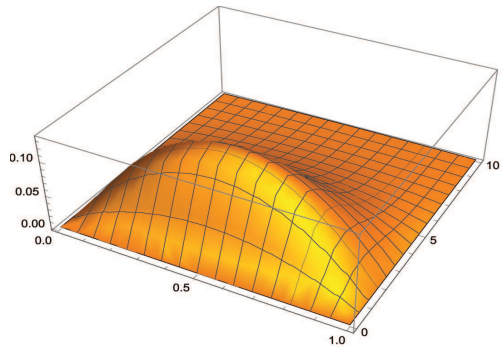


Figure 9. Graph of exact solution of equation (55).

4. Conclusion

The SGM is operated to determine the approximate solutions of second order PDEs. According to the obtained results in numerical examples, sinc-Galerkin method seems to be an efficient method in the sense that selection parameters and changing boundary conditions and also giving different problems to the algorithms. The accuracy of the solutions can be developed by increasing the number of grid points N. In this work, we improve a powerful algorithm for the solution with SGM via Maple. Various PDEs are solved in means of our technique in less than 20 seconds.

References

[1] Stenger, F. (1979). A Sinc-Galerkin Method of Solution of Boundary Value Problems. *Mathematics of Computation*, 33 85-109.

[2] Stenger, F. (2000). Summary of Sinc numerical methods. *Journal of Computational and Applied Mathematics*, 121 379-420.

[3] Alonso, N., Bowers, K.L. (2009). An Alternating-Direction Sinc-Galerkin Method for Elliptic Problems. *Journal of Complexity*, 25, 237–252.

[4] Bowers, K.L., Lund, J. (1987). Numerical Solution of Singular Poisson Problems via the Sinc-Galerkin Method. *SIAM Journal on Numerical Analysis*, 24(1) 36-51.

[5] Koonprasert, S. (2003). The Sinc-Galerkin Method for Problems in Oceanography. PhD thesis, Montana State University, Bozeman, Montana.

[6] Lewis, D.L., Lund, J., Bowers, K. L. (1987). The Space-Time Sinc-Galerkin Method for Parabolic Problems. *International Journal for Numerical Methods in Engineering*, 24 1629-1644.

[7] Lund, J. (1986). Symmetrization of the Sinc-Galerkin Method for Boundary Value Problems. *Mathematics of Computation*, 47, 571-588.

[8] Lund, J., Bowers, K.L., McArthur, K. M., (1989). Symmetrization of the Sinc-Galerkin Method with Block Techniques for Elliptic Equations. *IMA Journal of Numerical Analysis*, 9, 29-46.

[9] Lund, J., Bowers, K.L. (1992). Sinc Methods for Quadrature and Differential Equations. SIAM: Philadelphia.

[10] Lybeck, N.J. (1994). Sinc Domain Decomposition Methods for Elliptic Problems. PhD thesis, Montana State University, Bozeman, Montana.
[11] Lybeck, N.J., Bowers, K.L. (1996). Domain Decomposition in Conjunction with Sinc Methods for Poisson's Equation. *Numerical Methods for Partial Differential Equations*, 12, 461-487.
[12] McArthur, K.M., Bowers, K.L., Lund, J. (1987). Numerical Implementation of the Sinc-Galerkin Method for Second-Order Hyperbolic Equations. *Numerical Methods for Partial Differential Equations*, 3, 169-185.
[13] McArthur, K.M., Bowers, K.L., Lund, J. (1990). The Sinc Method in Multiple Space Dimensions: Model Problems. *NumerischeMathematik*, 56, 789-816.
[14] Morlet, A.C., Lybeck, N.J., Bowers, K.L. (1997). The Schwarz Alternating Sinc Domain Decomposition Method. *Applied Numerical Mathematics*, 25, 461-483.
[15] Morlet, A.C., Lybeck, N.J., Bowers, K.L. (1999). Convergence of the Sinc Overlapping Domain Decomposition Method. *Applied Mathematics and Computation*, 98, 209-227.
[16] Ng, M. (1999). Fast Iterative Methods for Symmetric Sinc-Galerkin Systems. *IMA Journal of Numerical Analysis*, 19 357-373.
[17] Ng, M., Bai, Z. (2003). A Hybrid Preconditioner of Banded Matrix Approximation and Alternating-Direction Implicit Iteration for Symmetric Sinc-Galerkin Linear Systems, Linear Systems, Linear Algebra and its Applications, 366 317-335.
[18] Stenger, F. (1981). Numerical Methods Based on Whittaker Cardinal, or Sinc Functions. *SIAM Review*, 23 165-224.
[19] Stenger, F. (1993). Numerical Methods Based on Sinc and Analytic Functions, Springer-Verlag, New York.
[20] Whittaker, E.T. (1915). On the Functions which are Represented by the Expansions of the Interpolation Theory. *Proceedings of the Royal Society of Edinburg*, 35, 181-194.
[21] Whittaker, J.M. (1961). Interpolation Function Theory, Cambridge Tracts in Mathematics and Mathematical Physics. No. 33, Cambridge University Press, London.
[22] Stenger, F. (1976). Approximations via Whittaker's Cardinal Function. *Journal of Approximation Theory*, 17, 222-240.
[23] Stenger, F., O'Reilly, M. J. (1998). Computing Solutions to Medical Problems via Sinc Convolution. *IEEE Transactions on Automatic Control*, 43, 843.
[24] Narasimhan, S., Majdalani, J., Stenger, F. (2002). A First Step In Applying The Sinc Collocation Method To The Nonlinear Navier Stokes Equations. *Numerical Heat Transfer*,41, 447-462.
[25] Mueller, J.L., Shores, T.S. (2004). A New Sinc-Galerkin Method for Convection-Diffusion Equations with Mixed Boundary Conditions. *Computers and Mathematics with Applications*, 47, 803-822.
[26] El-Gamel, M., Behiry, S.H., Hashish, H. (2003). Numerical method for the solution of special nonlinear fourth-order boundary value problems. *Applied Mathematics and Computation*, 145, 717–734.
[27] Alkan, S., Secer, A. (2015). Solving nonlinear boundary value problems by the Galerkin method with sinc functions. *Open Physics*, 13, 1, 389-394.
[28] Zamani, N.G. (1987). A finite element based collocation method for eigenvalue calculations. *Journal of the Franklin Institute*, 324, 205-217.
[29] Alkan, S., Secer, A. (2015). Application of Sinc-Galerkin method for solving space-fractional boundary value problems. *Mathematical Problems in Engineering*, 2015, 1-10.
[30] Lybeck, N.J., Bowers, K.L. (1996). Sinc methods for domain decomposition. *Applied Mathematics and Computation*, 75, 4-13.
[31] Secer, A., Alkan, S., Akinlar, M.A., Bayram, M. (2013). Sinc-Galerkin method for approximate solutions of fractional order boundary value problems. *Boundary Value Problems*, 2013, 1, 281.
[32] Secer, A., Kurulay M., Bayram, M., Akinlar, M.A. (2013). An efficient computer application of the sinc-Galerkin approximation for nonlinear boundary value problems, Boundary Value Problems, 2012, 1, 117.

Aydin Secer *is an Associate Professor of Mathematical Engineering at Yildiz Technical University. He completed his M.Sc. and Ph.D. at Ataturk University, in 2005 and 2011, respectively. Upon completion of his Ph.D. he worked as an assistant professor at department of Mathematical Engineering at Yildiz Technical University. He is now working as Associate Professor at the same department. His research interests are Numerical and Computational Mathematics, Computer Sciences and Graph Matching Algorithms.*

An International Journal of Optimization and Control: Theories & Applications
(http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.8, No.2, pp.259-265 (2018)
https://doi.org/10.11121/ijocta.01.2018.00565

RESEARCH ARTICLE

Some new integral inequalities for Lipschitzian functions

İmdat İşcan[a], Mahir Kadakal[a *] and Alper Aydın[b]

[a] *Department of Mathematics, Sciences and Arts Faculty, Giresun University, Güre Campus, Giresun, Turkey*
[b] *Instutite of Sciences, Giresun University, Güre Campus, Giresun, Turkey*
imdati@yahoo.com, mahirkadakal@gmail.com, mbatin28@gmail.com

ARTICLE INFO

Article history:
Received: 29 November 2017
Accepted: 18 February 2018
Available Online: 24 July 2018

Keywords:
Integral inequalities
Lipschitzian function
Hermite-Hadamard inequality
Integral means

AMS Classification 2010:
26A51, 26D10, 26D15

ABSTRACT

This paper is about obtaining some new type of integral inequalities for functions from the Lipschitz class. For this, some new integral inequalities related to the differences between the two different types of integral averages for Lipschitzian functions are obtained. Moreover, applications for some special means as arithmetic, geometric, logarithmic, p-logarithmic, harmonic, identric are given.

1. Introduction

Convexity theory has appeared as a powerful technique to study a wide class of unrelated problems in pure and applied sciences. It is well known that theory of convex sets and convex functions play an important role in mathematics and the other pure and applied sciences.

Definition 1. A function $f: I \subseteq \mathbb{R} \to \mathbb{R}$ is said to be convex if the inequality

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

is valid for all $x, y \in I$ and $t \in [0,1]$. If this inequality reverses, then f is said to be concave on interval $I \neq \emptyset$. This definition is well known in the literature.

The research of beautiful inequalities which have symmetry is very interesting and important to Analysis and PDE. A well-known example is the famous Hermite-Hadamard inequality which was first published in [1].

If $f: I \to \mathbb{R}$ is a convex function on the interval I, then for any $a, b \in I$ with $a \neq b$ we have the following double inequality

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2} \quad (1)$$

This double inequality is known as Hermite-Hadamard integral inequality for convex functions in the literature. Note that some of the classical inequalities for means can be derived from (1) for appropriate particular selections of the mapping f. Both inequalities hold in the reversed direction if mapping f is concave.

Definition 2. [2] Let $I \subset \mathbb{R}\backslash\{0\}$ be a real interval. A function $f: I \to \mathbb{R}$ is said to be harmonically convex, if

$$f\left(\frac{xy}{tx + (1-t)y}\right) \leq tf(y) + (1-t)f(x)$$

for all $x, y \in I$ and $t \in [0,1]$. If this inequality is reversed, then the function f is said to be harmonically concave.

Definition 3. [2] Let $f: I \subseteq \mathbb{R}\backslash\{0\} \to \mathbb{R}$ be a harmonically convex function and $a, b \in I$ with $a < b$, If $f \in L[a,b]$ then the following inequalities hold:

$$f\left(\frac{2ab}{a+b}\right) \leq \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \leq \frac{f(a)+f(b)}{2} \quad (2)$$

* Corresponding author

Definition 4. (Beta function) The beta function denoted by $\beta(m,n)$ is defined as

$$\beta(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx.$$

Definition 5. (Hypergeometric function) [3] The hypergeometric function denoted by ${}_2F_1\ (a,b;c;z)$ is defined by the integral equality

$${}_2F_1\ (a,b;c;z)$$
$$=\frac{1}{\beta(b,c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}\,(1-zt)^{-a}dt,$$
$$c>b>0,|z|<1.$$

Definition 6. (M-Lipschitz Condition) [4] $f:I\to\mathbb{R}$ is said to satisfy the Lipschitz condition if there is a constant $M>0$ such that

$$|f(x)-f(y)|\le M|x-y|\quad \forall x,y\in I.$$

Theorem 1. [4] If $f:I\to\mathbb{R}$ is convex, then f satisfies a Lipschitz condition on any closed interval $[a,b]$ contained in the interior I° of I. Consequently, f is absolutely continuous on $[a,b]$ and continuous on I°.

In [5], the inequalities related to left-hand side and right-hand side of the inequality (1) for Lipschitzian mappings as follow:

Theorem 2. [5] Let $f:I\subseteq\mathbb{R}\to\mathbb{R}$ be an M-Lipschitzian mapping on I and $a,b\in I$ with $a<b$. Then we have the inequalities

$$\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(x)dx\right|\le\frac{M}{4}(b-a)$$

and

$$\left|\frac{f(a)+f(b)}{2}-\frac{1}{b-a}\int_a^b f(x)dx\right|\le\frac{M}{3}(b-a).$$

Corollary 1. Let $f:\mathrm{I}\subseteq\mathbb{R}\to\mathbb{R}$ be a convex and differentiable function on interval I and $a,b\in I$ with $a<b$ and $M=sup_{t\in[a,b]}|f'(t)|<\infty$. Then we have the inequalities

$$0\le\frac{1}{b-a}\int_a^b f(x)dx-f\left(\frac{a+b}{2}\right)\le\frac{M}{4}(b-a)$$

and

$$0\le\frac{f(a)+f(b)}{2}-\frac{1}{b-a}\int_a^b f(x)dx\le\frac{M}{3}(b-a)$$

See [5-8] and references therein for more information about the Hadamard-type inequalities for the Lipschitzian functions.

2. Main results

In this section, we obtain some new inequalities related to integral means given in the inequalities (1) and (2) for Lipschitzian mappings.

Theorem 3. Let $f:I\subseteq(0,\infty)\to\mathbb{R}$ be an M-Lipschitzian mapping on interval I and $a,b\in I$ with $a<b$. Then following inequality holds:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{ab}{b-a}\int_a^b\frac{f(u)}{u^2}du\right|$$
$$\le\frac{M(b-a)^2}{6b}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right)$$

Proof: Since f is M-Lipschitzian function on interval I, for $\forall x,y\in[a,b]$

$$|f(x)-f(y)|\le M|x-y|.$$

Here, for arbitrary $t\in[0,1]$, if we take

$$x=tb+(1-t)a,\qquad y=\frac{ab}{ta+(1-t)b}$$

then

$$\left|f[(tb+(1-t)a]-f\left(\frac{ab}{ta+(1-t)b}\right)\right|$$
$$\le M\left|(tb+(1-t)a-\frac{ab}{ta+(1-t)b}\right|$$
$$=M\left|\frac{t^2ab+t(1-t)(b^2+a^2)+(1-t)^2ab-ab}{ta+(1-t)b}\right|$$
$$=\frac{Mt(1-t)(b-a)^2}{ta+(1-t)b}.$$

Consequently, we get the following inequality:

$$\left|f(tb+(1-t)a-f\left(\frac{ab}{ta+(1-t)b}\right)\right|$$
$$\le\frac{Mt(1-t)(b-a)^2}{b-t(b-a)}$$

If we take integral the last inequality on $t\in[0,1]$ and use property of modulus, we have

$$\left|\int_0^1 f[tb+(1-t)a]dt-\int_0^1 f\left(\frac{ab}{ta+(1-t)b}\right)dt\right|$$
$$\le\int_0^1\left|f(tb+(1-t)a-f\left(\frac{ab}{ta+(1-t)b}\right)\right|dt$$
$$\le M(b-a)^2\int_0^1\frac{t(1-t)}{b\left[1-t\left(1-\frac{a}{b}\right)\right]}dt$$

$$= \frac{M(b-a)^2}{b}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right)\beta(2,2)$$

If we make the change of variables $u = tb + (1-t)a$ and $u = \frac{ab}{ta+(1-t)b}$ in the integrals on the left side of the last inequality respectively, we have the following inequality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du - \frac{ab}{b-a}\int_a^b \frac{f(u)}{u^2}du\right| \leq \frac{M(b-a)^2}{6b}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right)$$

This completes the proof of theorem.

Proposition 1. Let $p \in (1,\infty)\backslash\{2\}$ and $a,b \in \mathbb{R}$ with $0 < a < b$. Then

$$\left|L_p^p - G^2 L_{p-2}^{p-2}\right| \leq \frac{pb^{p-2}(b-a)^2}{6}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right),$$

where $G = G(a,b)$ and $L_p = L_p(a,b)$ are geometric and p-logarithmic means respectively.

Proof: If the $f(x) = x^p$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 3, the inequality

$$\left|\frac{1}{b-a}\int_a^b f(u)du - \frac{ab}{b-a}\int_a^b \frac{f(u)}{u^2}du\right| = \left|\frac{1}{b-a}\int_a^b u^p du - \frac{ab}{b-a}\int_a^b \frac{u^p}{u^2}du\right|$$

is obtained. If the integral is calculated,

$$\left|\frac{1}{b-a}\int_a^b u^p du - \frac{ab}{b-a}\int_a^b \frac{u^p}{u^2}du\right| = \left|\frac{b^{p+1}-a^{p+1}}{(b-a)(p+1)} - \frac{ab(b^{p-1}-a^{p-1})}{(b-a)(p-1)}\right| = \left|L_p^p - G^2 L_{p-2}^{p-2}\right|.$$

From Corollary 1, if $M = sup_{t\in[a,b]}|f'(t)| < \infty$ is taken for the right side of the inequality, then $M = pb^{p-1}$. So, we get

$$\left|L_p^p - abL_{p-2}^{p-2}\right| \leq \frac{pb^{p-2}(b-a)^2}{6}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right).$$

Proposition 2. Let $p \in (1,\infty)\backslash\{2\}$ and $a,b \in \mathbb{R}$ with $0 < a < b$. Then

$$\left|L^{-1} - \frac{A}{G^2}\right| \leq \frac{(b-a)^2}{6ba^2}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right),$$

where $G = G(a,b)$, $A = A(a,b)$ and $L = L(a,b)$ are geometric, arithmetic and logarithmic means respectively.

Proof: If the $f(x) = \frac{1}{x}$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 3, we have the following equality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du - \frac{ab}{b-a}\int_a^b \frac{f(u)}{u^2}du\right| = \left|\frac{1}{b-a}\int_a^b u^{-1}du - \frac{ab}{b-a}\int_a^b u^{-3}du\right| = \left|L^{-1} - \frac{A}{G^2}\right|.$$

Using the Corollary 1, if $M = sup_{t\in[a,b]}|f'(t)| < \infty$ is taken for the right side of the inequality, then $M = \frac{1}{a^2}$. So, we get

$$\left|L^{-1} - \frac{A}{G^2}\right| \leq \frac{(b-a)^2}{6ba^2}\ {}_2F_1\left(1,2;4;1-\frac{a}{b}\right).$$

Theorem 4. Let $f: I \subset \mathbb{R} \to \mathbb{R}$ be an M-Lipschitzian function on interval I and $a,b,x,y \in I$ with $a \leq x < y$ and $a < b$. Then following inequality holds:

$$\left|\frac{1}{b-a}\int_a^b f(u)du - \frac{1}{y-x}\int_x^y f(u)du\right| \leq \frac{M}{2}[|b-y| + x - a].$$

Proof: Since f is an M-Lipschitzian function on interval I, for $\forall v,w \in I$

$$|f(v) - f(w)| \leq M|v-w|.$$

Here, for arbitrary $t \in [0,1]$, if we take

$$v = tb + (1-t)a, \qquad w = ty + (1-t)x,$$

then

$$|f[tb + (1-t)a] - f[ty + (1-t)x]| \leq M|t(b-y) + (1-t)(a-x)| \leq M[t|b-y| + (1-t)|a-x|].$$

If we take integral the last inequality on $t \in [0,1]$ and use the property of modulus, we have

$$\left|\int_0^1 f[tb+(1-t)a]dt-\int_0^1 f[ty+(1-t)x]dt\right|$$
$$\leq \left|\int_0^1 (f[tb+(1-t)a]-f[ty+(1-t)x])dt\right|$$
$$\leq \int_0^1 |f[tb+(1-t)a]-f[ty+(1-t)x]|dt$$
$$\leq M\int_0^1 [t|b-y|+(1-t)(x-a)]dt.$$

If we make the change of variables $u=tb+(1-t)a$ and $u=ty+(1-t)x$ in the integrals on the left side of the last inequality respectively, we have the following inequality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{1}{y-x}\int_x^y f(u)du\right| \leq \frac{M}{2}[|b-y|+x-a].$$

This completes the proof of theorem.

Proposition 3. Let $p>1$ and $a,b,x,y\in\mathbb{R}$ with $0<a\leq x<y$ and $a<b$. Then

$$\left|L_p(a,b)-L_p(x,y)\right| \leq \frac{pb^{p-1}}{2}[|b-y|+x-a],$$

where $L_p=L_p(a,b)$ is p-logarithmic mean.

Proof: If the $f(x)=x^p$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 4, we have the following equality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{1}{y-x}\int_x^y f(u)du\right|$$
$$=\left|\frac{1}{b-a}\int_a^b u^p du-\frac{1}{y-x}\int_x^y u^p du\right|$$
$$=\left|\frac{b^{p+1}-a^{p+1}}{(b-a)(p+1)}-\frac{y^{p+1}-x^{p+1}}{(y-x)(p+1)}\right|$$
$$=\left|L_p(a,b)-L_p(x,y)\right|$$
$$\leq \frac{pb^{p-1}}{2}[|b-y|+x-a],$$

where $M=pb^{p-1}$.

Proposition 4. Let $a,b,x,y\in\mathbb{R}$ with $0<a\leq x<y$ and $a<b$. Then

$$|L^{-1}(a,b)-L^{-1}(x,y)| \leq \frac{1}{2a^2}[|b-y|+x-a],$$

where $L=L(a,b)$ is logarithmic mean.

Proof: If the $f(x)=\frac{1}{x}$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 4, we have the following equality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{1}{y-x}\int_x^y f(u)du\right|$$
$$=\left|\frac{1}{b-a}\int_a^b \frac{1}{x}dx-\frac{1}{y-x}\int_x^y \frac{1}{x}dx\right|$$
$$=\left|\frac{1}{b-a}ln\frac{b}{a}-\frac{1}{y-x}ln\frac{y}{x}\right|$$
$$=|L^{-1}(a,b)-L^{-1}(x,y)|.$$

From Corollary 1, since $M=\frac{1}{a^2}$, the following inequality

$$|L^{-1}(a,b)-L^{-1}(x,y)| \leq \frac{1}{2a^2}[|b-y|+x-a]$$

is obtained.

Proposition 5. Let $a,b,x,y\in\mathbb{R}$ with $0<a\leq x<y$ and $a<b$. Then

$$\left|\frac{e^b-e^a}{b-a}-\frac{e^y-e^x}{y-x}\right| \leq \frac{e^b}{2}[|b-y|+x-a].$$

Proof: If the $f(x)=e^x$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 4, we have the following inequality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{1}{y-x}\int_x^y f(u)du\right|$$
$$=\left|\frac{1}{b-a}\int_a^b e^u du-\frac{1}{y-x}\int_x^y e^u du\right|$$
$$=\left|\frac{e^b-e^a}{b-a}-\frac{e^y-e^x}{y-x}\right|$$
$$\leq \frac{e^b}{2}[|b-y|+x-a].$$

Proposition 6. Let $a,b,x,y\in\mathbb{R}$ with $0<a\leq x<y$ and $a<b$. Then

$$|lnI(x,y)-lnI(a,b)| \leq \frac{1}{2a}[|b-y|+x-a],$$

where $I=I(a,b)$ is identric mean.

Proof: If the $f(x)=-lnx$ convex mapping defined

on interval $[a,b]$ is applied to the left side of the inequality in Theorem 4, we have the following equality:

$$\left|\frac{1}{b-a}\int_a^b f(u)du-\frac{1}{y-x}\int_x^y f(u)du\right|$$
$$=\left|\frac{1}{b-a}\int_a^b -lnudu-\frac{1}{y-x}\int_x^y -lnudu\right|$$

$$=\left|1-\frac{lnb^b-lna^a}{b-a}-\left(1-\frac{lny^y-lnx^x}{y-x}\right)\right|$$
$$=\left|-\left[ln\frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}}\right]+\left[ln\frac{1}{e}\left(\frac{y^y}{x^x}\right)^{\frac{1}{y-x}}\right]\right|$$
$$=|lnI(x,y)-lnI(a,b)|.$$

From Corollary 1, since $M=\frac{1}{a}$, the following inequality

$$|lnI(x,y)-lnI(a,b)|\le\frac{1}{2a}[|b-y|+x-a].$$

Theorem 5. Let $f:\mathrm{I}\subset(0,\infty)\to\mathbb{R}$ be an M-Lipschitzian function on interval I and $a,b,x,y\in I$ with $a\le x<y$ and $a<b$. Then following inequality holds:

$$\left|\frac{ab}{b-a}\int_a^b\frac{f(u)}{u^2}du-\frac{xy}{y-x}\int_x^y\frac{f(u)}{u^2}du\right|$$
$$\le\frac{M}{b-a}\{ax|b-y|[bL^{-1}(ay,bx)-L^{-1}(x,y)]$$
$$+by|a-x|[L^{-1}(x,y)-aL^{-1}(ay,bx)]\}$$

Proof: Since f is an M-Lipschitzian function on interval I, for $\forall v,w\in I$

$$|f(v)-f(w)|\le M|v-w|.$$

Here, for arbitrary $t\in[0,1]$, if we take

$$v=\frac{ab}{ta+(1-t)b},\qquad w=\frac{xy}{tx+(1-t)y}$$

then

$$\left|f\left(\frac{ab}{ta+(1-t)b}\right)-f\left(\frac{xy}{tx+(1-t)y}\right)\right|$$
$$\le M\left|\frac{ab}{ta+(1-t)b}-\frac{xy}{tx+(1-t)y}\right|$$
$$\le M\frac{tax|b-y|+(1-t)by|a-x|}{[ta+(1-t)b][tx+(1-t)y]}.$$

If we take integral the last inequality on $t\in[0,1]$ and use the property of modulus and changing variable, we have

$$\left|\frac{ab}{b-a}\int_a^b\frac{f(u)}{u^2}du-\frac{xy}{y-x}\int_x^y\frac{f(u)}{u^2}du\right|$$
$$\le M\left\{ax|b-y|\int_0^1\frac{t}{[ta+(1-t)b][tx+(1-t)y]}dt\right.$$
$$\left.+by|a-x|\int_0^1\frac{1-t}{[ta+(1-t)b][tx+(1-t)y]}dt\right\}$$
$$=M\left\{ax|b-y|\int_0^1\frac{tdt}{[b+t(a-b)][y+t(x-y)]}\right.$$
$$\left.+by|a-x|\int_0^1\frac{t}{[a+t(b-a)][x+t(y-x)]}dt\right\}$$
$$\le\left\{\frac{ax|b-y|}{(b-a)(y-x)}\int_0^1\frac{tdt}{\left[t+\frac{b}{a-b}\right]\left[t+\frac{y}{x-y}\right]}\right.$$
$$\left.+\frac{by|a-x|}{(b-a)(y-x)}\int_0^1\frac{tdt}{\left[t+\frac{a}{b-a}\right]\left[t+\frac{x}{y-x}\right]}\right\}\qquad(2)$$

If the integrals in (2) are calculated, we get

$$\int_0^1\frac{t}{\left[t+\frac{b}{a-b}\right]\left[t+\frac{y}{x-y}\right]}dt$$
$$=\frac{1}{(b-a)(y-x)}\left[ln\frac{x}{y}+\frac{b(y-x)}{ay-bx}ln\frac{ay}{bx}\right]$$
$$=\frac{1}{b-a}[bL^{-1}(ay,bx)-L^{-1}(x,y)],\qquad(3)$$

and

$$\int_0^1\frac{t}{\left[t+\frac{a}{b-a}\right]\left[t+\frac{x}{y-x}\right]}dt$$
$$=\frac{1}{b-a}[L^{-1}(x,y)-aL^{-1}(ay,bx)].\qquad(4)$$

By substituting (3) and (4) in (2), desired result can be obtained.

This completes the proof of theorem.

Proposition 7. Let $p\in(1,\infty)\backslash\{2\}$ and $a,b\in\mathbb{R}$ with $0<a<b$. Then

$$\left|G^2(a,b)L_{p-2}^{p-2}(a,b)-G^2(x,y)L_{p-2}^{p-2}(x,y)\right|$$
$$\le\frac{pb^{p-1}}{b-a}\{ax|b-y|[bL^{-1}(ay,bx)-L^{-1}(x,y)]$$

$+by|a-x|[L^{-1}(x,y)-aL^{-1}(ay,bx)]\},$

where $G=G(a,b)$, $L_p=L_p(a,b)$ and $L=L(a,b)$are geometric, p-logarithmic and logarithmic means respectively.

Proof: If the$f(x)=x^p$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 5, we have the following equality:

$$\left|\frac{ab}{b-a}\int_a^b\frac{f(u)}{u^2}du-\frac{xy}{y-x}\int_x^y\frac{f(u)}{u^2}du\right|$$
$$=\left|\frac{ab}{b-a}\int_a^b\frac{u^p}{u^2}dx-\frac{xy}{y-x}\int_x^y\frac{u^p}{u^2}dx\right|$$
$$=\left|\frac{ab(b^{p-1}-a^{p-1})}{(b-a)(p-1)}-\frac{xy(x^{p-1}-y^{p-1})}{(y-x)(p-1)}\right|$$

From Corollary 1, since $M=pb^{p-1}$, the following inequality

$$\left|\frac{ab(b^{p-1}-a^{p-1})}{(b-a)(p-1)}-\frac{xy(x^{p-1}-y^{p-1})}{(y-x)(p-1)}\right|$$
$$=\left|G^2(a,b)L_{p-2}^{p-2}(a,b)-G^2(x,y)L_{p-2}^{p-2}(x,y)\right|$$
$$\leq\frac{pb^{p-1}}{b-a}\{ax|b-y|[bL^{-1}(ay,bx)-L^{-1}(x,y)]$$
$$+by|a-x|[L^{-1}(x,y)-aL^{-1}(ay,bx)]\}$$

Proposition 8. Let $p\geq 1$ and $a,b\in\mathbb{R}$ with $0<a<b$. Then

$$|H^{-1}(a,b)-H^{-1}(x,y)|$$
$$\leq\frac{1}{a^2(b-a)}\{ax|b-y|[bL^{-1}(ay,bx)-L^{-1}(x,y)]$$
$$+by|a-x|[L^{-1}(x,y)-aL^{-1}(ay,bx)]\},$$

where $H=H(a,b)$ and $L=L(a,b)$ are harmonic and logarithmic means respectively.

Proof: If the $f(x)=\frac{1}{x}$ convex mapping defined on interval $[a,b]$ is applied to the left side of the inequality in Theorem 5, we have the following equality:

$$\left|\frac{ab}{b-a}\int_a^b\frac{f(u)}{u^2}du-\frac{xy}{y-x}\int_x^y\frac{f(u)}{u^2}du\right|$$
$$=\left|\frac{ab}{b-a}\int_a^b\frac{\frac{1}{x}}{x^2}dx-\frac{xy}{y-x}\int_x^y\frac{\frac{1}{x}}{x^2}dx\right|$$

So, we get

$$|H^{-1}(a,b)-H^{-1}(x,y)|$$
$$\leq\frac{1}{a^2(b-a)}\{ax|b-y|[bL^{-1}(ay,bx)-L^{-1}(x,y)]$$
$$+by|a-x|[L^{-1}(x,y)-aL^{-1}(ay,bx)]\}.$$

3. Conclusions

In this paper, some new type integral inequalities related to the differences between the two different types of integral averages for Lipschitzian functions are obtained. The significance of the obtained inequalities is that: some approaches of the same type averages to each other at different points are given in here first time. Similar studies can also be obtained for fractional integrals.

References

[1] Hadamard J. (1893). Étude sur les propriétés des fonctions entiéres et en particulier d'une fonction considérée par Riemann, J. Math. Pures Appl., 58 (1893), 171–215

[2] İşcan, İ. (2014). Hermite-Hadamard type inequalities for harmonically convex functions, *Hacettepe Journal of Mathematics and Statistics*, Volume 43 (6), 935–942.

[3] Kilbas, A.A. Srivastava, H.M. and Trujillo, J. (2006). *Theory and applications of fractional differential equations*, Elsevier, Amsterdam.

[4] Roberts, A.W. and Varberg, D.E. (1973) *Convex Functions*. Academic Press, 300pp, New York.

[5] Dragomir, S.S. Cho, Y.J. and Kim, S.S. (2000). Inequalities of Hadamard's type for Lipschitzian mappings and their applications, *Journal of Mathematical Analysis and Applications*, vol. 245, no. 2, pp. 489-501.

[6] Hwang, S.R. Hsu, K.C. and Tseng, K.L. (2013). Hadamard-type inequalities for Lipschitzian functions in one and two variables with applications, *Journal of Mathematical Analysis and Applications*, vol. 405, no. 2, 546-554.

[7] İşcan, İ. (2014). New general integral inequalities for Lipschitzian functions via Hadamard fractional integrals, *International Journal of Analysis*, volume 2014, Article ID 353924, 8 pages.

[8] İşcan, İ. (2016). Hadamard-type and Bullen-type inequalities for Lipschitzian functions via fractional integrals, *Mathematical Sciences and Applications E-Notes*, 4 (1), 77-87.

İmdat İşcan *is an Associate Professor at the Department of Mathematics, Art and Science Faculty, Giresun University. He received his BSc (1992) degree from Department of Mathematics Teaching, Faculty of Fatih Education, Karadeniz Technical University and MSc (1996) degree and PhD (2003) from Karadeniz Technical University, Turkey. He has many research papers about the theory of inequalities, fractional theory, ıntegral equations and transforms.*

Mahir Kadakal *is a Professor at the Department of Mathematics, Art and Science Faculty, Giresun University. He received his BSc (1993) degree from Department of Mathematics Teaching, Faculty of Education, Ondokuz Mayıs University and MSc (1996) degree and PhD (2000) from Ondokuz Mayıs University, Turkey. He has many research papers about the theory of inequalities, ıntegral equations and transforms, boundary value problems.*

Alper Aydın *is a mathematics teacher in the Ministry of National Education. He received his BSc (1998) degree from Department of Mathematics Teaching, Atatürk University, Turkey and MSc (2017) degree from Giresun University, Turkey. His research areas includes theory of inequalities and integral equations and transforms.*

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.266-275 (2018)
http://doi.org/10.11121/ijocta.01.2018.00457

RESEARCH ARTICLE

Optimization of lactic acid bacteria viability using fuzzy soft set modelling

Reyhan Irkin[a], Nihal Yılmaz Özgür[b*] and Nihal Taş[b]

[a] *Department of Nutrition and Dietetics, İzmir Democracy University, Turkey*
[b] *Department of Mathematics Balıkesir University, Turkey*
reyhan.irkin@idu.edu.tr, nihal@balikesir.edu.tr, nihaltas@balikesir.edu.tr

ARTICLE INFO

Article History:
Received 15 February 2017
Accepted 14 June 2018
Available 31 July 2018

Keywords:
Pickle
Salt concentration
Lactic acid bacteria
Optimization
Fuzzy soft set

AMS Classification 2010:
03E72, 06D72, 54A40, 97E60

ABSTRACT

Lactic acid fermented vegetables are important sources of vitamins and minerals. In recent years consumers demand for non-dairy based functional products has increased. Cabbage pickle has high enough concentrations of fiber and also it may show health effect with the containing high numbers of lactic acid bacteria. The aim of this study is to optimize mathematically cabbage-carrot pickle fermentation for the viability of *Lactobacillus acidophilus*, *Lactobacillus casei* cultures and the sensory scores in brine with 5% and 7% (w/v) salt concentrations. Viability optimization of lactic acid bacteria is done via the notion of "*fuzzy soft set*" method. *Lb. casei*, *Lb. acidophilus*, total lactic acid bacteria, *Enterobacteriaceae* sp., yeast-mould counts and pH values have been reported during the 30 days of storage. The results are compared with the control traditional fermented cabbage-carrot pickle. Organoleptic properties are evaluated. We conclude that the fermented pickle samples contain a significant number of beneficial lactic acid bacteria and high sensory marks at the end of the storage.

1. Introduction

Nowadays consumers pay a lot of attention to the relation between food and health. The market for foods with health promoting properties called functional foods has shown remarkable growth over the last years. In this context, functional foods have received considerable attention in recent years [1].

Functional lactic acid bacteria can reduce the number of undesired microorganisms in vegetable products [2]. It has some functions on antibiotic-associated diarrhea and immune system functions [3]. *Lb. acidophilus* and *Lb. casei* produce lactic acid as the main end product of fermentation. In addition, lactic acid bacteria produce hydrogen peroxide, diacetyl and bacteriocin as antimicrobial substances. *Lb. acidophilus* strains can exert anti-listerial bactericidal activity which could be of great technological importance (see [3] and [4] for more details). Also, *Lb. casei* decreased the severity of infection with *Salmonella enterica* serovar Typhimurium, a daily supplementation of *Lb. acidophilus* during post antibiotic therapy reduced the extent of disruption to the intestinal microbiota (see [3] and [5] for more details). These strains can prevent allergic disease, reducing lactose intolerance, enhancing bioavailability of nutrients. Researchers reported that *Lb. acidophilus* (acidolin , acidophilin producer) into the diet lowers the incidence of chemically induces colon tumors in rats (see [3] and [6] for more details).

Vegetables are good sources of natural antioxidants such as carotenoids, vitamins, minerals and dietary fibers. Vegetables may be preserved by fermentation, direct acidification, or a combination of other processing conditions and additives

*Corresponding Author

to yield products that are referred to as pickles. Cucumbers, cabbages, olives, peppers account for the largest volume and lesser quantities of onions, tomatoes, cauliflowers, carrots, melon rinds, okra, artichokes, beans etc. are also pickled. This method of food preservation has been used for many centuries [7]. Fermented vegetables are good sources of lactic acid bacteria (LAB). Representatives of some important genera such as *Leuconostoc*, *Lactobacillus*, *Lactococcus* and *Pediococcus* are found in fermented vegetables. It was reported that *Leuconostoc* sp. was the main species in the early stages of fermentation, while *Lactobacillus* sp. became predominant with the pH value gradually falling to 4.0 [8]. Plant fibers provide a promising alternative to suppress systemic inflammation, to reduce the risk of developing other chronic diseases and to considerably improve quality of life. Treatment with specific lactic acid bacteria and plant fibers has shown a unique ability to suppress inflammation in animal models and to prevent destruction of tissues [9].

The usage of starter culture (Lactic acid bacteria - LAB) performs food safety and causes rapid decrease of the pH. The controlled fermentation of pickles reduces economic losses and leads to get uniform quality product over a short period time. In the last years, *Lb. casei* and *Lb. acidophilus* were began to use in vegetable and fruit juice based products such as tomato, cabbage, beet, orange, pineapple, carrot, grape juices (see [10], [11] and [12] for more details). The high salt levels are used to select for naturally occurring, heterofermentative and homofermentative lactic acid bacteria (LAB) to carry out the fermentation and then to protect against spoilage after the active fermentation period. However, health authorities recommend a reduction of the salt content in food and nowadays consumers of industrialised countries demand low-salt foods for health reasons (see [6], [13], [14] and [15] for more details).

In this study our aim is to establish the growth prediction models under two different salinity conditions (5% and 7% w/v) for two critical microorganisms of *Lb. casei* and *Lb. acidophilus* in functional pickled cabbage-carrot processing involved. We analyze the obtained results using the notion of a fuzzy soft set. There exist some applications of the notions of "*soft set*" and "*fuzzy soft set*" in many areas of science (see [16], [17], [18], [19] and [20] for more details). This study is the first application of the fuzzy soft set theory in food engineering. By this approach, it can be easily optimize the results by using appropriate parameters and degree of membership functions. We see that this decision making method gives accurate and adequate predicts.

2. Materials and methods

In this section we explain our approach and methods.

2.1. Pickle production

Cabbage and carrots were bought from a local market in Balıkesir, Turkey. Cabbage outer leaves were separated and after cut the halves, all pieces shredded small pieces (about $2 \times 2 \times 0.02$ cm^3 dimensions), they washed with tap water before treatment rinsed with a disinfectant (Surfcera-Vegisafe, Japan) solution. Carrots were trimmed, washed and cut into slices. After that cabbage and carrot pieces were mixed and divided into the groups and filled into the glass-jars.

5% and 7% of (w/v) NaCl were added in to the tap water. Brine solutions were sterilized at 100°C/20 min and cooled to 20° C. Garlic 10 g/lt, Saccharose 5 g/lt, Grape vinegar 10 ml/lt were added into the per jar and brine solution were filled completely in pickle jars.

2.2. Bacterial cultures

Lactobacillus casei (NRRL B-1922) and *Lactobacillus acidophilus* (NRRL B-4495) freeze-dried strains were obtained from United States Department of Agriculture, Illinois, US. They activated in sterilized Liver Infusion Broth (Difco, US) and then activated young culture (10^{10} cfu/ml) of them added (5 ml/lt) into the pickle jars and mixed homogeneously. Lactic acid bacteria were not added into the control groups. *Lb. casei* and *Lb. acidophilus* as starter cultures were inoculated about 10^8/ml into the samples and the number of total lactic acid bacteria was 10^6/ml on raw vegetables before the fermentation process. *Lb. casei* is one of the normal microflora bacterium in pickles and it was determined as < 2 log cfu/g in control pickle groups. *Lb. acidophilus* was not found in control pickles.

At first, pickles were incubated in 35°C/8 h growing conditions for *Lb. acidophilus* and *Lb.casei* (see [21] and [22] for more details) for developing acidity. Then they were incubated at 20 °C/15 d [23] after that they stored in refrigerator at 4°C. Acidity, viable numbers of *Lb. casei*, *Lb. acidophilus*, total lactic acid bacteria and yeast and mould counts were determined in 1., 4., 10., 17. and 30. days.

2.3. Microbiological analysis

Samples were diluted in (1:10) in Buffered Peptone Water (BPW) and homogenized for 20 s in a stomacher (Bag mixer, Interscience, FR). Subsequently, a decimal dilution series made in BPW and enumeration was performed by pour plate or spread plating techniques (see Table 1).

Table 1. Media and incubation conditions used for determining microorganism numbers.

Microorganism	**Media**	**Incubation conditions**
Total lactic acid bacteria numbers	Man Rogosa Sharpe agar	30°C/2-3 d
Lb. casei number	MRS-Vancomycine agar 2 ml/l vancomycine (0.05 g vancomycine/100 ml) in MRS.	30° C/2-3 d
Lb. acidophilus number	MRS-Sorbitol agar (10 ml, 10% (w/v), D-Sorbitol/90 ml agar)	35 °C/2-3 d
Enterobacteriaceae sp.	Violet Red Bile Dextrose Agar	37°C/ 24 h
Yeast-Mold numbers	Rose Bengal Agar	30 °C/3-5 d

2.4. Sensory analysis

After the 30 days of storage, the products were sensorially evaluated for taste, odor, color, texture, and overall acceptability by a panel consisting of 7 sensory experts. The panelists, who have considerable background knowledge in sensory evaluation, were selected from the staff, researchers. The score given by the panel varied from 1 (dislike extremely) to 5 (like extremely). The sensory characteristics of odor, texture, and taste were assessed on each test sample. Sensory analysis of the pickle sample was evaluated explaining the following descriptors – salty, acidic sourness, sweet acidic, smelly, bitter, kraut sulfur flavor, raw cabbage flavor, discoloration, pleasant aroma and overall acceptability. A sample was considered as unacceptable for a sensory characteristic if the score was less than 2.5 (see [24] and [25] for more details).

2.5. Decision making via "Fuzzy Soft Sets"

We optimize the lactic acid bacteria viability and sensory acceptability in the production of white cabbage-carrot pickle fermentation in different salinity conditions by means of "*fuzzy soft sets* " [26] modelling. At first, the membership functions which through all of the results using MATLAB (Matlab R 2015 a and Curve Fitting Toolbox -Version 8.5, The Mathworks, Inc., Natick, Massachusetts, United States) for pH, *Lb. casei*, *Lb. acidophilus*, lactic acid bacteria and yeast–mold numbers were constructed. It was used the notion of "*fuzzy soft set*" to optimize the results. To do this it was defined a universal set and a parameter set. Fuzzy soft sets were obtained by use of the membership functions, the universal set and the parameter set. Finally, Table 3 was constructed using the values of membership functions and matrix theory [18]. Using this table it can be decided the most appropriate salt concentration in pickles and compared with the sensory scores.

3. Results

In this section we present the obtained results.

3.1. Acidity results

The usage of the starter culture performs food safety and causes rapid decrease of the pH. Decline of the pH resulted in all pickle samples with 5% and 7% salt (w/v) brine conditions. Rapid increases of active culture in the pickles caused optimum fermentation conditions and decreasing of the pH.

3.2. Microbiological analysis results

In our study, *Lb. casei* in 7% (w/v) salt concentration was affected from low pH (3.32) and high salt concentration conditions and the numbers were low (7.85 log cfu/gr) than the others at the end of the storage (see Figures 1-4). At the first day of fermentation *Lb. casei* numbers were 7.74 and 9.02 log cfu/g in pickles with 5% (w/v) and 7% (w/v) brine solution, respectively. However, at the end of the storage *Lb. casei* numbers were reached to 8.7 log cfu/g in pickles with 5% (w/v) brine solution and 7.85 log cfu/g with 7% (w/v) brine solution.

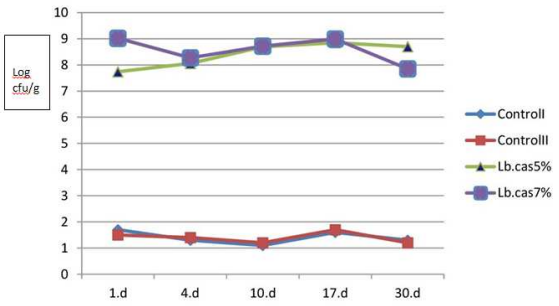


Figure 1. *Lb. casei* numbers (log CFU/gr).

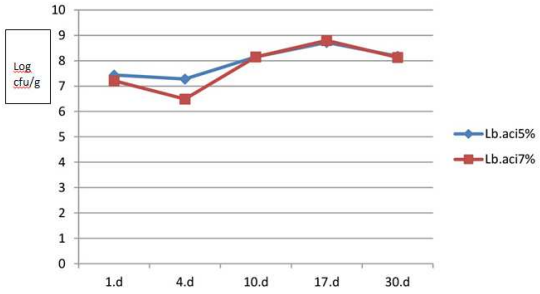


Figure 2. *Lb. acidophilus* numbers (log CFU/gr).

Total lactic acid bacteria numbers were determined as max. 9.5 and 8.88 log cfu/g for *Lb. casei* and *Lb. acidophilus*, respectively on the 17^{th} day of storage of 7% (w/v) brine solutions. However, it was found that 8.66 and 8,72 log cfu/g for *Lb. casei* and *Lb. acidophilus* numbers in 5% (w/v) brine solutions in the same day. After the 30 days of storage *Lb. casei* numbers decreased to 7.77 and 7.6 log cfu/g in 7% and 5% (w/v) salt concentration brine, respectively (see Figure 3).

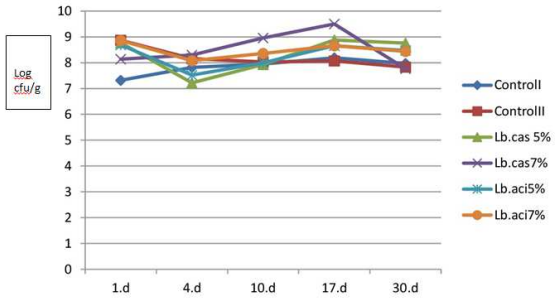


Figure 3. Total lactic acid bacteria (LAB) numbers (log CFU/gr).

Moulds didn't grow in any pickles. Only some yeasts were observed. Yeast numbers showed a sharply decline during the 30 days (see Figure 4). Maximum yeast numbers were 5.16 log cfu/g in *Lb. casei* containing pickles with 5% (w/v) salt concentrations brine on the first day. Then, minimum yeast numbers were determined as 0.28 log cfu/g in *Lb. acidophilus* containing pickles with 7% (w/v) salt concentrations brine on the 30^{th} day.

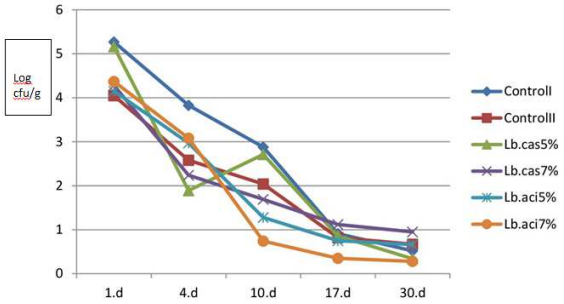


Figure 4. Total yeast numbers (log CFU/gr) during the storage of pickles.

In research, *Enterobacteriacea* spp. were not determined in pickles and it can be thought that because of washing vegetables with a disinfectant solution before treatments cause inhibition effects on these groups of bacteria.

3.3. Sensory analysis results

From the Figure 5, we can see that the sensory analysis results. Using these results we deduce that the pickles with *Lb. casei* culture in 5% salt concentration were preferred according to the overall, colour, taste and odour quality properties. But the texture of the pickles with 7% salt concentrations have higher scores than the others.

Figure 5. Sensory profile of pickle samples during the storage of 30 days.

3.4. Microbial analysis via fuzzy soft sets

Now we analyze the obtained results using the notion of "*fuzzy soft set*". The notion of a fuzzy soft set was introduced in [26] as follows:

Let U be a universal set, E be a set of parameters and $A \subset E$. Let $P(U)$ denotes the set of all fuzzy subsets of U. Then a pair (F, A) is called a fuzzy soft set over U, where F is mapping from A to $P(U)$.

Following algorithm can be presented for optimization of the salt concentration :

Step 1. Defining a universal set X and a parameter set E.

Step 2. Using MATLAB (Curve Fitting Toolbox- 2015), choose appropriate functions through all of the results obtained for *ph*, *lc*, *la*, *lab* and *ym*, respectively.

Step 3. Using the functions chosen in Step 2, define the membership functions for *ph*, *lc*, *la*, *lab* and *ym*, respectively.

Step 4. Constructing the fuzzy soft sets $(F_{x,i}, E_i)$, $i = 1$ to 6 where $x \in \{1, 4, 10, 17, 30\}$.

Step 5. Computing the fuzzy soft sets $(F_{30,i}, E_i)$, $i = 1$ to 6.

Step 6. Constructing a table of the fuzzy soft sets $(F_{30,i}, E_i)$ and insert "0" or "1" for each E_i, $i = 1$ to 6.

Step 7. Adding sensory scores column obtained by a panel consisting of 7 sensory experts to the above table.

Step 8. Making decision is the E_i, $(i = 1$ to $6)$ which has the top score "1".

Then, data can be established by fuzzy soft modeling to optimize the results as follows:

It can be defined the following notations:

ph : *pH value,*
lc : *Lb. casei number,*
la : *Lb. acidophilus number,*
lab : *Lactic acid bacteria numbers,*
ym : *Yeast - mold number,*

and

E_1 : Control %5,
E_2 : Control %7,
E_3 : *Lb. casei* %5,
E_4 : *Lb. casei* %7,
E_5 : *Lb. acido* %5,
E_6 : *Lb. acido* %7.

A universal set and a parameter set were defined as follows, respectively:

$$X = \{ph, lc, la, lab, ym\} ,$$

and

$$E = \{E_1, E_2, E_3, E_4, E_5, E_6\}.$$

Table 2. pH ± standard deviations (S.D.) of pickles during the storage at (1., 4., 7., 10., 17. and 30. days) +4 °C.

Products	**1**	**4**	**10**	**17**	**30**
Control I 5%	3.44±0.4	3.47±0.5	3.40±0.4	3.35±0.2	3.34±0.8
Control II 7%	3.58±0.6	3.61±0.9	3.55±0.5	3.41±0.5	3.39±0.4
Lb.casei 5%	3.83±0.3	3.70±0.7	3.52±0.7	3.45±0.5	3.39±0.5
Lb. casei 7%	3.73±0.5	3.59±0.6	3.44±0.9	3.33±0.6	3.32±0.7
Lb. acidophilus 5%	3.62±0.7	3.55±0.3	3.48±0.6	3.40±0.4	3.36±0.2
Lb.acidophilus 7%	3.63±0.4	3.55±0.8	3.46±0.3	3.42±0.9	3.39±0.5

Appropriate functions such as ph_i, lc_i, la_i, lab_i, ym_i were chosen using MATLAB (Curve Fitting Toolbox- 2015) where $i \in \{1, 2, 3, 4, 5, 6\}$. Then it was given membership functions for pH, *Lb. casei* numbers, *Lb. acidophilus* numbers, lactic acid bacteria numbers and yeast – mold numbers as the following cases:

Case1: Using MATLAB – Curve Fitting Toolbox and the results given in Table 2, following functions for pH value (see Figure 6) were chosen:

$$ph_1(x) = 0.0004x^3 - 0.0068x^2 + 0.0371x + 3.4094,$$
$$ph_2(x) = 0.0001x^3 - 0.0030x^2 + 0.0223x + 3.5619,$$
$$ph_3(x) = 0.0026x^2 - 0.0580x + 3.8882,$$
$$ph_4(x) = 0.0017x^2 - 0.0491x + 3.7719,$$
$$ph_5(x) = 0.0004x^2 - 0.0205x + 3.6352,$$
$$ph_6(x) = 0.0015x^2 - 0.0329x + 3.6606,$$

where $x \in \{1, 4, 10, 17, 30\}$.

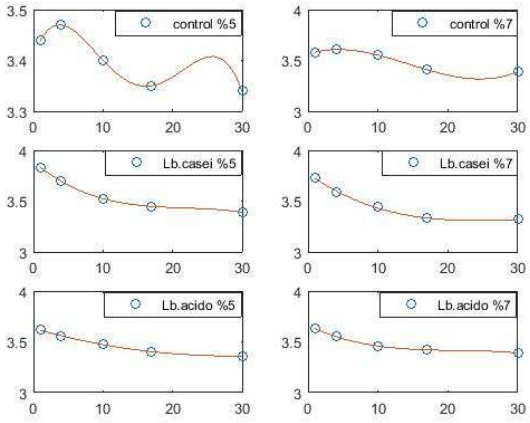


Figure 6. The graphics of the membership functions for pH value (the x-axis represent the days and the y-axis represent the pH values).

The following membership function was defined

$$\mu_i^{ph}(x) = \frac{ph_i(x)}{10^n},$$

for all $i \in \{1,2,3,4,5,6\}$, where n is the number of digits of the integer part of $ph_i(x)$.

Case 2: By a similar way, using the results given in Figure 1, following functions for *Lb. casei* number (see Figure 7) were chosen:

$$lc_3(x) = -0.0010x^3 + 0.0123x^2 + 0.0656x + 7.6632,$$
$$lc_4(x) = 0.0001x^4 - 0.0049x^3 + 0.0963x^2 - 0.6350x + 9.5735,$$

where $x \in \{1,4,10,17,30\}$.

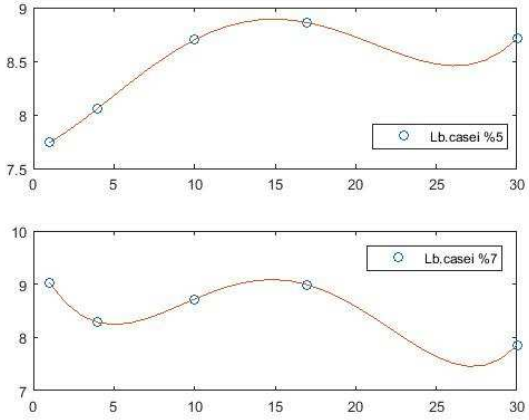


Figure 7. The graphics of the membership functions for *Lb. casei* number (the x-axis represent the days and the y-axis represent the *Lb. casei* numbers).

The following membership function was defined

$$\mu_i^{lc}(x) = \begin{cases} \frac{lc_i(x)}{10^n} \text{ if } i \in \{3,4\} \\ 0 \text{ if otherwise} \end{cases},$$

for all $i \in \{1,2,3,4,5,6\}$, where n is the number of digits of the integer part of $lc_i(x)$.

Case 3: Using the results given in Figure 2, following functions for *Lb. acidophilus* number were chosen (see Figure 8):

$$la_5(x) = 0.0001x^4 - 0.0035x^3 + 0.0657x^2 - 0.3156x + 7.7034,$$
$$la_6(x) = 0.0002x^4 - 0.0098x^3 + 0.1760x^2 - 0.9286x + 7.9723,$$

where $x \in \{1,4,10,17,30\}$.

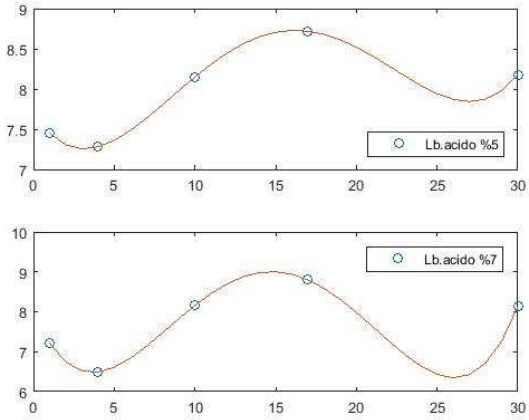


Figure 8. The graphics of the membership functions for *Lb. acidophilus* number (the x-axis represent the days and the y-axis represent the *Lb. acidophilus* numbers).

The following membership function was defined

$$\mu_i^{la}(x) = \begin{cases} \frac{la_i(x)}{10^n} \text{ if } i \in \{5,6\} \\ 0 \text{ if otherwise} \end{cases},$$

for all $i \in \{1,2,3,4,5,6\}$, where n is the number of digits of the integer part of $la_i(x)$.

Case 4: Using the results given in Figure 3, it can be chosen the following functions for lactic acid bacteria numbers (see Figure 9):

$$lab_1(x) = 0.0023x^3 - 0.0430x^2 + 0.3338x + 7.0369,$$
$$lab_2(x) = -0.0028x^3 + 0.0592x^2 - 0.4768x + 9.2904,$$
$$lab_3(x) = 0.0001x^4 - 0.0087x^3 + 0.1778x^2 - 1.2342x + 9.8350,$$
$$lab_4(x) = -0.0008x^3 + 0.0172x^2 - 0.0160x + 8.1396,$$
$$lab_5(x) = 0.0001x^4 - 0.0064x^3 + 0.1315x^2 - 0.9209x + 9.4857,$$
$$lab_6(x) = 0.0001x^4 - 0.0045x^3 + 0.0891x^2 - 0.6211x + 9.4065,$$

where $x \in \{1,4,10,17,30\}$.

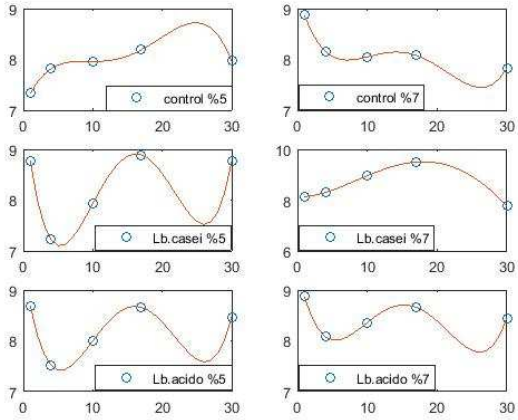


Figure 9. The graphics of the membership functions for lactic acid bacteria numbers (the x-axis represent the days and the y-axis represent the lactic acid bacteria numbers).

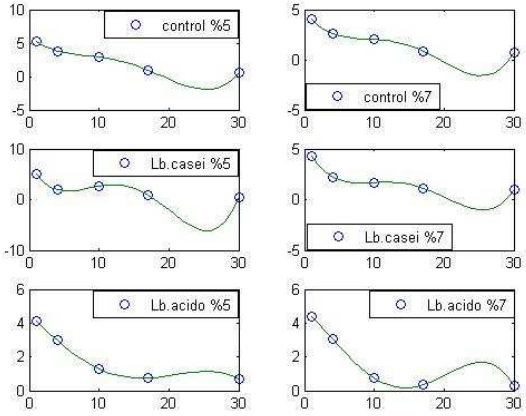


Figure 10. The graphics of the membership functions for yeast and mold numbers (the x-axis represent the days and the y-axis represent the yeast and mold numbers).

The following membership function was defined

$$\mu_i^{lab}(x) = \frac{lab_i(x)}{10^n},$$

for all $i \in \{1,2,3,4,5,6\}$, where n is the number of digits of the integer part of $lab_i(x)$.

Case 5: Using the results given in Figure 4, the following functions for yeast – mold numbers were chosen (see Figure 10):

$$ym_1(x) = 0.0001x^4 - 0.0069x^3 + 0.1173x^2 - 0.9326x + 6.0921,$$
$$ym_2(x) = 0.0001x^4 - 0.0073x^3 + 0.1321x^2 - 1.0078x + 4.9329,$$
$$ym_3(x) = 0.0004x^4 - 0.0238x^3 + 0.4215x^2 - 2.7341x + 7.4959,$$
$$ym_4(x) = 0.0001x^4 - 0.0086x^3 + 0.1700x^2 - 1.3615x + 5.4800,$$
$$ym_5(x) = 0.0010x^3 + 0.0018x^2 - 0.4209x + 4.5682,$$
$$ym_6(x) = -0.0001x^4 + 0.0038x^3 - 0.0392x^2 - 0.3070x + 4.7125,$$

where $x \in \{1,4,10,17,30\}$.

The following membership function was defined

$$\mu_i^{ym}(x) = \frac{ym_i(x)}{10^n},$$

for all $i \in \{1,2,3,4,5,6\}$, where n is the number of digits of the integer part of $ym_i(x)$.
The fuzzy soft sets $(F_{x,i}, E_i)$, $i = 1$ to 6 were constructed by considering the membership values $\mu_i^{ph}(x)$, $\mu_i^{lb}(x)$, $\mu_i^{la}(x)$, $\mu_i^{lab}(x)$ and $\mu_i^{ym}(x)$. It can be defined

$$(F_{x,i}, E_i) = F_i = \left\{ \frac{ph}{\mu_i^{ph}(x)}, \frac{lc}{\mu_i^{lc}(x)}, \frac{la}{\mu_i^{la}(x)}, \frac{lab}{\mu_i^{lab}(x)}, \frac{ym}{\mu_i^{ym}(x)} \right\},$$

where $i \in \{1,2,3,4,5,6\}$ and $x \in \{1,4,10,17,30\}$.
Now it can be investigated fuzzy soft sets for $x = 30$ and $i = 1$ to 6.

$$(F_{30,1}, E_1) = F_1(\text{control } \%5) = \left\{ \frac{ph}{0.334}, \frac{lc}{0}, \frac{la}{0}, \frac{lab}{0.797}, \frac{ym}{0.52} \right\}.$$

$$(F_{30,2}, E_2) = F_2(\text{control } \%7) = \left\{ \frac{ph}{0.339}, \frac{lc}{0}, \frac{la}{0}, \frac{lab}{0.783}, \frac{ym}{0.67} \right\}.$$

$$(F_{30,3}, E_3) = F_3(\text{Lb.casei } \%5) = \left\{ \frac{ph}{0.339}, \frac{lc}{0.871}, \frac{la}{0}, \frac{lab}{0.876}, \frac{ym}{0.34} \right\}.$$

$$(F_{30,4}, E_4) = F_4(\text{Lb.casei \%7})$$
$$= \left\{ \frac{ph}{0.332}, \frac{lc}{0.785}, \frac{la}{0}, \frac{lab}{0.777}, \frac{ym}{0.95} \right\}.$$

$$(F_{30,5}, E_5) = F_5(\text{Lb.acido \%5})$$
$$= \left\{ \frac{ph}{0.336}, \frac{lc}{0}, \frac{la}{0.818}, \frac{lab}{0.847}, \frac{ym}{0.66} \right\}.$$

$$(F_{30,6}, E_6) = F_6(\text{Lb.acido \%7})$$
$$= \left\{ \frac{ph}{0.339}, \frac{lc}{0}, \frac{la}{0.813}, \frac{lab}{0.843}, \frac{ym}{0.28} \right\}.$$

By a similar way, other fuzzy soft sets can be defined for $x \in \{1, 4, 10, 17\}$ and $i = 1$ to 6.

A results' table can be constructed to see the most accurate and adequate predictions in the research. It can be written "1" for the largest value of the membership function for each parameter E_i, $i = 1$ to 6 and "0" for other values. Then the row which has the top "1" score is chosen for the prediction. It can be seen from the Table 3 parameter E_3 is the optimum salt concentration for our research.

As it is seen from the Table 3, this model could accurately and adequately predict the growths of *Lb. casei*, *Lb. acidophilus*, total lactic acid bacteria and yeasts in pickles. And it can be shown that pickles with 5% (w/v) salt concentrations and *Lb. casei* culture gave the highest scores in this study.

Table 3. The results for $x = 30$ and $i = 1$ to 6.

	ph	*lc*	*la*	*lab*	*ym*	*sensory scores*
E_1	0	0	0	0	0	0
E_2	1	0	0	0	0	0
E_3	1	1	0	1	0	1
E_4	0	0	0	0	0	0
E_5	0	0	1	0	0	0
E_6	1	0	0	0	1	0

4. Discussions

Heterofermentative LAB are more sensitive to high salt concentrations than homofermentatives. Therefore high salt levels favor the growth of homofermentative LAB and resulting in an accelerated production of lactic acid. Maintaining the viability (minimum numbers of probiotic cultures present in the final product recommended to be 10^6 cfu/ml or higher) and the activity of lactic acid bacteria in foods to the end of shelf life are two important criteria [27].

Salt concentration can affect the growth of the naturally present microorganisms and the sensory properties of the pickles [28]. In the following studies they have shown potential benefits of using starter cultures in low-salt pickle fermentations. In Xiong et al. [29] study unfavorable conditions resulting from low pH contributed to the rapid decline of the lactococci as fermentation progressed. Similarly, Weng et al. [30] found that NaCl concentration affects the growth of *L. citreum* L-33 in pickle and the specific growth rate decreases with the increase of NaCl concentration. It also demonstrated that less salt addition can lead to a rapid growth for the important lactic acid bacteria in pickling production. Beganovic et al. [28] showed that the application of the probiotic strain *Lb. plantarum* L4 together with *Lc. mesenteroides* LMG 7954, positively influenced the fermentations by improving the quality of the final product with added probiotic properties, considerably shortening the fermentation time and offering the possibility of low salt fermentations (2.5 w/v). The rapid increase in acidity minimizes the influence of spoilage bacteria. Reducing the influence of spoilage bacteria would most probably improve the microbiological and sensory quality of the fermented end product significantly (see [31] and [32] for more details).

Yoon et al. [33] researched red beet juice fermentation with *Lb. acidophilus*, *Lb. casei*, *Lb. delbrueckii* and *Lb. plantarum* in their study. They found that *Lb. acidophilus* in fermented beet juice could be remained at $10^6 - 10^8$ cfu/ml after 4 weeks of cold storage and the others lost their viability. In this study, all viable numbers after the storage were $> 10^6$ cfu/g and it can be considered pickles have probiotic meaning health of view (Figures 1 and 2). *Lb. acidophilus* numbers reached max. 8.8 log cfu/g on the 17^{th} day of storage in pickles with 5% (w/v) and 7% (w/v) salt concentration brine. On the 30^{th} day, bacteria numbers were very close to each other and declined to 8.13 log cfu/g.

LAB produce several antimicrobials, including organic acids (lactic, acetic, formic, phenyllactic, caproic acids) carbondioxide, hydrogen peroxide, diacetyl, ethanol, bacteriocins, reuterin and reutericyclin and they can prevent mould spoilage and growth of some pathogenic bacteria (see [8], [34] and [35] for more details). It can be concluded that these specific lactic acid bacteria in pickles were able to produce bacteriocins that inhibit yeasts at the last of storage.

Using bacterial cultures must not only aim at expressing functional properties upon microbial growth, but also at the impact on other quality

changes in the pickles. Among these changes, the sensory or organoleptic property is the most important [13]. In Weon and Lee [36] study, perception and preference of low salt Korean pickle "Jangachi" were evaluated. Low sodium Jangachi was found safe, sanity, safekeeping and the most preferred by the consumers.

5. Conclusions

Researches' results provided a useful basis for further studies of the development of optimization sodium chloride concentration level in brine during fermentation of pickles. It was demonstrated that concentration of sodium chloride in brine solution have significant effects on the growths of *Lb. acidophilus*, *Lb. casei*, total lactic acid bacteria and yeast counts during fermentation. The population dynamics during cabbage-carrot pickle fermentation can be predicted by the use of the notion of fuzzy soft set at the same production conditions. Application of "*fuzzy soft set*" in optimization of pickle fermentation helps and allows better understanding of the interaction between the variables.

Acknowledgement 1. *The authors would like to thank the anonymous referees for their comments that helped us improve this article.*

References

[1] Tokatlı, M., Gülgör, G., Elmacı, S.B., İşleyen, N.A., Özçelik, F. (2015). In vitro properties of potential probiotic indigenous lactic acid bacteria originating from traditional pickles. *BioMed. Res. Int.*, 1, 1-8.

[2] Tamang, J.P., Tamang, B., Schillinger, U., Guigas, C., Holzapfel, W.H. (2009). Functional properties of lactic acid bacteria isolated from ethnic fermented vegetables of the Himalayas. *Int. J. Food Microbiol.*, 135, 28-33.

[3] Divya, J.B., Varsha, K.K., Madhavan, K., Ismail, N.B., Pandey, A. (2012). Probiotic fermented foods for health benefits. *Eng. Life Sci.*, 12, 377-390.

[4] Kos, B., Suskovic, J., Beganovic, J., Gjuracic, K., Frece, J., Iannaccone, C., Canganella, F. (2008). Characterization of the three selected probiotic strains for the application in food industry. *World J. Microbiol. Biotechnol.*, 24, 699-707.

[5] Surh, J., Kim, Y.K.L., Kwon, H. (2008). Korean fermented foods: Kimchi and Doenjang. *In Handbook of Fermented Functional Foods*. Edward, R.; Farnworth, T., Eds.; CRC Press, US 333-353.

[6] Nagpal, R., Kumar, A., Kumar, M., Behare, P.V., Jain, S., Yadav, H. (2012). Probiotics, their health benefits and applications for developing healthier foods: a review. *FEMS Microbiol. Lett.*, 334, 1-15.

[7] Wacher, C., Diaz-Ruiz, G., Tamang, J.P. (2010). *Fermented Vegetable Products. In Fermented Foods and Beverages of the World*; Tamang J.P. ; Kaialasapathy, K., Eds.; CRC Press, US 149-190.

[8] Zhou, F., Zhao, H., Bai, F., Dziugan, P., Liu, Y., Zhang, B. (2014). Purification and characterisation of the bacteriocin produced by *Lactobacillus plantarum*, isolated from Chinese pickle. *Czech J. Food Sci.*, 32, 430-436.

[9] Bengmark, S. (2010). *Lactic acid Bacteria and Plant Fibers: Treatment in acute and chronic human disease. In Prebiotics and Probiotics ingredients, Health Benefits and Food Applications*; Cho, S.S.; Finocchiaro, E.T., Eds.; CRC Press, US 163-192.

[10] Montet, D., Ray, R.C., Zakhia-Rozis, N. (2015). *Lactic acid fermentation of vegetables and fruits. In Microorganisms and Fermentation of Traditional Foods*; Ray R.C.; Montet D., Eds.; CRC Press, US 108-141.

[11] Furtado-Martins, E.M., Ramos, A.M., Vanzela, E.S.L., Stringheta, P.C., Pinto, C.L.O., Martins, J.M. (2013). Products of vegetable origin: A new alternative for the consumption of probiotic bacteria. *Food Res. Int.*, 51, 764770.

[12] Yoon, K.Y., Woodams, E.E., Hang, Y.D. (2009). Production of probiotic cabbage juice by lactic acid bacteria. *Biores. Technol.*, 97, 14271430.

[13] Zhao, D., Ding, X. (2008). Studies on the low-salt Chinese potherb mustard (Brassica juncea, Coss.) pickle. I—The effect of a homofermentative L(+) -lactic acid producer Bacillus coagulans on starter culture in the low-salt Chinese potherb mustard pickle fermentation. *LWT-Food Sci. Technol.*, 41(3), 474-482.

[14] Penas, E., Frias, J., Sidro,B., Vidal-Valverde, C. (2010). Impact of fermentation conditions and refrigerated storage on microbial quality and biogenic amine content of sauerkraut. *Food Chem.*. 123, 143-150.

[15] Lin, S.H., Li, Y. H., Leung, K., Huang, C.Y., Wang, X.R. (2014). Salt processed food and gastric cancer in a Chinese population. *As. Pac. J. Cancer Prev.*, 15(13), 5293-5298.

[16] Kalaichelvi, A., Malini, P. H. (2011). Application of fuzzy soft sets to investment decision making problem. *Int. J. Math. Sci. Appl.*, 1(3), 1583-1586.

[17] Yüksel, S., Dizman, T., Yıldızdan, G., Sert, U. (2013). Application of soft sets to diagnose the prostate cancer risk. *J. Inequal Appl.*, 1, 229-240.

[18] Özgür, N.Y., Tas, N. (2015). A note on "Application of Fuzzy Soft Sets to Investment Decision Making Problem". *J. New Theory.*, 1(7), 1-10.

[19] Maji, P.K., Biswas, R., Roy, A.R. (2002). An Application of Soft Set in a Decision Making Problem. *Comput. Math. Appl.* 44(8-9), 1077-1083.

[20] Roy, A.R., Maji, P.K. (2007). A Fuzzy Soft Set Theoretic Approach to Decision Making Problems. *J. Comput. Appl. Math.*, 203(2), 412-418.

[21] Ng, E.W.Y. (2009). Effect of starter cultures on Lactobacillus acidophilus and gene expression in yogurt. Thesis of California Polytechnic State University, San Luis Obispo, US 108 p.

[22] Oh, S., Rheem, S., Sim, J., Kim, S., Baek, Y.J. (1995). Optimizing Conditions for the Growth of Lactobacillus casei YIT 9018 in Tryptone-Yeast Extract-Glucose Medium by Using Response Surface Methodology. *Appl. Environ. Microbiol.*, 61(11), 3809-3814.

[23] Zhao, D., Tang, J., Ding, X. (2007). Analysis of volatile components during potherb mustard (Brassica juncea, Coss.) pickle fermentation using SPME–GC-MS. *LWT-Food Sci. Technol.*, 40, 439-447.

[24] Inatsu, Y., Bari, M.L., Kawasaki, S., Kawamoto, S. (2005). Effectiveness of some natural antimicrobial

compounds in controlling pathogen or spoilage bacteria in lightly fermented Chinese cabbage. *J. Food Sci.*, 70(9), 393-397.

[25] Jagannath, P., Raju, P.S., Bawa, A. S. (2012). A Two-step controlled lactic fermentation of cabbage for improved chemical and microbiological qualities. *J. Food Qual.*, 35, 13-20.

[26] Maji, P.K., Biswas, R., Roy, A.R. (2001). Fuzzy Soft Sets. *J. Fuzzy Math.*, 9, 589-602.

[27] Kearney, N., Stanton, C., Desmond, C., Coakley, M., Collins, J. K., Fitzgerald, G., Ross, R.P. (2008). *Challenges associated with the development of probiotic-containing Functional foods. In Handbook of Fermented Functional Foods;* Edward, R.; Farnworth, T., Eds.; CRC Press, US 25-71.

[28] Beganovic, J., Pavunc, A. L., Gjuracic, K., Spoljarec, M., Suskovic, J., Kos, B. (2011). Improved sauerkraut production with probiotic strain Lactobacillus plantarum L4 and Leuconostoc mesenteroides LMG 7954. *J. Food Sci.* 76(2), 124-129.

[29] Xiong, T., Guan, Q., Song, S., Hao, M., Xie, M. (2012). Dynamic changes of lactic acid bacteria flora during Chinese sauerkraut fermentation. *Food Control.* 26, 178-181.

[30] Weng, P.F., Wu, Z.F., Lei, L.L. (2013). Predictive models for growth of Leuconostoc citreum and its Dynamics in pickled vegetables with low salinity. *J. Food Proces. Eng.*, 36, 284-291.

[31] Viander, B., Maki, M.M., Palva, A. (2003). Impact of low salt concentration, salt quality on natural large-scale sauerkraut fermentation. *Food Microbiol.*, 20, 391-395.

[32] Cvetkovic, B.R., Pezo, L. L., Tasic, T., Saric, L., Kevresan, Z., Mastilovic, J. (2015). The optimisation of traditional fermentation process of white cabbage (in relation to biogenic amines and polyamines content and microbiological profile). *Food Chem.* 168, 471-477.

[33] Yoon, K.Y., Woodams, E. E., Hang, Y. D. (2005). Fermentation of beet juice by beneficial lactic acid bacteria. *LWT- Food Sci. Technol.*, 38, 73-75.

[34] Irkin, R., Songun, G.E. (2012). Applications of probiotic bacteria to the vegetable pickle products. *Sci. Rev. Chem. Com.*, 2(4), 562-567.

[35] Heperkan, D. (2013). Microbiota of table olive fermentations and criteria of selection for their use as starters. *Front. Microbiol.*, 4(143), 1-11.

[36] Weon, M.K., Lee, Y.J. (2013). Consumer's perception, preference and intake frequency of Jangachi (Korean Pickle) by age for developing low salt Jangachi. *Korean J. Cul. Res.*, 19(5), 249-263.

Reyhan Irkin *is currently an associate professor at İzmir Democracy University in Turkey. Her research interests include food microbiology.*

Nihal Yılmaz Özgür *is currently a professor at Balıkesir University in Turkey. Her research interests include complex functions with one variable, soft set theory and fixed point theory.*

Nihal Taş *is currently a research assistant at Balıkesir University in Turkey. Her research interests include topological spaces, soft set theory and fixed point theory.*

An International Journal of Optimization and Control: Theories & Applications
(http://ijocta.balikesir.edu.tr)

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.8, No.2, pp.276-287 (2018)
http://doi.org/10.11121/ijocta.01.2018.00572

RESEARCH ARTICLE

A rich vehicle routing problem arising in the replenishment of automated teller machines

Çağrı Koç[a], Mehmet Erbaş[b], Eren Özceylan[c*]

[a] *Department of Business Administration, Social Sciences University of Ankara, Ankara, Turkey*
[b] *General Directorate of Mapping, Ministry of National Defense, Ankara, Turkey*
[c] *Department of Industrial Engineering, Gaziantep University, Gaziantep, Turkey*

cagri.koc@asbu.edu.tr, mehmet.erbas@hgk.msb.gov.tr, erenozceylan@gmail.com

ARTICLE INFO

Article History:
Received 29 December 2017
Accepted 26 July 2018
Available 31 July 2018

Keywords:
Vehicle routing
GIS
Tabu search
Replenishment of automated teller machines

AMS Classification 2010:
90-08, 90B06, 90C11

ABSTRACT

This paper introduces, models, and solves a rich vehicle routing problem (VRP) motivated by the case study of replenishment of automated teller machines (ATMs) in Turkey. In this practical problem, commodities can be taken from the depot, as well as from the branches to efficiently manage the inventory shortages at ATMs. This rich VRP variant concerns with the joint multiple depots, pickup and delivery, multi-trip, and homogeneous fixed vehicle fleet. We first mathematically formulate the problem as a mixed-integer linear programming model. We then apply a Geographic Information System (GIS)-based solution method, which uses a tabu search heuristic optimization method, to a real dataset of one of the major bank. Our numerical results show that we are able to obtain solutions within reasonable solution time for this new and challenging practical problem. The paper presents computational and managerial results by analyzing the trade-offs between various constraints.

1. Introduction

In logistics operations, fulfilling consumer demands for diverse and premium products is an important challenge [1]. The classical vehicle routing problem (VRP) aims to determine an optimal routing plan for a fleet of homogeneous vehicles to serve a set of customers, such that each vehicle route starts and ends at the depot, each customer is visited once by one vehicle, and some side constraints are satisfied. Many variants and extensions of the VRP have intensively studied in the literature. For further details about the VRP and its variants, we refer the reader to Laporte [2] and Toth and Vigo [3].

Over the last years, several variants of multi-constrained VRPs have been studied, forming a class of problems known as Rich VRPs. Lahyani et al. [4] presented a comprehensive and relevant taxonomy for the literature devoted to Rich VRPs. The authors have investigated 41 articles devoted to rich VRPs in detail, and developed an elaborate definition of RVRPs.

Karaoglan et al. [5] studied aircraft routing and scheduling for cargo transportation in an airline company in Turkey. Karagul and Gungor [6] studied the mixed fleet VRP to optimize the distribution of the tourists who have traveled between the airport and the hotels in Turkey. The authors developed a Savings algorithm, a Sweep algorithm and a random permutation alignment. Furthermore, a genetic algorihm and random search algorithms algorithms are also developed. Van Anholt et al. [7] introduced, modeled, and solved a rich multiperiod inventory-routing problem with pickups and deliveries motivated by the replenishment of automated teller machines (ATMs) in

*Corresponding Author

the Netherlands. The authors first decomposed the problem into several more manageable subproblems by means of a clustering procedure, and they simplified the subproblems by fixing some variables. Valid inequalities are then generated to strength the resulting subproblems. An efficient branch-and-cut algorithm is then developed. Karagul et al. [8] used 2-Opt based evolution strategy for travelling salesman problem.

A variant of the VRP known as multiple depots VRP, in which more than one depot is considered, studied by many researchers. Crevier et al. [9] considered the multiple depots VRP with interdepot routes. Braekers et al. [10] proposed exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. Contardo and Martinelli [11] developed a new exact algorithm for the multiple depots VRP under capacity and route length constraints. For further details about the multiple depots VRP and its variants, we refer the reader to the review paper of Montoya et al. [12]. Another interesting variant is multi-trip VRP, in which vehicles can perform several trips per day, because of their limited number and capacity [13–15].

An important family of routing problems is pickup-and-delivery problems (PDPs) in which goods have to be transported from different origins to different destinations. In one-to-one variant of PDPs, each customer demand consists of transporting a load from one pickup node to one destination node. Many exact and heuristic methods are developed for this problem variant which is usually referred to as the pickup-and-delivery VRP. Xu et al. [16] studied a rich PDP with many side constraints. Sigurd et al. [17] considered the transportation of live animals. For further details about the PDPs and its variants, we refer the reader to Battarra et al. [18], Berbeglia et al. [19], Koç and Laporte [20], and Parragh et al. [21, 22].

In recent years, Geographical Information System (GIS)-based solution methods used to solve several optimization problems. Casas et al. [23] developed an automated network generation procedure for routing of unmanned aerial vehicles in a GIS environment. Bozkaya et al. [24] studied the competitive multi-facility location-routing problem and presented a hybrid heuristic algorithm. The method is applied on a case study arising at a supermarket store chain in the city of Istanbul. The authors used genetic algorithm for the location part, and tabu search of GIS-based solution method for the VRP part. Samanlioglu [25] developed a multi-objective location-routing problem and described a mathematical model. The author used a GIS software to obtain the data related to the Marmara region of Turkey. Yanik et al. [26] considered the capacitated VRP with multiple pickup, single delivery and time windows, and proposed a hybrid metaheuristic approach. The method integrates a genetic algorithm for vendor selection and allocation, and a GIS-based solution method which uses a modified savings algorithm for the routing part. Krichen et al. [27] studied the VRP with loading and distance constraints and used a GIS solution method to solve the problem.

Our study is motivated by the problem faced by one of the major bank of Turkey operating in the city of Gaziantep. We consider a multi-depot, multi-trip, pick-up and delivery with homogenous vehicle fleet. We first define this new problem and presented a mathematical formulation. We then used a GIS-based solution approach employs a tabu search algorithm which can be used to store, analyze and visualize all data as well as model solutions in geographic format. We considered a real dataset of the bank and applied our GIS-based solution approach. We finally provide several managerial and policy insights by exploring the trade-offs between various constraints.

The remainder of this paper is structured as follows. Section 2 presents the problem definition and mathematical formulation. Section 3 describes the solution approach. Section 4 presents a case study with input data. Section 5 presents the solutions we propose. Finally, Section 6 provides our conclusions.

2. Problem definition and mathematical formulation

The problem is defined on a complete directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$. The location of each ATM, branch, and the district office is represented by a node. $\mathcal{N} = \{0\} \cup \mathcal{N}_b \cup \mathcal{N}_c$ is a set of nodes in which "$\{0\}$" is the district office node, $\mathcal{N}_b$ is a set of branch nodes, and $\mathcal{N}_c$ is a set of ATM nodes. $\mathcal{A} = \{(i,j) : i \in \mathcal{N}, j \in \mathcal{N}, i \neq j\}$ is the set of arcs. Each arc $(i,j) \in \mathcal{A}$ has a nonnegative distance d_{ij}. Each arc $(i,j) \in \mathcal{A}$ has a nonnegative travel time c_{ij}. Each ATM $i \in \mathcal{N}_c$ has a demand q_i and a service time p_i. A fixed number of limited homogeneous vehicle fleet m is available. The index set of routes is denoted by $\mathcal{R} = \{1, \ldots, r, \ldots\}$. The capacity of a vehicle is denoted by Q. The maximum allowed working duration is T_{max} for each vehicle. We use the real network distances when we computing the d_{ij} values on each arc $(i,j) \in A$. Therefore, it is possible that $d_{ij} \neq d_{ji}$,

i.e., asymmetric, which are illustrated in Figure 1.

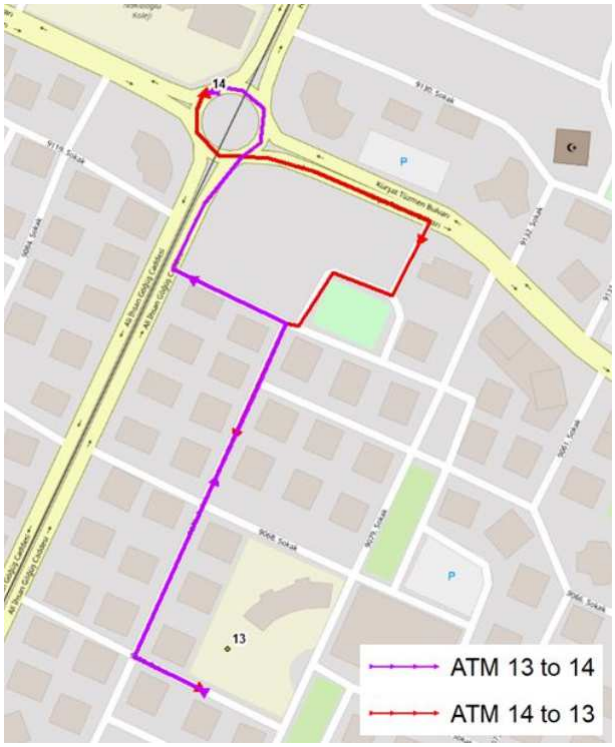


Figure 1. An example of asymmetric case from ATM 13 to 14, and from ATM 14 to 13.

To formulate the problem, we define the following decision variables. Let x_{ijr} be equal to 1 if a vehicle travels directly from node i to node j on route $r \in \mathcal{R}$. Let f_{ijr} be the amount of commodity flowing on arc $(i,j) \in \mathcal{A}$ by a vehicle on route $r \in \mathcal{R}$.

In our problem, one considers a homogeneous fixed fleet of vehicles, as well as a set of ATMs with known demands. The demand of each ATM is expressed as money tray and each vehicle are designed to satisfy these specific ATM demands. The bank has variable number of orders from ATMs which can be fulfilled by both district office and several branches. The journey of a vehicle which carries demanded money starts from district office and it begins to visit ATMs to load ordered money. If the money of a vehicle is finished before meeting the demand of ATMs, vehicle has two options. While the first option is to go back to district office, the second option is to go a branch to get money. Vehicles can perform several tours per day because of their limited number and capacity. The objective is to minimise the total en-route time of vehicles. Due to the minimization of total en-route time time in this problem, the vehicle visits district office or branch which is closer to it. This rich VRP variant is concerned with the joint multiple depots, pickup and delivery problems, multi-trip, and homogeneous fixed vehicle fleet.

The mathematical formulation of the problem is given as follows:

$$\text{Minimize} \sum_{(i,j)\in\mathcal{A}} \sum_{r\in\mathcal{R}} c_{ij} x_{ijr} \tag{1}$$

subject to

$$\sum_{j\in\mathcal{N}} x_{0j1} \leq m \tag{2}$$

$$\sum_{j\in\mathcal{N}} \sum_{r\in\mathcal{R}} x_{ijr} = 1 \qquad i \in \mathcal{N}_c \tag{3}$$

$$\sum_{i\in\mathcal{N}} \sum_{r\in\mathcal{R}} x_{ijr} = 1 \qquad j \in \mathcal{N}_c \tag{4}$$

$$\sum_{j\in\mathcal{N}_c} x_{0jr} \geq \sum_{j\in\mathcal{N}_c} x_{0j,r+1} \qquad r \in \mathcal{R} : r < |\mathcal{R}| \tag{5}$$

$$\sum_{j\in\mathcal{N}} \sum_{r\in\mathcal{R}} f_{jir} - \sum_{j\in\mathcal{N}} \sum_{r\in\mathcal{R}} f_{ijr} = q_i \qquad i \in \mathcal{N}_c \tag{6}$$

$$q_j x_{ijr} \leq f_{ijr} \leq (Q - q_i) x_{ijr} \qquad (i,j) \in \mathcal{A}, r \in \mathcal{R} \tag{7}$$

$$\sum_{(i,j)\in\mathcal{A}} \sum_{r\in\mathcal{R}} c_{ij} x_{ijr} \leq T_{max} \tag{8}$$

$$x_{ijr} \in \{0,1\} \qquad (i,j) \in \mathcal{A}, r \in \mathcal{R} \tag{9}$$

$$f_{ijr} \geq 0 \qquad (i,j) \in \mathcal{A}, r \in \mathcal{R}. \tag{10}$$

The objective function (1) minimizes the total en-route time. Constraints (2) bounds the number of vehicles. Constraints (3) and (4) ensure that each customer is visited exactly once. Constraints (5) impose that a vehicle cannot start route $r+1$ before finish route r. Constraints (6) and (7) define the flows. Constraints (8) ensure that the total travel time cannot exceed the maximum allowed working duration. Finally, constraints (9) and (10) enforce the integrality and nonnegativity restrictions on the variables.

3. Solution approach

The mathematical formulation of the problem is a member of a rich VRP family [4], which is hard to solve optimally as it requires the joint solution several difficult subproblems. To overcome this barrier, we now present a GIS-based solution approach.

In practice, there are several commercial programs are available to solve the VRP and its

extensive variations. GIS is a kind of system that provides spatial analyses and supports the decision-making activities by using various geographic data. It can also support logistic and marketing managers to evaluate placement options. Thus, GIS is used for replenishment of automated teller machines in Gaziantep [28]. We used the ArcGIS 10.2 commercial package to solve our optimization problem and also for building our GIS-based decision support framework. The ArcGIS is frequently used in many broad areas where spatially-enabled data need to be stored, retrieved, analyzed, visualized and even served online [29]. The ArcGIS is first used as a platform to store all problem data in geographic format. It visualizes all data as well as the solutions we obtain through our heuristic approach.

The software platform commercial package ArcGIS uses a tabu search heuristic algorithm to solve our defined problem. The solution method follows the classical tabu search principles such as non-improving solutions are accepted along the way. However, cycling of solutions are avoided using tabu lists and tabu tenure parameters [30]. In the last decades, tabu search heuristics are commonly used in VRP and its variants. It obtains quite competetive solutions and it is still an highly effective heuristic method [31–33]. Initialization phase creates an origin-destination matrix of shortest travel costs between all locations that must be visited by a route. A feasible initial solution is then generated by inserting each location one at a time into the most suitable route. The improvement phase aims to obtain high quality solution by applying the following procedures.

- Changing the sequence nodes on a single route.
- Moving a single node from its current route to a better route.
- Swapping two nodes between their respective routes.

Figure 2 shows the framework of the system proposed in a form of a diagram.

4. A case study

We now present a case study arising in one of the major bank operating in Turkey. The considered bank group is an integrated financial services group operating in every segment of the banking sector including corporate, commercial, small and medium-sized enterprises, payment systems, retail, private and investment banking together with its subsidiaries in pension and life insurance, leasing, factoring, brokerage, and asset management. As of September 2017, the bank group provides a wide range of financial services to its tens of million customers through an extensive distribution network of 942 domestic branches with 4,769 ATMs.

To manage the money flow between branches and ATMs, the bank group aims to speed up decision-making and implementation processes by establishing a well-designed logistic network. To do so, one district office, 12 branches and 53 ATMs which are located in the city of Gaziantep are considered to be designed. The locations of the ATMs, and district office and branches are illustrated in Figures 3 and 4, respectively.

Gaziantep with its 1,975,302 population in 2016 is the 8th most crowded city of Turkey and it is an important commercial and industrial center for Turkey. The considered stores are located in two districts which cover 85% of total population of Gaziantep. In total, there are 5 benchmark instances, i.e., GB-1, GB-2, GB-3, GB-4, and GB-5, which include all ATMs with different demands range from 5 to 45 money trays. Solving a network analysis problem in ArcGIS software, several parameters shown below have to be utilized in our study. Figure 5 shows an example of the user interface of ArcGIS of parameter entry. Table 1 presents the detailed information about benchmark instances. The first column shows the ATM number, while others present the daily demand.

- Vehicle number : Bank group has four vehicles.
- Vehicle Capacity: Each vehichle has the same capacity (200 money trays) and type.
- Each ATM has a service time which includes the loading money and handling paperwork for shipment.
- Start Depot: Each vehicle starts from district office.
- Maximum Travel Time: Each vehicles duration time is fixed at 6 hours.
- Vehicle Speed: Speed is fixed at 50 km/h.
- Distance Attribute: Road length is selected as distance attribute.
- Restrictions: One-way traffic is set as road restrictions.

5. Computational experiments and analyses

This section presents the results of the computational experiments. All experiments were conducted on a server with an Intel Core i7 CPU 3.07

Table 1. The detailed information about benchmark instances.

ATM No	GB1	GB2	GB3	GB4	GB5
1	11	12	10	12	13
2	22	23	21	23	24
3	22	23	21	23	24
4	11	12	10	12	13
5	22	23	21	23	24
6	11	12	10	12	13
7	11	12	10	12	13
8	22	23	21	23	24
9	43	44	42	44	45
10	17	18	16	18	19
11	11	12	10	12	13
12	11	12	10	12	13
13	11	12	10	12	13
14	17	18	16	18	19
15	17	18	16	18	19
16	11	12	10	12	13
17	22	23	21	23	24
18	22	23	21	23	24
19	6	7	5	7	8
20	22	23	21	23	24
21	33	34	32	34	35
22	22	23	21	23	24
23	33	34	32	34	35
24	33	34	32	34	35
25	22	23	21	23	24
26	22	23	21	23	24
27	11	12	10	12	13
28	33	34	32	34	35
29	6	7	5	7	8
30	11	12	10	12	13
31	11	12	10	12	13
32	11	12	10	12	13
33	11	12	10	12	13
34	11	12	10	12	13
35	6	7	5	7	8
36	6	7	5	7	8
37	11	12	10	12	13
38	11	12	10	12	13
39	6	7	5	7	8
40	6	7	5	7	8
41	6	7	5	7	8
42	11	12	10	12	13
43	11	12	10	12	13
44	17	18	16	18	19
45	27	28	26	28	29
46	27	28	26	28	29
47	22	23	21	23	24
48	22	23	21	23	24
49	17	18	16	18	19
50	33	34	32	34	35
51	17	18	16	18	19
52	17	18	16	18	19
53	11	12	10	12	13

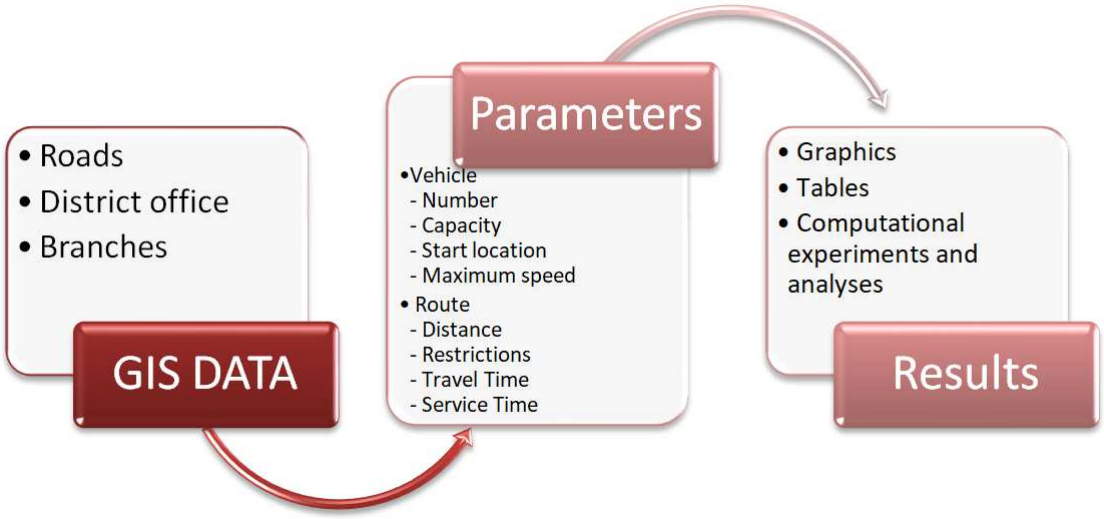


Figure 2. The framework of the system proposed in a form of a diagram.

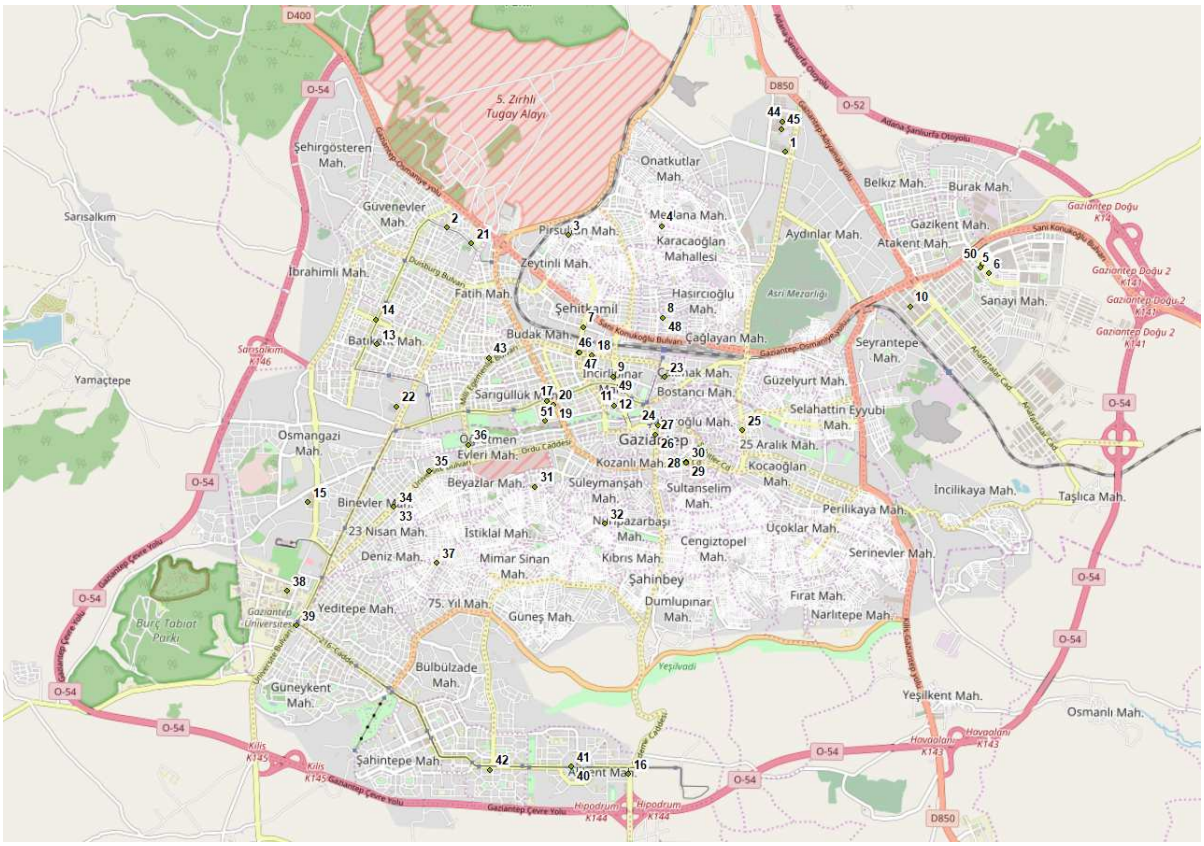


Figure 3. Locations of ATMs.

Ghz processor. The analysis of the ArcGIS for the case study has considered fixed parameters.

Tables 2–6 presents the results obtained on 5 benchmark instances of the bank. The illustrations of solutions are given in Figures 6–10. In our experiments, we first relax the starting from the district office and returning to the district office constraint and presents its results in the first part of the table. We then present the results of the considered problem in the second part of the table. In each table, the first column shows the vehicle and its tour number. For example, "1/1" indicates that the first vehicle's first tour. The second and third columns show the start and end nodes of the vehicle tour, respectively. The other columns show the total number of orders, total

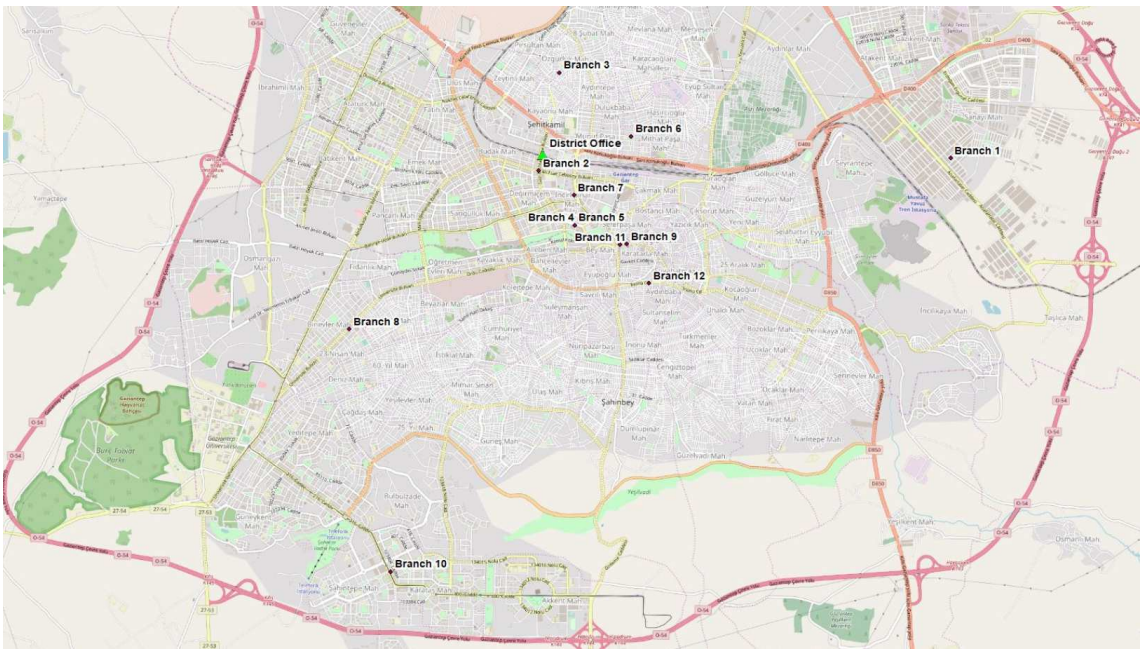


Figure 4. Locations of district office and branches.

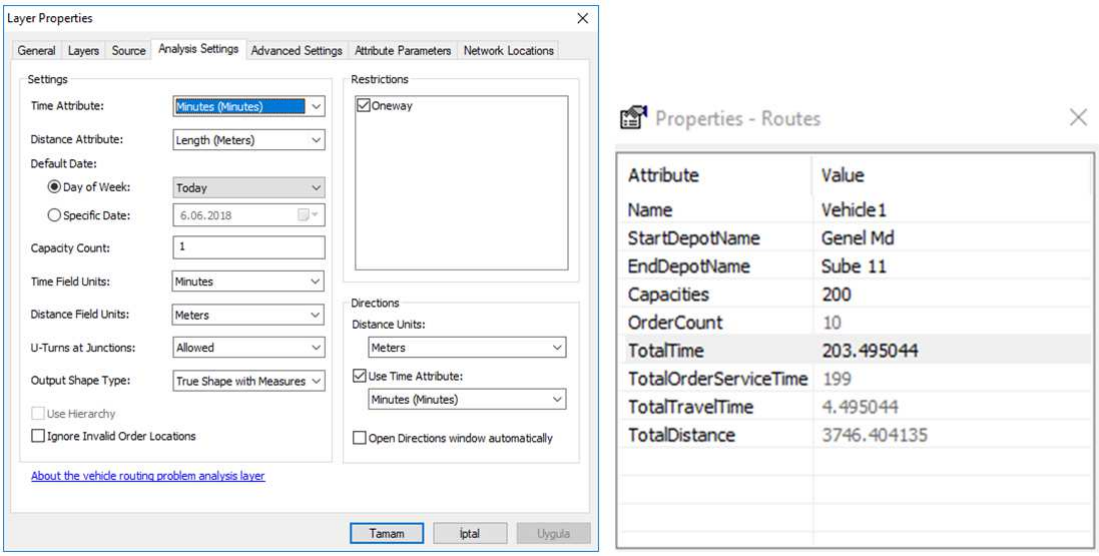


Figure 5. An example of the user interface of ArcGIS of parameter entry.

travel time (seconds), the total distance (km), total service time (seconds), and total en-route time (seconds), respectively.

Solution times for each instance are less than two seconds. Table 2 shows that all vehicles are used once. Tables 3, 4, and 5 show that vehicle 1 used two times, but vehicles 2, 3, and 4 only used once. Table 6 shows that vehicle 1 and 3 used two times, vehicles 2 and 4 used only once. Total distances are 84.32, 88.79, 84.76, 85.35, and 92.52 km for GB-1, GB-2, GB-3, GB-4, and GB-5, respectively. Total travel times are 101.17, 106.53, 101.70, 102.40, and 111.01 seconds for GB-1, GB-2, GB-3, GB-4, and GB-5, respectively. Total en-route times are 879.17, 937.53, 985.70, 1039.40, and 1101.01, respectively.

When we remove the each vehicle route starts and ends at the depot constraint, Tables 2–6 show that vehicle 1 used two times, vehicle 2 used two times, and vehicle 3 used only once. Total distances are 69.67, 71.71, 70.74, 79.44, and 82.52 km for GB-1, GB-2, GB-3, GB-4, and GB-5, respectively. Total travel times are 83.59, 86.04, 84.88, 95.31, and 99.01 seconds for GB-1, GB-2, GB-3, GB-4, and GB-5, respectively. Total en-route times are 861.59, 917.04, 968.88, 968.88, 1032.31,

and 1089.01 seconds for GB-1, GB-2, GB-3, GB-4, and GB-5, respectively.

When we compared our results with current one used by the bank for its daily operation, our results provided better solutions. On average, in terms of total distance, total travel times, and total en-route times, our method obtained 9.52%, 10.51%, and 10.65% better solutions. These results show that total distances are reduced when we relax each vehicle route starts and ends at the depot constraint. Similarly, total travel times and total en-route times are also reduced. Our results indicate that four vehicle are enough for satisfying ATM demands, and in general more than one vehicle tour is not necessary.

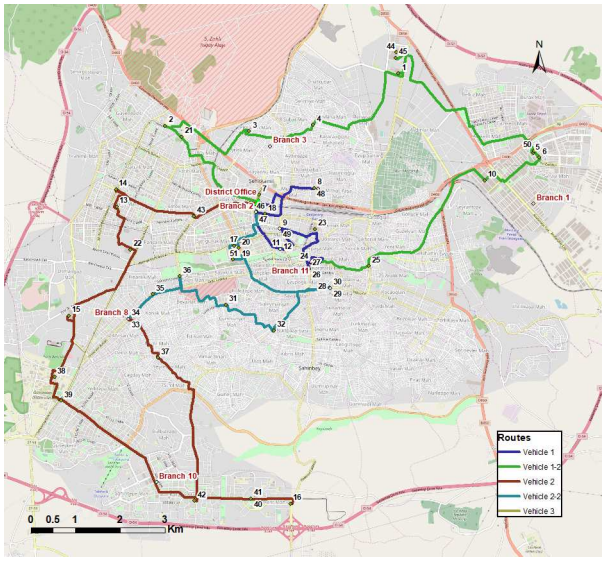


Figure 6. Solution of GB-1.

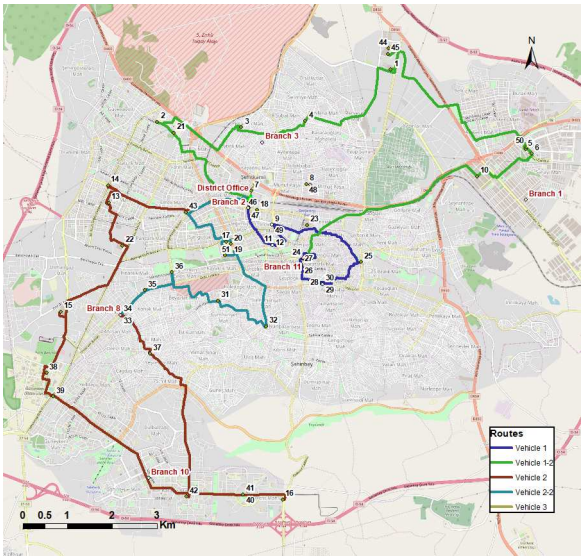


Figure 7. Solution of GB-2.

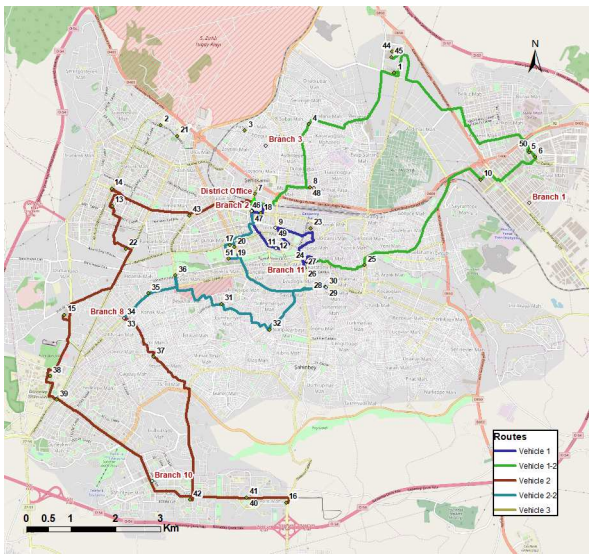


Figure 8. Solution of GB-3.

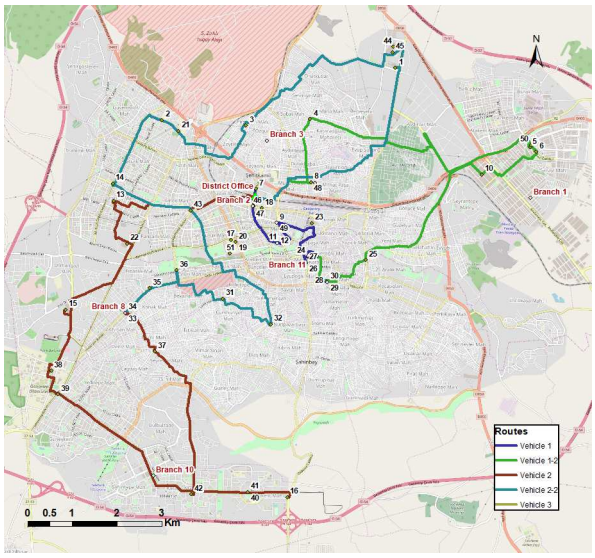


Figure 9. Solution of GB-4.

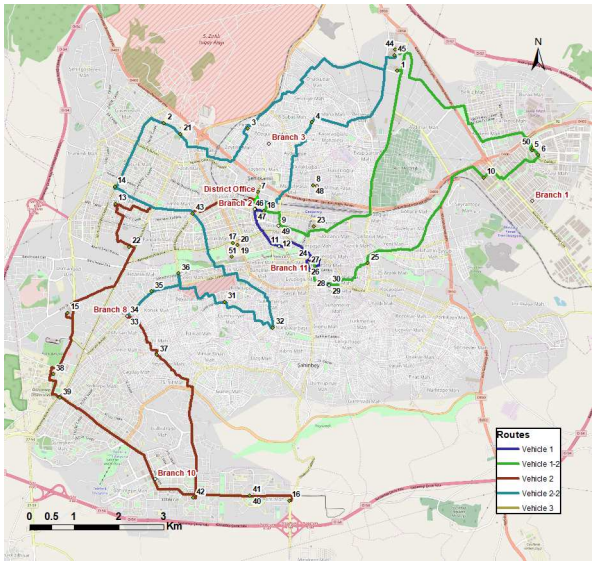


Figure 10. Solution of GB-5.

Table 2. The detailed results of instance GB-1.

Vehicle/ Tour	Starting node	End node	Total order	Total distance	Total travel time	Total service time	Total en-route time
1/1	District office	Branch 11	14	8.20	9.84	189	198.84
1/2	Branch 11	District Office	12	24.03	28.83	192	220.83
2/1	District Office	Branch 8	12	25.04	30.04	196	226.04
2/2	Branch 8	District Office	14	12.01	14.41	186	200.41
3/1	District Office	District Office	1	0.39	0.47	15	15.47
Total			53	69.67	83.59	778	861.59
1/1	District Office	District Office	18	23.60	28,32	196	224.32
2/1	District Office	District Office	13	17.31	20,77	192	212.77
3/1	District Office	District Office	13	25.06	30,07	200	230.07
4/1	District Office	District Office	9	18.34	22,01	190	212.01
Total			53	84.32	101.17	778	879.17

Table 3. The detailed results of instance GB-2.

Vehicle/ Tour	Starting node	End node	Total order	Total distance	Total travel time	Total service time	Total en-route time
1/1	District Office	Branch 11	14	7.86	9.43	181	190.43
1/2	Branch 11	District Office	11	23.54	28.25	199	227.25
2/1	District Office	Branch 8	11	25.01	30.01	187	217.01
2/2	Branch 8	District Office	11	11.50	13.80	175	188.80
3/1	District Office	District Office	6	3.80	4.56	89	93.56
Total			53	71.71	86.04	831	917.04
1/1	District Office	District Office	14	26.57	31.88	168	199.88
1/2	District Office	District Office	9	15.73	18.87	157	175.87
2/1	District Office	District Office	12	16.89	20.26	184	204.26
3/1	District Office	District Office	13	25.59	30.70	192	222.70
4/1	District Office	District Office	5	4.02	4.82	130	134.82
Total			53	88.79	106.53	831	937.53

Table 4. The detailed results of instance GB-3.

Vehicle/ Tour	Starting node	End node	Total order	Total distance	Total travel time	Total service time	Total en-route time
1/1	District Office	Branch 11	12	5.20	6.24	195	201.24
1/2	Branch 11	District Office	12	19.61	23.53	200	223.53
2/1	District Office	Branch 8	11	25.01	30.01	198	228.01
2/2	Branch 8	District Office	13	11.96	14.35	197	211.35
3/1	District Office	District Office	5	8.96	10.75	94	104.75
Total			53	70.74	84.88	884	968.88
1/1	District Office	District Office	12	6.70	8.04	195	203.04
1/2	District Office	District Office	10	20.37	24.44	197	221.44
2/1	District Office	District Office	13	28.76	34.51	193	227.51
3/1	District Office	District Office	7	11.63	13.95	127	140.95
4/1	District Office	District Office	11	17.30	20.76	172	192.76
Total			53	84.76	101.70	884	985.70

Table 5. The detailed results of instance GB-4.

Vehicle/ Tour	Starting node	End node	Total order	Total distance	Total travel time	Total service time	Total en-route time
1/1	District Office	Branch 11	10	5.08	6,09	166	172.09
1/2	Branch 11	District Office	12	18.40	22,08	195	217.08
2/1	District Office	Branch 8	11	24.43	29,31	198	227.31
2/2	Branch 8	District Office	13	27.12	32,54	197	229.54
3/1	District Office	District Office	7	4.41	5.29	181	186.29
Total			53	79.44	95.31	937	1032.31
1/1	District Office	District Office	11	6.70	8.04	189	197.04
1/2	District Office	District Office	11	16.16	19.39	194	213.39
2/1	District Office	District Office	11	17.59	21.11	197	218.11
3/1	District Office	District Office	8	13.55	16.26	167	183.26
4/1	District Office	District Office	12	31.35	37.61	190	227.61
Total			53	85.35	102.40	937	1039.40

Table 6. The detailed results of instance GB-5.

Vehicle/ Tour	Starting node	End node	Total order	Total distance	Total travel time	Total service time	Total en-route time
1/1	District Office	Branch 11	10	3.75	4.50	199	203.50
1/2	Branch 11	District Office	11	20.77	24.92	199	223.92
2/1	District Office	Branch 8	10	24.43	29.31	196	225.31
2/2	Branch 8	District Office	13	2664	31.96	199	230.96
3/1	District Office	District Office	9	6.94	8.33	197	205.33
Total			53	82.52	99.01	990	1089.01
1/1	District Office	District Office	12	13.02	15.63	175	190.63
1/2	District Office	District Office	9	30.10	36.11	190	226.11
2/1	District Office	District Office	10	17.57	21.08	175	196.08
3/1	District Office	District Office	11	25.26	30.31	198	228.31
3/2	District Office	District Office	3	1.04	1.24	62	63.24
4/1	District Office	District Office	8	5.53	6.64	190	196.64
Total			53	92.52	111.01	990	1101.01

6. Conclusions

This paper has been motivated by the problem faced by one of the major banks of Turkey operating in the city of Gaziantep. We have defined a new rich vehicle routing problem which is concerned with the joint multiple depots, pickup and delivery, multi-trip, and homogeneous fixed vehicle fleet. We have presented a mathematical formulation for the problem. To obtain fast and good quality solutions, we have then used a GIS-based solution approach employs a tabu search algorithm which can be used to store, analyze and visualize all data as well as model solutions in geographic format. We have considered a real dataset of the bank and have applied our GIS-based solution approach. We have finally provided several managerial and policy insights on results by exploring the problem.

Our results indicated that four vehicles are enough to satisfy the demand of ATMs for the bank and one vehicle tour is also enough for each vehicle in general. We have also shown that total distances are reduced if we do not consider each vehicle route starts and ends at the depot constraint. In a similar manner, total travel times and total en-route times are also reduced. Furthermore, the running times of the algorithm are so small that it can be used in practical bank operations.

For future studies, stochasticity and dynamism can be taken account in the problem definition, instead of using deterministic parameters. This

would require new mathematical models and solution algorithms, such as stochastic optimization. Furthermore, new effective exact methods can be developed, such as Lagrangean relaxation to obtain lower bounds, or decomposition techniques to solve large size benchmark instances to optimality.

Acknowledgements

The authors thank the three anonymous referees for their insightful comments and suggestions that helped improve the content and the presentation of the paper.

References

[1] Özceylan, E., Uslu, A., Erbaş, M., Çetinkaya, C., İşleyen, S. K. (2017). Optimizing the location-allocation problem of pharmacy warehouses: A case study in Gaziantep. An International Journal of Optimization and Control: Theories & Applications (IJOCTA), 7, 117–129.

[2] Laporte, G. (2009). Fifty years of vehicle routing. Transportation Science, 43, 408–416.

[3] Toth, P. Vigo, D., eds. (2014). Vehicle routing: Problems, methods, and applications. MOS-SIAM Series on Optimization, Philadelphia.

[4] Lahyani, R., Khemakhem, M., Semet, F. (2015). Rich vehicle routing problems: From a taxonomy to a definition. European Journal of Operational Research, 241, 1–14.

[5] Karaoglan, A. D., Gonen, D., Ucmus, E. (2011). Aircraft routing and scheduling: A case study in an airline company. An International Journal of Optimization and Control: Theories & Applications (IJOCTA), 1, 27–43.

[6] Karagul, K., Gungor, I. (2014). A case study of heterogeneous fleet vehicle routing problem: Touristic distribution application in Alanya. An International Journal of Optimization and Control: Theories & Applications (IJOCTA), 4, 67–76.

[7] Van Anholt, R. G., Coelho, L. C., Laporte, G., Vis, I. F. (2016). An inventory-routing problem with pickups and deliveries arising in the replenishment of automated teller machines. Transportation Science, 50, 1077–1091.

[8] Karagul, K., Aydemir, E., Tokat, S. (2016). Using 2-Opt based evolution strategy for travelling salesman problem. An International Journal of Optimization and Control: Theories & Applications (IJOCTA), 6, 103–113.

[9] Crevier, B., Cordeau, J. F., Laporte, G. (2007). The multi-depot vehicle routing problem with inter-depot routes. European Journal of Operational Research, 176, 756–773.

[10] Braekers, K., Caris, A., Jenssens, G. K. (2014). Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. Transportation Research Part B, 67, 166–186.

[11] Contardo, C., Martinelli, R. (2014). A new exact algorithm for the multi-depot vehicle routing problem under capacity and route length constraints. Discrete Optimization, 12, 129–146.

[12] Montoya-Torres, J. R., Franco, J. L., Isaza, S. N., Jimnez, H. F., Herazo-Padilla, N. (2015). A literature review on the vehicle routing problem with multiple depots. Computers & Industrial Engineering, 79, 115–129.

[13] Koç, Ç., Karaoğlan, İ. (2012). A mathematical model for the vehicle routing problem with time windows and multiple use of vehicles. Journal of the Faculty of Engineering and Architecture of Gazi University, 27, 569–576.

[14] Cattaruzza, D., Absi, N., Feillet, D., Vidal, T. (2014). A memetic algorithm for the multi trip vehicle routing problem. European Journal of Operational Research, 236, 833–848.

[15] Olivera, A., Viera, O. (2007). Adaptive memory programming for the vehicle routing problem with multiple trips. Computers & Operations Research, 34, 28–47.

[16] Xu, H., Chen, Z. L., Rajagopal, S., Arunapuram, S. (2003). Solving a practical pickup and delivery problem. Transportation Science, 37, 347–364.

[17] Sigurd, M., Pisinger, D., Sig, M. (2004). Scheduling transportation of live animals to avoid the spread of diseases. Transportation Science, 38, 197–209.

[18] Battarra, M., Cordeau, J-F., Iori, M. (2014). Pickup-and-delivery problems for goods transportation. In Toth, P. Vigo, D., eds. Vehicle Routing: Problems, Methods, and Applications (pp. 161–192). MOS-SIAM Series on Optimization, Philadelphia.

[19] Berbeglia, G., Cordeau, J.-F., Gribkovskaia, I., Laporte, G. (2007). Static pickup and delivery problems: A classification scheme and survey. TOP: An Official Journal of the Spanish Society of Statistics and Operations Research, 15, 1–31.

[20] Koç, Ç., Laporte, G. (2018). Vehicle routing with backhauls: Review and research perspectives. Computers & Operations Research, 91, 79–91.

[21] Parragh, S. N., Doerner, K. F., Hartl, R. F. (2008a). A survey on pickup and delivery problems. Part I: Transportation between customers and depot. Journal für Betriebswirtschaft, 58, 21–51.

[22] Parragh, S. N., Doerner, K. F., Hartl, R. F. (2008b). A survey on pickup and delivery problems. Part II: Transportation between pickup and delivery locations. Journal für Betriebswirtschaft, 58, 81–117.

[23] Casas, I., Malik, A., Delmelle, E. M., Karwan, M. H., Batta, R. (2007). An automated network generation procedure for routing of unmanned aerial vehicles (UAVs) in a GIS environment. Networks and Spatial Economics, 7, 153–176.

[24] Bozkaya, B., Yanik, S., Balcisoy, S. (2010). A GIS-based optimization framework for competitive multi-facility location-routing problem. Networks and Spatial Economics, 10, 297–320.

[25] Samanlioglu, F. (2013). A multi-objective mathematical model for the industrial hazardous waste location-routing problem. European Journal of Operational Research, 226, 332–340.

[26] Yanik, S., Bozkaya, B., de Kervenoael, R. (2014). A new VRPPD model and a hybrid heuristic solution approach for e-tailing. European Journal of Operational Research, 236, 879–890.

[27] Krichen, S., Faiz, S., Tlili, T., Tej, K. (2014). Tabu-based GIS for solving the vehicle routing problem. Expert Systems with Applications, 41, 6483–6493.

[28] Vlachopoulou, M., Silleos, G., Manthou, V. (2001). Geographic information systems in warehouse site selection decisions. International Journal of Production Economics, 71, 205–212.
[29] ArcGIS, (2017). ArcGIS Network Analyst Tutorial. `http://desktop.arcgis.com/en/arcmap/`.
[30] Glover, F.W., Laguna, M. (1998). Tabu search. Kluwer Academic, Massachusetts.
[31] Brandão, J. (2009). A deterministic tabu search algorithm for the fleet size and mix vehicle routing problem. European Journal of Operational Research, 195, 716–728.
[32] Brandão, J. (2011). A tabu search algorithm for the heterogeneous fixed fleet vehicle routing problem. Computers & Operations Research, 38, 140–151.
[33] Gendreau, M., Hertz, A., Laporte, G. (1994). A tabu search heuristic for the vehicle routing problem. Management Science, 40, 1276–1290.

Çağrı Koç *is an Assistant Professor in Department of Business Administration at Social Sciences University of Ankara. He was a Postdoctoral Fellow at HEC Montreal, Canada Research Chair in Distribution Management, and at CIRRELT (Interuniversity Research Center on Enterprise Networks, Logistics and Transportation). He received his Ph.D. degree (2015) in Management Science from the Southampton Business School of University of Southampton. He is the recipient of the Operational Research Society 2015 Doctoral Dissertation Award. His research mainly focuses on transportation and logistics, supply chain management, vehicle routing and scheduling, and mathematical and metaheuristic optimization.*

Mehmet Erbaş *is working in General Directorate of Mapping, Ministry of National Defense. He worked as a Lecturer in the Geomatics Division of the Department of Civil Engineering Geomatics Division at the Turkish Military Academy in Ankara. His research and teaching interests focus on the topics of the geographic information systems and remote sensing applications.*

Eren Özceylan *is an Associate Professor in Department of Industrial Engineering at Gaziantep University. He received his Ph.D. degree in Computer Engineering from Seluk University in 2013. His research focuses on logistics and supply chain management. In particular, he focuses on supply chain network design, environmental conscious production/distribution, multi-criteria decision-making and fuzzy logic.*

INSTRUCTIONS FOR AUTHORS

Aims and Scope

This journal shares the research carried out through different disciplines in regards to optimization, control and their applications.

The basic fields of this journal are linear, nonlinear, stochastic, parametric, discrete and dynamic programming; heuristic algorithms in optimization, control theory, game theory and their applications. Problems such as managerial decisions, time minimization, profit maximizations and other related topics are also shared in this journal.

Besides the research articles expository papers, which are hard to express or model, conference proceedings, book reviews and announcements are also welcome.

Journal Topics

- Applied Mathematics,
- Financial Mathematics,
- Control Theory,
- Game Theory,
- Fractional Calculus,
- Fractional Control,
- Modeling of Bio-systems for Optimization and Control,
- Linear Programming,
- Nonlinear Programming,
- Stochastic Programming,
- Parametric Programming,
- Conic Programming,
- Discrete Programming,
- Dynamic Programming,
- Optimization with Artificial Intelligence,
- Operational Research in Life and Human Sciences,
- Heuristic Algorithms in Optimization,
- Applications Related to Optimization on Engineering.

Submission of Manuscripts

New Submissions

Solicited and contributed manuscripts should be submitted to IJOCTA via the journal's online submission system. You need to make registration prior to submitting a new manuscript (please click here to register and do not forget to define yourself as an "Author" in doing so). You may then click on the "New Submission" link on your User Home.

IMPORTANT: If you already have an account, please click here to login. It is likely that you will have created an account if you have reviewed or authored for the journal in the past.

On the submission page, enter data and answer questions as prompted. Click on the “Next” button on each screen to save your work and advance to the next screen. The names and contact details of at least four internationally recognized experts who can review your manuscript should be entered in the "Comments for the Editor" box.

You will be prompted to upload your files: Click on the “Browse” button and locate the file on your computer. Select the description of the file in the drop down next to the Browse button. When you have selected all files you wish to upload, click the “Upload” button. Review your submission before sending to the Editors. Click the “Submit” button when you are done reviewing. Authors are responsible for verifying all files have uploaded correctly.

You may stop a submission at any phase and save it to submit later. Acknowledgment of receipt of the manuscript by IJOCTA Online Submission System will be sent to the corresponding author, including an assigned manuscript number that should be included in all subsequent correspondence. You can also log-

on to submission web page of IJOCTA any time to check the status of your manuscript. You will receive an e-mail once a decision has been made on your manuscript.

Each manuscript must be accompanied by a statement that it has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere.

Manuscripts can be prepared using LaTeX (.tex) or MSWord (.docx). However, manuscripts with heavy mathematical content will only be accepted as LaTeX files.

Preferred first submission format (for reviewing purpose only) is Portable Document File (.pdf). Please find below the templates for first submission.

Click here to download Word template for first submission (.docx)

Click here to download LaTeX template for first submission (.tex)

Revised Manuscripts

Revised manuscripts should be submitted via IJOCTA online system to ensure that they are linked to the original submission. It is also necessary to attach a separate file in which a point-by-point explanation is given to the specific points/questions raised by the referees and the corresponding changes made in the revised version.

To upload your revised manuscript, please go to your author page and click on the related manuscript title. Navigate to the "Review" link on the top left and scroll down the page. Click on the "Choose File" button under the "Editor Decision" title, choose the revised article (in pdf format) that you want to submit, and click on the "Upload" button to upload the author version. Repeat the same steps to upload the "Responses to Reviewers/Editor" file and make sure that you click the "Upload" button again.

To avoid any delay in making the article available freely online, the authors also need to upload the source files (Word or LaTex) when submitting revised manuscripts. Files can be compressed if necessary. The two-column final submission templates are as follows:

Click here to download Word template for final submission (.docx)

Click here to download LaTeX template for final submission (.tex)

Authors are responsible for obtaining permission to reproduce copyrighted material from other sources and are required to sign an agreement for the transfer of copyright to IJOCTA.

Article Processing Charges

There are no charges for submission and/or publication.

English Editing

Papers must be in English. Both British and American spelling is acceptable, provided usage is consistent within the manuscript. Manuscripts that are written in English that is ambiguous or incomprehensible, in the opinion of the Editor, will be returned to the authors with a request to resubmit once the language issues have been improved. This policy does not imply that all papers must be written in "perfect" English, whatever that may mean. Rather, the criteria require that the intended meaning of the authors must be clearly understandable, i.e., not obscured by language problems, by referees who have agreed to review the paper.

Presentation of Papers

Manuscript style

Use a standard font of the **11-point type: Times New Roman** is preferred. It is necessary to single line space your manuscript. Normally manuscripts are expected not to exceed 25 single-spaced pages including text, tables, figures and bibliography. All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.

During the submission process you must enter: (1) the full title, (2) names and affiliations of all authors and (3) the full address, including email, telephone and fax of the author who is to check the proofs. Supply a brief **biography** of each author at the end of the manuscript after references.

- Include the name(s) of any **sponsor(s)** of the research contained in the paper, along with **grant number(s).**
- Enter an **abstract** of no more than 250 words for all articles.

Keywords

Authors should prepare no more than 5 keywords for their manuscript.

Maximum five **AMS Classification number** (http://www.ams.org/mathscinet/msc/msc2010.html) of the study should be specified after keywords.

Writing Abstract

An abstract is a concise summary of the whole paper, not just the conclusions. The abstract should be no more than 250 words and convey the following:

1. An introduction to the work. This should be accessible by scientists in any field and express the necessity of the experiments executed.
2. Some scientific detail regarding the background to the problem.
3. A summary of the main result.
4. The implications of the result.
5. A broader perspective of the results, once again understandable across scientific disciplines.

It is crucial that the abstract conveys the importance of the work and be understandable without reference to the rest of the manuscript to a multidisciplinary audience. Abstracts should not contain any citation to other published works.

Reference Style

Reference citations in the text should be identified by numbers in square brackets "[ ]". All references must be complete and accurate. Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Online citations should include date of access. References should be listed in the following style:

Journal article

Author, A.A., & Author, B. (Year). Title of article. Title of Journal, Vol(Issue), pages.

Castles, F.G., Curtin, J.C., & Vowles, J. (2006). Public policy in Australia and New Zealand: The new global context. Australian Journal of Political Science, 41(2), 131–143.

Book

Author, A. (Year). Title of book. Publisher, Place of Publication.

Mercer, P.A., & Smith, G. (1993). Private Viewdata in the UK. 2nd ed. Longman, London.

Chapter

Author, A. (Year). Title of chapter. In: A. Editor and B. Editor, eds. Title of book. Publisher, Place of publication, pages.

Bantz, C.R. (1995). Social dimensions of software development. In: J.A. Anderson, ed. Annual review of software management and development. CA: Sage, Newbury Park, 502–510.

Internet document

Author, A. (Year). Title of document [online]. Source. Available from: URL [Accessed (date)].

Holland, M. (2004). Guide to citing Internet sources [online]. Poole, Bournemouth University. Available from: http://www.bournemouth.ac.uk/library/using/guide_to_citing_internet_sourc.html [Accessed 4 November 2004].

Newspaper article

Author, A. (or Title of Newspaper) (Year). Title of article. Title of Newspaper, day Month, page, column.

Independent (1992). Picking up the bills. Independent, 4 June, p. 28a.

Thesis

Author, A. (Year). Title of thesis. Type of thesis (degree). Name of University.

Agutter, A.J. (1995). The linguistic significance of current British slang. PhD Thesis. Edinburgh University.

Illustrations

Illustrations submitted (line drawings, halftones, photos, photomicrographs, etc.) should be clean originals or digital files. Digital files are recommended for highest quality reproduction and should follow these guidelines:

- 300 dpi or higher
- Sized to fit on journal page
- TIFF or JPEG format only
- Embedded in text files and submitted as separate files (if required)

Tables and Figures

Tables and figures (illustrations) should be embedded in the text at the appropriate points, rather than at the end. A short descriptive title should appear above each table with a clear legend and any footnotes suitably identified below.

Proofs

Page proofs are sent to the designated author using IJOCTA EProof system. They must be carefully checked and returned within 48 hours of receipt.

Offprints/Reprints

Each corresponding author of an article will receive a PDF file of the article via email. This file is for personal use only and may not be copied and disseminated in any form without prior written permission from IJOCTA.

Submission Preparation Checklist

As part of the submission process, authors are required to check off their submission's compliance with all of the following items, and submissions may be returned to authors that do not adhere to these guidelines.

1. The submission has not been previously published, nor is it before another journal for consideration (or an explanation has been provided in Comments for the Editor).
2. The submission file is in Portable Document Format (.pdf).
3. Where available, URLs for the references have been provided.
4. The text is single line spaced; uses a 11-point font; employs italics, rather than underlining (except with URL addresses); and all illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.
5. The text adheres to the stylistic and bibliographic requirements outlined in the Author Guidelines, which is found in "About the Journal".
6. Maximum five AMS Classification number (http://www.ams.org/mathscinet/msc/msc2010.html) of the study have been provided after keywords.
7. After the acceptance of manuscript (before copy editing), Word (.docx) or LaTeX (.tex) version of the paper will be presented.
8. The names and email addresses of at least four (4) possible reviewers have been indicated in "Comments for the Editor" box in Paper Submission Step 1. Please note that at least two of the recommendations should be from different countries. Avoid suggesting reviewers who are at arms-length from you or your co-authors. This includes graduate advisors, people in your current department, or any others with a conflict of interest.

Peer Review Process

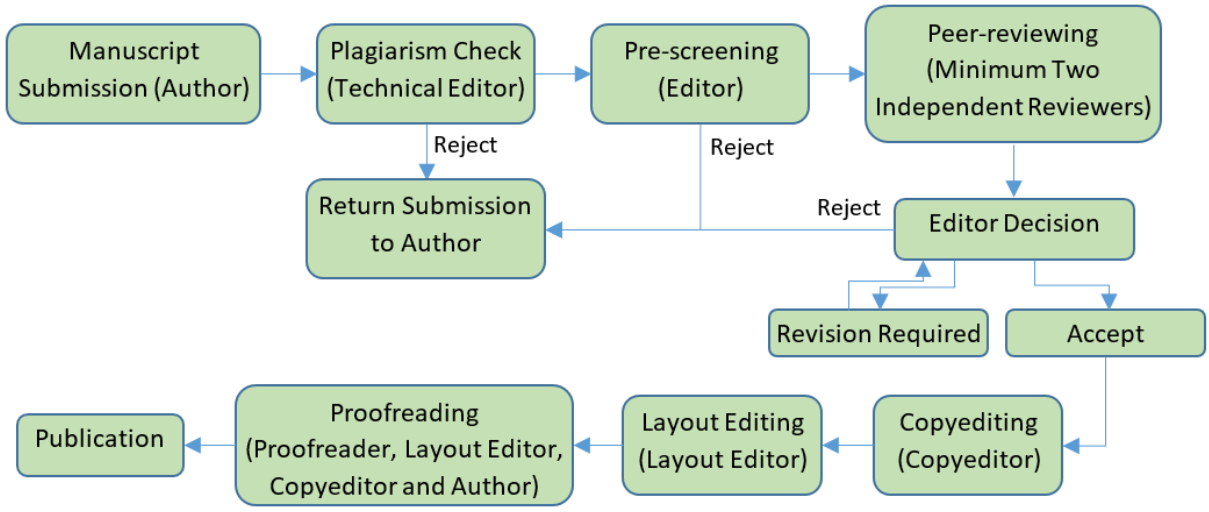
All contributions, prepared according to the author guidelines and submitted via IJOCTA online submission system are evaluated according to the criteria of originality and quality of their scientific content. The corresponding author will receive a confirmation e-mail with a reference number assigned to the paper, which he/she is asked to quote in all subsequent correspondence.

All manuscripts are first checked by the Technical Editor using plagiarism detection software (iThenticate) to verify originality and ensure the quality of the written work. If the result is not satisfactory (i.e. exceeding the limit of 30% of overlapping), the submission is rejected and the author is notified.

After the plagiarism check, the manuscripts are evaluated by the Editor-in-Chief and can be rejected without reviewing if considered not of sufficient interest or novelty, too preliminary or out of the scope of the journal. If the manuscript is considered suitable for further evaluation, it is first sent to the Area Editor. Based on his/her opinion the paper is then sent to at least two independent reviewers. Each reviewer is allowed up to four weeks to return his/her feedback but this duration may be extended based on his/her availability. IJOCTA has instituted a blind peer review process where the reviewers' identities are not known to authors. When the reviews are received, the Area Editor gives a decision and lets the author know it together with the reviewer comments and any supplementary files.

Should the reviews be positive, the authors are expected to submit the revised version usually within two months the editor decision is sent (this period can be extended when the authors contact to the editor and let him/her know that they need extra time for resubmission). If a revised paper is not resubmitted within the deadline, it is considered as a new submission after all the changes requested by reviewers have been made. Authors are required to submit a new cover letter, a response to reviewers letter and the revised manuscript (which ideally shows the revisions made in a different color or highlighted). If a change in authorship (addition or removal of author) has occurred during the revision, authors are requested to clarify the reason for change, and all authors (including the removed/added ones) need to submit a written consent for the change. The revised version is evaluated by the Area editor and/or reviewers and the Editor-in-Chief brings a decision about final acceptance based on their suggestions. If necessary, further revision can be asked for to fulfil all the requirements of the reviewers.

When a manuscript is accepted for publication, an acceptance letter is sent to the corresponding author and the authors are asked to submit the source file of the manuscript conforming to the IJOCTA two-column final submission template. After that stage, changes of authors of the manuscript are not possible. The manuscript is sent to the Copyeditor and a linguistic, metrological and technical revision is made, at which stage the authors are asked to make the final corrections in no more than a week. The layout editor prepares the galleys and the authors receive the galley proof for final check before printing. The authors are expected to correct only typographical errors on the proofs and return the proofs within 48 hours. After the final check by the layout editor and the proofreader, the manuscript is assigned a DOI number, made publicly available and listed in the forthcoming journal issue. After printing the issue, the corresponding metadata and files published in this issue are sent to the databases for indexing.



Publication Ethics and Malpractice Statement

IJOCTA is committed to ensuring ethics in publication and quality of articles. Conforming to standards of expected ethical behavior is therefore necessary for all parties (the author, the editor(s), the peer reviewer) involved in the act of publishing.

International Standards for Editors

The editors of the IJOCTA are responsible for deciding which of the articles submitted to the journal should be published considering their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors. The editors may be guided by the policies of the journal's editorial board and constrained by such legal requirements as shall then be in force regarding libel, copyright infringement and plagiarism. The editors may confer with other editors or reviewers in making this decision. As guardians and stewards of the research record, editors should encourage authors to strive for, and adhere themselves to, the highest standards of publication ethics. Furthermore, editors are in a unique position to indirectly foster responsible conduct of research through their policies and processes.

To achieve the maximum effect within the research community, ideally all editors should adhere to universal standards and good practices.

- Editors are accountable and should take responsibility for everything they publish.
- Editors should make fair and unbiased decisions independent from commercial consideration and ensure a fair and appropriate peer review process.
- Editors should adopt editorial policies that encourage maximum transparency and complete, honest reporting.
- Editors should guard the integrity of the published record by issuing corrections and retractions when needed and pursuing suspected or alleged research and publication misconduct.
- Editors should pursue reviewer and editorial misconduct.
- Editors should critically assess the ethical conduct of studies in humans and animals.
- Peer reviewers and authors should be told what is expected of them.
- Editors should have appropriate policies in place for handling editorial conflicts of interest.

Reference:

Kleinert S & Wager E (2011). Responsible research publication: international standards for editors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 51 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment. Imperial College Press / World Scientific Publishing, Singapore (pp 317-28). (ISBN 978-981-4340-97-7) [Link].

International Standards for Authors

Publication is the final stage of research and therefore a responsibility for all researchers. Scholarly publications are expected to provide a detailed and permanent record of research. Because publications form the basis for both new research and the application of findings, they can affect not only the research community but also, indirectly, society at large. Researchers therefore have a responsibility to ensure that their publications are honest, clear, accurate, complete and balanced, and should avoid misleading, selective or ambiguous reporting. Journal editors also have responsibilities for ensuring the integrity of the research literature and these are set out in companion guidelines.

- The research being reported should have been conducted in an ethical and responsible manner and should comply with all relevant legislation.
- Researchers should present their results clearly, honestly, and without fabrication, falsification or inappropriate data manipulation.
- Researchers should strive to describe their methods clearly and unambiguously so that their findings can be confirmed by others.
- Researchers should adhere to publication requirements that submitted work is original, is not plagiarised, and has not been published elsewhere.
- Authors should take collective responsibility for submitted and published work.
- The authorship of research publications should accurately reflect individuals' contributions to the work and its reporting.

- Funding sources and relevant conflicts of interest should be disclosed.
- When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the journal's Editor-in-Chief and cooperate with them to either retract the paper or to publish an appropriate erratum.

Reference:

Wager E & Kleinert S (2011) Responsible research publication: international standards for authors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 50 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment. Imperial College Press / World Scientific Publishing, Singapore (pp 309-16). (ISBN 978-981-4340-97-7) [Link].

Basic principles to which peer reviewers should adhere

Peer review in all its forms plays an important role in ensuring the integrity of the scholarly record. The process depends to a large extent on trust and requires that everyone involved behaves responsibly and ethically. Peer reviewers play a central and critical part in the peer-review process as the peer review assists the Editors in making editorial decisions and, through the editorial communication with the author, may also assist the author in improving the manuscript.

Peer reviewers should:

- respect the confidentiality of peer review and not reveal any details of a manuscript or its review, during or after the peer-review process, beyond those that are released by the journal;
- not use information obtained during the peer-review process for their own or any other person's or organization's advantage, or to disadvantage or discredit others;
- only agree to review manuscripts for which they have the subject expertise required to carry out a proper assessment and which they can assess within a reasonable time-frame;
- declare all potential conflicting interests, seeking advice from the journal if they are unsure whether something constitutes a relevant conflict;
- not allow their reviews to be influenced by the origins of a manuscript, by the nationality, religion, political beliefs, gender or other characteristics of the authors, or by commercial considerations;
- be objective and constructive in their reviews, refraining from being hostile or inflammatory and from making libellous or derogatory personal comments;
- acknowledge that peer review is largely a reciprocal endeavour and undertake to carry out their fair share of reviewing, in a timely manner;
- provide personal and professional information that is accurate and a true representation of their expertise when creating or updating journal accounts.

Reference:

Homes I (2013). COPE Ethical Guidelines for Peer Reviewers, March 2013, v1 [Link].

Copyright Notice

An International Journal of Optimization and Control: Theories & Applications

Volume: 8 Number: 2
July 2018

CONTENTS

www.ijocta.com
info@ijocta.com

ISSN 2146-0957