

An open access, peer reviewed, international journal of science. Biannual (June & December). ISSN 2147-1630 | e-ISSN 2146-586X. Publisher: Adiyaman University.

Publication language: English (with Turkish title and abstract)

Issue published date: 31.12.2021

Privilege owner: On Behalf of Rectorate of Adiyaman University, Prof. Dr. Mehmet TURGUT (Rector)

Web site: EN: <https://dergipark.org.tr/en/pub/adyujsci>
TR: <https://dergipark.org.tr/tr/pub/adyujsci>

EDITORIAL BOARD

Editor-in-Chief : Deniz SUNAR ÇERÇİ, Ph.D.

Editors:

Biology : Serdar SÖNMEZ, Ph.D.
: Ertan YOLOĞLU, Ph.D.
Chemistry : Cumhur KIRILMIŞ, Ph.D.
: Gökhan ELMACI, Ph.D.
Mathematics : Selcen YÜKSEL PERKTAŞ, Ph.D.
Physics : Salim ÇERÇİ, Ph.D.
: Özge ERKEN, Ph.D.

Statistics Editor: : Tayfun SERVİ, Ph.D.

Section Editors

Biology:

Aydın AKBUDAK, Ph.D.
Bahadır AKMAN, Ph.D.
Birgül ÖZCAN, Ph.D.
Deniz AYAS, Ph.D.
Hasan YILDIZ, Ph.D.
Olga SAK, Ph.D.
Özkan ASLANTAŞ, Ph.D.
Si Hong PARK, Ph.D.
Süphan KARAYTUĞ, Ph.D.

Chemistry:

Sezgin BAKIRDERE, Ph.D.
H. Mehmet KAYILI, Ph.D.
Önder METİN, Ph.D.
Zeynel SEFEROĞLU, Ph.D.
Lokman UZUN, Ph.D.

Mathematics:

Aynur KESKİN KAYMAKÇI, Ph.D.
Bilge İNAN, Ph.D.
Eylem GÜZEL KARPUS, Ph.D.

Feyza Esra ERDOĞAN, Ph.D.
James F. PETERS, Ph.D.
Mehmet GÜLBAHAR, Ph.D.
Mehmet Onur FEN, Ph.D.
Murat CANDAN, Ph.D.
Mustafa Çağatay KORKMAZ, Ph.D.
Öznur GÖLBAŞI, Ph.D.
Ramazan AKGÜN, Ph.D.
Tahsin ÖNER, Ph.D.

Physics:

Ahmet EKİCİBİL, Ph.D.
Didar DOBUR, Ph.D.
Faruk KARADAĞ, Ph.D.
Hakan ÖZTÜRK, Ph.D.
Kalvir DHUGA, Ph.D.,
Kristina RUSIMOVA, Ph.D.
Latife ŞAHİN YALÇIN, Ph.D.
Mustafa GÜNEŞ, Ph.D.
Paolo GUNNELLINI, Ph.D.

Technical Contact: Serdar SÖNMEZ, Ph.D., ssonmez@adiyaman.edu.tr, sonmezserdar@gmail.com

Language Editors: Münevver AKBAŞ, İbrahim KAYA, Hakkı ŞİMŞEK



The articles published in this journal are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Table of Contents (İçindekiler)

Volume (Cilt): 11 Number (Sayı): 2

December (Aralık) 2021

BIOLOGY

Physiological and Molecular Effects of Exogenous Gibberellin (GA3) Treatment on Germination of Barley Seeds under Salt Stress

Arpa Tohumlarının Tuz Stresi Altında Çimlenmesine Eksojen Gibberellin (GA3) Tedavisinin Fizyolojik ve Moleküler Etkileri
Cüneyt UÇARLI 227-243

Identification of Damaging SNPs and Their Effects on Alzheimer's Disease-Associated PSEN1 Protein: Computational Analysis

Zarar Verici SNP'lerin ve Alzheimer Hastalığıyla İlişkili PSEN1 Proteinine Etkilerinin Tanımlanması: Hesaplamalı Analiz
Orcun AVSAR 321-349

Comparatively Investigation of Textile Dye Decolorization by a White Rot Fungus and Various Bacterial Strains

Bir Beyaz Çürükçül Fungus ve Çeşitli Bakteriyel Suşlar ile Tekstil Boyasının Renginin Gideriminin Karşılaştırmalı Olarak Araştırılması
Emre BİRHANLI, Özfer YEŞİLADA, Ahmet ÇABUK, Filiz BORAN, Eray TATLICI 350-361

CHEMISTRY

The Effective and Eco-friendly Tea Fungus for the Biosorption of Dye Pollutant from Aqueous Solutions

Sulu Çözeltilerdeki Boyarmadde Kirliliğinin Biyosorpsiyonu için Etkili ve Çevre Dostu Çay Mantarı
Tuğba ALP ARICI 370-384

Determination of Prooxidant Activities of Several Fruit Juices and Herbs

Çeşitli Meyve Suları ve Şifalı Bitkilerin Prooksidan Aktivitelerinin Altın Nanoküme Biyosensörleri ve Karbonil Yöntemi ile Karşılaştırılması
Esin AKYÜZ 474-486

MATHEMATICS

On Surfaces in Pseudo-Galilean Space with Prescribed Mean Curvature

Yarı-Galileo Uzayında Belli Bir Ortalama Eğriliğe Sahip Yüzeyler
Muhittin Evren AYDIN, Alper Osman ÖĞRENMİŞ 215-226

On The Convergence Properties of Kantorovich-Szasz Type Operators Involving Tangent Polynomials

Tanjant Polinomlarını İçeren Kantorovich-Szasz Tipli Operatörlerin Yakınsama Özellikleri Üzerine
Erkan AĞYÜZ 244-252

Fixed Points for Generalized Type Contractions in Partially Ordered Metric Spaces

Kısmen Sıralı Metrik Uzaylarda Genelleştirilmiş Tip Büzülmeler için Sabit Noktalar
Seher Sultan YEŞİLKAYA 253-262

A New Approach to Fibonacci Tessarines with Fibonacci and Lucas Number Components

Fibonacci ve Lucas Sayı Bileşenleri ile Fibonacci Tessarinelere Yeni Bir Yaklaşım
Faik BABADAĞ, Merve USLU 263-275

The Modified Exp (-vartheta(sigma)) -Expansion Function Method For Exact Solutions of The Simplified MCH Equation and The Getmanou Equation

Modifiye Basitleştirilmiş MCH Denklemi ve Getmanou Denklemi Tam Çözümleri için Modifiye Edilmiş Exp (- $\vartheta(\sigma)$)-Açılım Fonksiyon Metodu
Seyma TULUCE DEMIRAY, Sevgi KASTAL 276-289

Solitary Wave Solutions of the (3+1)-dimensional Khokhlov–Zabolotskaya–Kuznetsov Equation by using the (G'/G,1/G)-Expansion Method	
(3+1)-Boyutlu Khokhlov–Zabolotskaya–Kuznetsov Denklemi (G'/G , $1/G$)-Açılım Metodu Yardımıyla Solitary Dalga Çözümleri	290-301
<i>Serbay DURAN, Hülya DURUR, Asif YOKUŞ</i>	
Linear Convex Combination Estimators and Comparisons	
Lineer Konveks Kombinasyon Tahmin Ediciler ve Karşılaştırmalar	302-320
<i>Selahattin KAÇIRANLAR, Issam DAWOUD, Dünya KARAPINAR</i>	
Estimation of Risk Measures for Transmuted Weibull Distribution	
Dönüştürülmüş Weibull Dağılımı için Risk Ölçülerinin Tahmini	362-369
<i>Caner TANIŞ</i>	
Optimization Based Undersampling for Imbalanced Classe	
Dengesiz Sınıflamada Optimizasyona Dayalı Azörnekleme	385-409
<i>Fatih SAĞLAM</i>	
Prediction of Human Development Index with Health Indicators Using Tree-Based Regression Models	
Ağaç Tabanlı Regresyon Modelleri Kullanılarak Sağlık Göstergeleri ile İnsani Gelişme Endeksinin Tahmini	410-420
<i>Tuba KOC, Pelin AKIN</i>	
Invariant and Lacunary Invariant Statistical Equivalence of Order β for Double Set Sequences	
Çift Küme Dizileri için β ıncı Meriteden İnvaryant ve Lacunary İnvaryant İstatistiksel Denklik	421-430
<i>Uğur ULUSU, Erdiñ DÜNDAR, Fatih NURAY</i>	
<hr/> <hr/>	
PHYSICS	
An Educational Higgs Study with CMS Open Data	
CMS Açık Verisi ile Eğitici Higgs Çalışması	431-443
<i>Bora AKGUN</i>	
Exchange Bias Effect in NiMnSbB Ferromagnetic Shape Memory Alloys Depending on Mn Content	
Bor Eklenmiş NiMnSb Ferromanyetik Şekil Hatırlamalı Alaşımlarda Exchange Bias Etkisi	444-455
<i>Gökhan KIRAT</i>	
A Comparative Study of DFT/B3LYP/6-31 G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++ G(d,p) and HSEH1PBE/6-31G(d,p) Methods Applied to Molecular Geometry and Electronic Properties of C_s-C₆₀ Cl₆ Molecule	
C _s -C ₆₀ Cl ₆ Molekülünün Moleküler Geometri ve Elektronik Özelliklerine Uygulanan Metot Ve Baz Seti, DFT/B3LYP/6-31 G(d,p), RM062X/6-31G (d,p), B3LYP/6-311++ G(d,p) ve HSEH1PBE/6-31G(d,p), Yöntemlerinin Karşılaştırmalı İncelemesi	456-473
<i>Ebru KARAKAŞ SARIKAYA, Ömer DERELİ, Semiha BAHÇELİ</i>	
Density Functional Theory Studies of Some Barbiturates on Lipophilicity	
Bazı Barbitüratların Lipofilikliği Üzerine Yoğunluk Fonksiyonel Teori Çalışmaları	487-503
<i>Sümeyya SERİN, Ali BAYRİ</i>	
Evaluation of Radiation Attenuation Properties of Some Cancer Drug	
Bazı Kanser İlaçlarının Radyasyon Soğurma Özelliklerinin Değerlendirilmesi	504-522
<i>İlyas ÇAĞLAR, Gülçin BİLGİCİ CENGİZ</i>	
<hr/>	



On Surfaces in Pseudo-Galilean Space with Prescribed Mean Curvature

Muhittin Evren AYDIN¹, Alper Osman ÖĞRENMİŞ^{2,*}

¹Firat University, Faculty of Science, Department of Mathematics, 23100, Elazığ, Türkiye
meaydin@firat.edu.tr, ORCID: 0000-0001-9337-8165

²Firat University, Faculty of Science, Department of Mathematics, 23100, Elazığ, Türkiye
aogrenmis@firat.edu.tr, ORCID: 0000-0001-5008-2655

Received: 01.04.2021

Accepted: 04.08.2021

Published: 31.12.2021

Abstract

In this study, we consider certain classes of surfaces in the pseudo-Galilean space, the translation and factorable surfaces. We obtain these surfaces that satisfy the equation $H = v^\perp$, where H is the mean curvature and v^\perp is the normal component of an isotropic vector v .

Keywords: Translating soliton; Manifolds with density; Mean curvature; Pseudo-Galilean space.

Yarı-Galileo Uzayında Belli Bir Ortalama Eğriliğe Sahip Yüzeyle Üzerine

Öz

Bu çalışmada; yarı-Galileo uzayında, öteleme ve ayrışabilir yüzeyler denilen iki belirgin sınıf ele alınmıştır. v^\perp , bir v izotropik vektörün normal bileşeni olmak üzere bu yüzeylerden ortalama eğriliği $H = v^\perp$ denklemini sağlayanlar elde edilmiştir.

Anahtar Kelimeler: Ötelenen soliton; Yoğunluklu manifoldlar; Ortalama eğrilik; Yarı-Galileo uzayı.



1. Introduction

We are interested in the pseudo-Galilean geometry which is one of the real Cayley-Klein geometries. Let G_3^1 denote the pseudo-Galilean 3-space, $S \subset G_3^1$ an admissible surface, H and N are the mean curvature and unit normal vector field on S , respectively. Moreover, let $(L^2, \langle \cdot, \cdot \rangle_L)$ denote the Lorentzian 2-space. We consider the following:

$$H = v^\perp, \quad (1)$$

where v^\perp is the normal component of a unit isotropic vector $v \in G_3^1$. Note that $v^\perp = \langle N, v \rangle_L$ is the Lorentzian angle function of S between N and v . Up to the absolute figure of G_3^1 , since N is completely isotropic and orthogonal to all non-isotropic vectors, some minimal surface obeys to Eqn. (1) if v is non-isotropic. This is the justification why we take v as isotropic in Eqn. (1).

The importance of Eqn. (1) is due to the theories of manifolds with density and mean curvature flow. A surface whose mean curvature holds Eqn. (1) is called *translating soliton* of the mean curvature flow [1-5]. In the Euclidean setting, besides straight lines, one-dimensional solution to Eqn. (1) is the curve $s \mapsto -\log \cos s$, which is called *grim reaper* and known for moving upwards with constant speed under the flow, see [6, 7]. The hyperbolic versions of those functions are the so-called *Lorentzian grim reapers*, $s \mapsto \log \sinh s$ and $s \mapsto \log \cosh s$ [8]. In the Galilean setting, the situation is different. More explicitly, let κ be the curvature of a smooth curve γ in the Galilean plane G_2 and $\langle \cdot, \cdot \rangle_G$ the Galilean scalar product in G_2 . Then, Eqn. (1) writes $\kappa = \langle (0,1), v \rangle_G$, admitting solutions as straight lines ($\kappa = 0$) and parabolic circles ($\kappa = 1$).

Let (x, y, z) denote the affine coordinates in G_3^1 and $\varphi(x, y, z) = kx + py + qz, k, p, q \in \mathbb{R}$. From theory of manifolds with density, a surface satisfying Eqn. (1) is indeed a *minimal surface with density* e^φ [3, 9, 10]. Meanwhile, since $v = \text{grad } \varphi = ke_1 + pe_2 + qe_3$ is isotropic for standard basis vectors $\{e_1, e_2, e_3\}$ of G_3^1 , k must vanish in our case. More generally, a φ -mean curvature (or *weighted mean curvature*) H_φ with density e^φ is given by $2H_\varphi = 2H - (d\varphi/dN)$.

One of the basic classes of surfaces in differential geometry is the translation surfaces generated by translating two curves up to isometry of ambient space. Let $S \subset G_3^1$ be a translation surface and $s \mapsto \alpha(s)$ and $t \mapsto \beta(t)$ two parametric curves, $s \in I \subset \mathbb{R}$, $t \in J \subset \mathbb{R}$. Then, S is locally given by

$$x(s, t) = \alpha(s) + \beta(t), \quad (2)$$

where $\alpha := \alpha(s)$ and $\beta := \beta(t)$ are called *generating curves*. The other class of the surfaces in which we are interested is the one associated with the product of two single variable functions,

namely the *factorable* (or *homothetical*) surfaces. Up to the absolute figure a factorable surface is given by one of the explicit forms

$$x = f(s)g(t) \text{ and } z = f(s)g(t), \quad (3)$$

for smooth functions $f(s)$ and $g(t)$. Those surfaces in Galilean and pseudo-Galilean geometries have been considered in several research articles from different geometrical point of views. For example, the results on these surfaces in terms of Gaussian and mean curvatures can be found in [11-19], while the ones in terms of the Laplacian associated with the fundamental forms are in [20-23]. Some surfaces satisfying Eqn. (1) in G_3^1 were already considered from the manifolds with density point of view, [24-27].

In some sense, solving Eqn. (1) is a problem of finding prescribed mean curvature surfaces, which is our main interest. In this paper, we firstly study translation surfaces Eqn. (2) in G_3^1 , whose mean curvature satisfies Eqn. (1). When both α and β are planar, the problem was already solved in [26] and for this reason, we deal with the only case that one of α or β is planar and the other spatial. Under this condition, we solve Eqn. (1) completely. In Section 4, we also classify the surfaces given by Eqn. (3) which satisfy Eqn. (1).

2. Preliminaries

In this section, we recall some basics on the curves and surfaces in the pseudo-Galilean geometry from [18, 28-33]. We also refer to [34, 35] for the Lorentzian arguments.

Let $P_3(R)$ denote a real projective 3-space and $(u_0 : u_1 : u_2 : u_3)$ the homogeneous coordinates. The *pseudo-Galilean 3-space* G_3^1 is a Cayley-Klein space $P_3(R)$ with the *absolute figure* $\{\omega, f, l\}$, where ω is the *absolute plane* $u_0 = 0$, f the *absolute line* $u_0 = u_1 = 0$ and l the *fixed hyperbolic involution* of points of f . The hyperbolic involution is $(0 : 0 : u_2 : u_3) \mapsto (0 : 0 : u_3 : u_2)$ and then $u_2^2 - u_3^2 = 0$ is the *absolute conic*.

Let us introduce the affine coordinates $(u_0 : u_1 : u_2 : u_3) = (1 : x : y : z)$. Up to the absolute figure, the *pseudo-Galilean distance* between the points $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$ is

$$d(p, q) = \begin{cases} |q_1 - p_1|, & \text{if } p_1 \neq q_1 \\ \sqrt{|(q_2 - p_2)^2 - (q_3 - p_3)^2|}, & \text{if } p_1 = q_1. \end{cases}$$

The six-parameter group of motions of G_3^1 leaves invariant the absolute figure and the pseudo-Galilean distance, given in terms of affine coordinates as follows:

$$\begin{aligned} \bar{x} &= a_1 + x \\ \bar{y} &= a_2 + a_3x + y \cosh \phi + z \sinh \phi \\ \bar{z} &= a_4 + a_5x + y \sinh \phi + z \cosh \phi, \end{aligned}$$

where a_1, \dots, a_5, ϕ are some constants.

There are two sorts of lines and planes in G_3^1 . We call a line *isotropic* when its intersection with the absolute line f is non-empty and *non-isotropic* otherwise. A plane is said to be *isotropic* if it does not involve f , otherwise it is said to be *non-isotropic*. The non-isotropic planes are so-called *Lorentzian* since its induced geometry is Lorentzian. In the affine model of G_3^1 , the Lorentzian planes are in the form $x = \text{const}$.

A vector $v = (v_1, v_2, v_3)$ is said to be *isotropic (non-isotropic)* if $v_1 = 0 (\neq 0)$. Let $w = (w_1, w_2, w_3)$ and $\langle \cdot, \cdot \rangle_G$ denote the *pseudo-Galilean dot product*. Then, $\langle v, w \rangle_G$ is the Lorentzian scalar product if both v and w are isotropic. Otherwise, $v_1^2 + w_1^2 \neq 0$, it is defined by $\langle v, w \rangle_G = v_1 w_1$. The *pseudo-Galilean angle* between v and w is defined as the Lorentzian angle if v and w are isotropic. Otherwise, it is given by the pseudo-Galilean distance. We call v and w *orthogonal* if $\langle v, w \rangle_G = 0$.

An isotropic vector v is called *spacelike* if $\langle v, v \rangle_L > 0$; *timelike* if $\langle v, v \rangle_L < 0$ and *lightlike* if $\langle v, v \rangle_L = 0$. We call the spacelike and timelike vectors *non-degenerate*. Let $\{e_1, e_2, e_3\}$ be standard basis vectors and v and w no both isotropic vectors. Then, the *pseudo-Galilean cross-product* is

$$v \times_G w = \begin{vmatrix} 0 & -e_2 & e_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

Therefore we have $\langle v \times_G w, z \rangle_G = -\det(v, w, \tilde{z})$, where \tilde{z} is the projection of z onto the yz -plane. Note that the vector $v \times_G w$ is orthogonal to both v and w .

Let C be a curve given in the parametric form

$$s \mapsto r(s) = (x(s), y(s), z(s)), s \in I \subset R.$$

The curve C is said to be *admissible* if the following conditions hold: for each $s \in I$,

- 1) $r' = \frac{dr}{ds}$ is non-isotropic;
- 2) no where C has no inflection points, i.e. r' and $r'' = \frac{d^2r}{ds^2}$ are linearly independent;
- 3) \tilde{r}' and \tilde{r}'' are non-degenerate.

Then an admissible curve C is said to be *parameterized by arc-length* if the function x is the identity, up to a translation of G_3^1 . Let C be such a curve. Then we call $t = r'$ *unit tangent* to C and $\kappa = \sqrt{\langle r'', r'' \rangle_L}$ *curvature* of C . The *normal* and *binormal* to C are defined by

$$n = \frac{1}{\kappa(s)}(0, y'', z'') \text{ and } b = \frac{1}{\kappa(s)}(0, z'', y'').$$

The *torsion* of C is introduced by

$$\tau = \frac{\det(r', r'', r''')}{\kappa^2}.$$

We call the admissible curve C *spatial* provided $\tau \neq 0$ for each $s \in I$. We call an admissible curve *isotropic planar* if it fully lies in an isotropic plane and in such case τ vanishes identically. We also call a curve *Lorentzian planar* if it fully lies in a Lorentzian plane. For a Lorentzian planar curve the Frenet apparatus are well known.

Let S be a surface in G_3^1 locally given by a regular map

$$(u_1, u_2) \mapsto x(u_1, u_2) = (x(u_1, u_2), y(u_1, u_2), z(u_1, u_2)), (u_1, u_2) \in D \subset R^2.$$

Let $x_{,i} = \frac{\partial x}{\partial u_i}$ and $x_{,ij} = \frac{\partial^2 x}{\partial u_i \partial u_j}$ and etc., $1 \leq i, j \leq 2$. Then, S is said to be *admissible* if $x_{,i} \neq 0$, for some $i = 1, 2$. For such an admissible surface S , the *first fundamental form* is

$$\langle dx, dx \rangle_G = Edu_1^2 + 2Fdu_1du_2 + Gdu_2^2,$$

where $E = (x_{,1})^2$, $F = x_{,1}x_{,2}$, $G = (x_{,2})^2$. Since nowhere an admissible surface has Lorentzian tangent plane, up to the absolute figure, the isotropic vector $x_{,1} \times_G x_{,2}$ is *normal* to S . Let

$$W = \langle x_{,1} \times_G x_{,2}, x_{,1} \times_G x_{,2} \rangle_L.$$

Then the surface S is called *spacelike* if $W < 0$; *timelike* if $W > 0$; and *lightlike* if $W = 0$. The spacelike and timelike surfaces are so-called *non-degenerate* and then, the *unit normal vector* to the non-degenerate surface S is

$$N = \frac{x_{,1} \times_G x_{,2}}{\sqrt{|W|}}.$$

Let $\varepsilon = \langle N, N \rangle_L = \pm 1$ and

$$L_{ij} = \varepsilon \frac{1}{x_{,1}} \langle x_{,1} \tilde{x}_{,ij} - (x_{,i})_{,j} \tilde{x}_{,1}, N \rangle_L = \varepsilon \frac{1}{x_{,2}} \langle x_{,2} \tilde{x}_{,ij} - (x_{,i})_{,j} \tilde{x}_{,2}, N \rangle_L,$$

where one of $x_{,1}$ and $x_{,2}$ is always nonzero due to the admissibility. Then the *second fundamental form* of S is

$$II = Ldu_1^2 + 2Mdu_1du_2 + Ndu_2^2,$$

where $L = L_{11}$, $M = L_{12}$, $N = L_{22}$. Thereby, the Gaussian and mean curvatures are defined as

$$K = -\varepsilon \frac{LN-M^2}{|W|} \text{ and } H = -\varepsilon \frac{GL-2FM+EN}{2|W|}.$$

We call a surface *minimal* if H vanishes identically. Throughout this study, we deal with only non-degenerate admissible surfaces.

3. Translation Surfaces

Let $S \subset G_3^1$ a translation surface whose one generating curve lies in a Lorentzian plane and the other admissible. Then, it locally parameterizes

$$x(s, t) = \alpha(s) + \beta(t),$$

in which we may assume $\beta(v)$ fully lies in the Lorentzian yz -plane. Then the unit normal vector field and mean curvature are

$$N = n_\beta \text{ and } H = \frac{1}{2}k_\beta,$$

where n_β and k_β the principal normal and Frenet curvature of β . Then, for the surface S , Eqn. (1) is now

$$k_\beta = 2\varepsilon \langle n_\beta, v \rangle_L,$$

which means that β is one dimensional solution in L^2 to translating soliton Eqn. (1). As can be seen the generating curve α does not play a role. Therefore, we may state that

Proposition 1. Let S be a translation surface in G_3^1 given by $x(s, t) = \alpha(s) + \beta(t)$, where α is some admissible curve and β is Lorentzian planar. Then, S holds Eqn. (1) if and only if β is one dimensional solution in L^2 to Eqn. (1).

We next consider the translation surface S whose one generating curve is isotropic planar, say β , and the other spatial. Let

$$s \mapsto \alpha(s) = (s, f(s), h(s)) \text{ and } t \mapsto \beta(t) = (t, 0, g(t)),$$

where $(s, t) \in I \times J \subset \mathbb{R}^2$. Then, S is locally given by

$$x(s, t) = (s + t, f(s), h(s) + g(t)). \tag{4}$$

Denote a prime the derivative with respect to the related variable. Since α is assumed to be spatial, the following holds

$$f''h''' - f'''h'' \neq 0, \tag{5}$$

implying that f'' and g'' must be linearly independent. The unit normal vector field and mean curvature are

$$N = \frac{1}{\sqrt{|(g'-h')^2 - f'^2|}} (0, g' - h', -f')$$

and

$$H = -\frac{f''(g'-h') + f'(h''+g'')}{2|(g'-h')^2 - f'^2|^{3/2}}.$$

Let $v = (0, p, q)$. Then, Eqn. (1) is now

$$f''(g' - h') + f'(h'' + g'') = 2\varepsilon(p(h' - g') - qf') (f'^2 - (g' - h')^2). \tag{6}$$

The successive derivatives of Eqn. (6) with respect to s and t yield

$$f'''g'' + f''g''' = 4\varepsilon g'' \left(-pf'f'' - 3(g' - h')h'' + q(f''g' - (f'h')') \right). \tag{7}$$

Assume that $g'' \neq 0$ in Eqn. (7), for each $t \in J$. Dividing Eqn. (7) with g'' and then taking derivative with respect to t gives

$$f'' \left(\frac{g'''}{g''} \right)' = 4\varepsilon g'' (-3h'' + qf''), \tag{8}$$

in which the right-hand side of Eqn. (8) is non-vanishing due to Eqn. (5). Then, there exists a nonzero constant c such that

$$\left(\frac{g'''}{g''} \right)' = 4\varepsilon c g''. \tag{9}$$

Hence, Eqn. (8) reduces to $(q - c)f'' = 3h''$, which contradicts with Eqn. (5). This discussion allows our assumption to be false, namely there exists $t_0 \in J$ such that $g'' = 0$ in a neighborhood of t_0 in J . In such a case the generating curve $\beta(t)$ is a non-isotropic straight line parallel to $(1, 0, g_0)$, where $g' = g_0 \in R$, namely the surface given by Eqn. (4) is a cylinder with non-isotropic rulings.

Therefore, we have proved

Theorem 1. The cylinders with non-isotropic rulings are the only translating solitons of translation type whose one generating curve is isotropic planar and the other is spatial.

There is a class of translation solitons of translation type that we do not consider: both α and β are spatial, which remains an open problem.

4. Factorable Surfaces

Let $s \mapsto f(s)$ and $t \mapsto g(t)$, $s \in I, t \in J$, be smooth functions and $S \subset G_3^1$ locally the graph of the product of $f := f(s)$ and $g := g(t)$. Assume that f and g are non-vanishing on $I \times J$. Up to the absolute figure, the geometric properties of S depend on if it is the graph on an isotropic or Lorentzian planes. Thereby, we consider the surfaces $z = f(s)g(t)$ (or equivalently $y = f(s)g(t)$) and $x = f(s)g(t)$, separately.

Let S be locally the surface $z = f(s)g(t)$ which is parameterized by

$$(s, t) \mapsto x(s, t) = (s, t, f(s)g(t)).$$

The unit normal and mean curvature are given by

$$N = \frac{1}{\sqrt{|(fg)'^2 - 1|}} (0, fg', 1)$$

and

$$H = \frac{fg''}{2|(fg)'^2 - 1|^{3/2}}.$$

Let $v = (0, p, q)$ and then Eqn. (1) is

$$fg'' = 2\varepsilon(pfg' - q)((fg)'^2 - 1). \tag{10}$$

Assume that f is a non-constant function. Then, (10) turns to a polynomial of degree 2 in f

$$-2\varepsilon q + (2\varepsilon pg' + g'')f + 2\varepsilon g'^2 f^2 - 2\varepsilon pg'^2 f^3 = 0,$$

in which the coefficients must vanish, giving $q = 0$ and $g' = 0$. We then deduce that $v = (0,1,0)$ and

$$x(s, t) = (s, 0, g_0 f(s)) + t(0,1,0), \tag{11}$$

where $g' = g_0 \in R - \{0\}$. Eqn. (11) is a parameterization of a cylinder with isotropic rulings. If $f = f_0 \in R - \{0\}$, then

$$x(s, t) = s(1,0,0) + (0, t, f_0 g(t)),$$

which is a parameterization of a cylinder with non-isotropic rulings.

Therefore, we have proved

Theorem 2. The cylinders are the only translating solitons of the form $z = f(s)g(t)$.

We next take the surface $x = f(s)g(t)$, parameterized by

$$(s, t) \mapsto x(s, t) = (f(s)g(t), s, t). \tag{12}$$

The unit normal and mean curvature are given by

$$N = \frac{1}{\sqrt{|(f'g)^2 - (fg')^2|}} (0, f'g, -fg')$$

and

$$H = \frac{(f'g)^2 fg'' - 2fg(f'g')^2 + (fg')^2 f''g}{2|(f'g)^2 - (fg')^2|^{3/2}}.$$

Then Eqn. (1) with $v = (0, p, q)$ writes

$$fg\{f'^2(gg'' - g'^2) + g'^2(ff'' - f'^2)\} = 2\varepsilon(pf'g + qfg')((f'g)^2 - (fg')^2). \tag{13}$$

The functions f and g play symmetric roles in Eqn. (13) and we only concentrate for f .

Case (a). $f = f_0 \in R - \{0\}$. Then Eqn. (13) implies $v = (0, 1, 0)$ and

$$x(s, t) = s(0, 1, 0) + (f_0g(t), 0, t),$$

which is a cylinder with isotropic rulings.

Case (b). f and g are both non-constant functions. We divide Eqn. (13) with $fg(f'g')^2$ and write

$$\left(\frac{f'}{f}\right)' + \left(\frac{g'}{g}\right)' = 2\varepsilon\left(p\frac{f'}{f} + q\frac{g'}{g}\right)\left(\left(\frac{g'}{g}\right)^2 - \left(\frac{f'}{f}\right)^2\right). \tag{14}$$

After successive derivatives of Eqn. (14) with respect to s and t we may deduce

$$p\left(\frac{f'}{f}\right)' \left(\frac{g'}{g}\right) \left(\frac{g'}{g}\right)' - q\left(\frac{g'}{g}\right)' \left(\frac{f'}{f}\right) \left(\frac{f'}{f}\right)' = 0. \tag{15}$$

Assume that f and f' are linearly independent and, by symmetry, so are g and g' . Then Eqn. (15) reduces to

$$p\left(\frac{f'}{f}\right)^3 - q\left(\frac{g'}{g}\right)^3 = 0,$$

which implies $p = q = 0$. This is a contradiction. Therefore, f and f' must be linearly dependent and put $f' = b_1f$, namely $f(s) = b_2e^{b_1s}$ for nonzero constants b_1 and b_2 . Therefore, the surface becomes $x = b_2e^{b_1s}g(t)$ or, up to a translation, Eqn. (12) turns to

$$(x, t) \mapsto \left(x, \frac{1}{b_1} \log \left|\frac{x}{g(t)}\right|, t\right), \tag{16}$$

which is a translation surface. Its generating curves are

$$x \mapsto \alpha(x) := \left(x, \frac{1}{b_1} \log|x|, 0\right), \text{ and } t \mapsto \beta(t) := \left(0, -\frac{1}{b_1} \log|g(t)|, t\right).$$

Then, a surface $x = f(s)g(t)$ satisfying Eqn.(1) has to be of form Eqn. (16) if it is non-cylindrical. Since $\beta(t)$ is fully in the Lorentzian yz –plane, it has to be one dimensional solution to Eqn. (1) in L^2 due to Proposition 1.

To sum up, we have proved

Theorem 3. A surface $x = f(s)g(t)$ satisfying Eqn. (1) is either a cylinder with isotropic rulings or a translation surface of the form Eqn. (16), where one generating curve is one dimensional solution to Eqn. (1) in L^2 .

Example 1. Let $a^2 - b^2 = \pm 1$, $a, b \in \mathbb{R} - \{0\}$. Consider the surface $x = e^{as+bt}$ and take $v = (0, b, -a)$. This surface is indeed minimal and its normal is $N = (0, a, -b)$ such that Eqn. (1) holds obviously. Notice that it is also given by

$$(x, t) \mapsto \left(x, \frac{1}{a} \log|x| - \frac{b}{a}t, t\right).$$

Acknowledgement

The authors are thankful to the referees for their valuable comments and suggestions helping improvement of the paper.

References

- [1] Bueno, A., *Translating solitons of the mean curvature flow in the space $H^2 \times \mathbb{R}$* , Journal of Geometry, 109, Article number 42, 2018. <https://doi.org/10.1007/s00022-018-0447-x>, 2018.
- [2] López, R., *Separation of variables in equations of mean-curvature type*, Proceedings of the Royal Society of Edinburgh Section A, 146(5), 1017-1035, 2016.
- [3] López, R., *Some geometric properties of translating solitons in Euclidean space*, Journal of Geometry, 109, Article number 40, 2018. <https://doi.org/10.1007/s00022-018-0444-0>
- [4] Aydin, M.E., López, R., *Ruled translating solitons in Minkowski 3-space*, arXiv:2109.05254v1, 2021.
- [5] Aydin, M.E., López, R., *Translating solitons by separation of variables*, arXiv:2109.05253v1, 2021.
- [6] Angenent, S., *On the formation of singularities in the curve shortening flow*, Journal of Differential Geometry, 33, 601-633, 1991.
- [7] Halldorsson, H., *Self-similar solutions to the curve shortening flow*, Transactions of the American Mathematical Society, 364(10), 5285-5309, 2012.
- [8] Castro, I., Castro-Infantes, I., Castro-Infantes, J., *Curves in the Lorentz-Minkowski plane: elasticae, catenaries and grim-reapers*, Open Mathematics, 16,747-766, 2018.
- [9] Altin, M., Kazan, A., Karadag, H.B., *Ruled surfaces in E^3 with density*, Honam

Mathematical Journal, 41(4), 683-695, 2019.

[10] Hieu, D.T., Hoang, N.M., *Ruled minimal surfaces in R^3 with density e^z* , Pacific Journal of Mathematics, 243(2), 277-285, 2009.

[11] Abdel-Aziz, H.S., Saad, M.K., Ali, H.A., *Affine factorable surfaces in pseudo-Galilean space*, arXiv:1812.00765v1 [math.GM].

[12] Aydın, M.E., Öğrenmiş, A.O., Ergut, M., *Classification of factorable surfaces in the pseudo-Galilean space*, Glasnik Matematički Series III, 50(70), 441-451, 2015.

[13] Aydın, M.E., Kulahcı, M.A., Öğrenmiş, A.O., *Non-zero constant curvature factorable surfaces in pseudo-Galilean space*, Communications of the Korean Mathematical Society, 33(1), 247-259, 2018.

[14] Aydın, M.E., Kulahcı, M.A., Öğrenmiş, A.O., *Constant curvature translation surfaces in Galilean 3-space*, International Electronic Journal of Geometry, 12(1), 9-19, 2019.

[15] Dede, M., Ekici, C., *On minimal surfaces in Galilean space*, Conference Proceedings of Science and Technology, 2(2), 142-147, 2019.

[16] Kelleci, A., *Translation-factorable surfaces with vanishing curvatures in Galilean 3-spaces*, International Journal of Maps in Mathematics, 4(1), 14-26, 2021.

[17] Milin-Sipus, Z., Divjak, B., *Translation surface in the Galilean space*, Glasnik Matematički Series III, 46(66), 455-469, 2011.

[18] Milin-Sipus, Z., *On a certain class of translation surfaces in a pseudo-Galilean space*, International Mathematical Forum, 6(23), 1113-1125, 2012.

[19] Yoon, D.W., *Some classification of translation surfaces in Galilean 3-space*, International Journal of Mathematical Analysis, 28(6), 1355-1361, 2012.

[20] Abdel-Baky, R.A., Unluturk, Y., *A study on classification of translation surfaces in pseudo-Galilean 3-space*, Journal of Coupled Systems and Multiscale Dynamics, (6)3, 233-240, 2018.

[21] Bansal, P., Shahid, M.H., *On classification of factorable surfaces in Galilean space G^3* , Jordan Journal of Mathematics and Statistics, 12(3), 289-306, 2019.

[22] Cakmak, A., Karacan, M.K., Kiziltug, S., Yoon, D.W., *Corrigendum to translation surfaces in the 3-dimensional Galilean space satisfying $\Delta^I x_i = \lambda_i x_i$* , Bulletin of the Korean Mathematical Society, 56(2), 549-554, 2019.

[23] Lone, M.S., *Homothetical surfaces in three dimensional pseudo-Galilean spaces satisfying $\Delta^I x_i = \lambda_i x_i$* , Advances in Applied Clifford Algebras, 29(5), 92, 2019. <https://doi.org/10.1007/s00006-019-1007-7>

[24] Kazan, A., Karadag, H.B., *Weighted minimal and weighted flat surfaces of revolution in Galilean 3-space with density*, International Journal of Analysis and Applications, 16(3), 414-426, 2018.

[25] Mosa, S., Elzawy, M., *Helicoidal surfaces in Galilean space with density*, Frontiers in Physics, 8, Article 81, 2020. <https://doi.org/10.3389/fphy.2020.00081>.

[26] Yoon, D.W., *Weighted minimal translation surfaces in the Galilean space with density*, Open Mathematics, 15, 459-466, 2017.

[27] Yoon, D.W., Lee, J.W., Lee, C.W., *φ -minimal rotational surfaces in pseudo-Galilean space with density*, Annales Polonici Mathematici, 120, 183-196, 2017.

[28] Altın, M., Unal, I., *Surface family with common line of curvature in 3-dimensional*

Galilean space, Facta Universitatis: Series Mathematics and Informatics, 35(5), 1315-1325, 2020.

[29] Divjak, B., *Curves in pseudo-Galilean geometry*, Annales Universitatis Scientiarum Budapestinesis, 41,117-128,1998.

[30] Milin-Sipus, Z., Divjak, B., *Surfaces of constant curvature in the pseudo-Galilean space*, International Journal of Mathematics and Mathematical Sciences, Art ID375264, 28pp., 2012.

[31] Mólnar, E., *The projective interpretation of the eight 3-dimensional homogeneous geometries*, Beiträge zur Algebra und Geometrie, 38(2), 261-288, 1997.

[32] Onishchick, A., Sulanke, R., *Projective and Cayley-Klein Geometries*, Springer, 2006.

[33] Röschel, O., *Die Geometrie des Galileischen Raumes*, Habilitationsschrift, Leoben, 1984.

[34] López, R., *Differential geometry of curves and surfaces in Lorentz-Minkowski space*, International Electronic Journal of Geometry, 7(1), 44-107, 2014.

[35] O'Neill, B., *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, 1983.



Physiological and Molecular Effects of Exogenous Gibberellin (GA3) Treatment on Germination of Barley Seeds under Salt Stress

Cüneyt UÇARLI^{1,*}

¹*Istanbul University, Faculty of Science, Department of Molecular Biology and Genetics, 34134,
Istanbul, Türkiye*

ucarlicu@istanbul.edu.tr, ORCID: 0000-0002-9526-576X

Received: 27.03.2020

Accepted: 18.08.2021

Published: 31.12.2021

Abstract

Salinity is considered as one of the most important factors restricting germination parameters including rate and percentage of seed germination in crops. This study focused on the alleviating role of exogenously applied 10 mg/L gibberellic acid (GA3) on barley (*Hordeum vulgare* L.) seeds during germination under salt stress (120 mM NaCl). Physiological and morphological changes, and differential gene expression at 3 days after imbibition (DAI) were determined and compared with or without gibberellic acid (GA3) under salt stress. Exogenous GA3 was found to increase the shoot and root length of germinated barley seeds under salt stress by 67 and 15%, respectively, compared to those treated with salinity alone. On the other hand, exogenous GA3 treatment significantly reduced ion leakage, osmolyte accumulation, and proline content under salinity. NaCl was found to decrease the expression of the *HvABI5*, *HvABA7* and *HvKOl* by 3, 10, and 33 fold, respectively, at 3 DAI, whereas addition of GA3 in root medium rescued the expression of these genes to control levels. Besides, exogenous GA3 significantly decreased mRNA level of *HvGA2ox4* under salinity during germination. This study may give insight into the relationship between salinity stress and the genes involved in GA3 metabolism and germination.



Keywords: Barley; Germination; GA3; Salinity; Proline; *HvKOl*; *HvGA2ox4*.

Eksojen Gibberellin (GA3) Uygulamasının Tuz Stresi Altındaki Arpa Tohumlarının Çimlenmesine Fizyolojik ve Moleküler Etkileri

Öz

Tuzluluk, bitkilerde tohum çimlenme oranı ve yüzdesi gibi çimlenme parametrelerini kısıtlayan en önemli faktörlerden biri olarak kabul edilmektedir. Bu çalışma, tuz stresi (120 mM NaCl) altında çimlendirilen arpa (*Hordeum vulgare* L.) tohumlarına eksojen olarak uygulanan 10 mg/L giberellik asidin (GA3) stresi hafifletici rolüne odaklanmıştır. İmbibisyondan 3 gün sonra tuz stresi altında giberellik asit (GA3) ile veya onsuz gerçekleşen fizyolojik ve morfolojik değişiklikler ve farklılaşmış gen anlatımı belirlendi ve karşılaştırıldı. Eksojen GA3'ün, tuz stresi altında filizlenmiş arpa tohumlarının sürgün ve kök uzunluğunu, yalnızca tuz ile muamele edilenlere kıyasla sırasıyla %67 ve %15 artırdığı bulundu. Öte yandan, eksojen GA3 uygulaması tuzluluk altında iyon sızıntısını, ozmolit birikimini ve prolin içeriğini anlamlı şekilde azaltmıştır. NaCl'nin imbibisyondan 3 gün sonra *HvABI5*, *HvABA7* ve *HvKOl* genlerinin anlatımını sırasıyla 3, 10 ve 33 kat azalttığı bulunurken, kök besiyerine GA3 eklenmesinin bu genlerin anlatım seviyelerini kontrol grubunun seviyesine çıkarmıştır. Ayrıca, eksojen GA3 tuz stresi altında çimlenen örneklerde *HvGA2ox4*'ün mRNA seviyesini anlamlı derecede azaltmıştır. Bu çalışma, tuzluluk stresi ile GA3 metabolizması ve çimlenme ile ilgili genler arasındaki ilişki hakkında fikir verebilir.

Anahtar Kelimeler: Arpa; Çimlenme; GA3; Tuzluluk; Prolin; *HvKOl*; *HvGA2ox4*.

1. Introduction

Salinity is regarded as one of the most important stress factors lowering crop productivity by preventing or delay seed germination as well as seedling establishment in crops [1]. Salt stress exhibits adverse effects on plants by accumulation of Na⁺ and Cl⁻ causing an ionic imbalance and toxicity, and osmotically reducing water uptake in arid and semi-arid regions [2]. Salt stress causes detrimental physiological and biochemical changes including certain enzymatic or hormonal activities, protein mechanism, and gene expression profile in germinating seeds; therefore, led to delay or preventing seed germination [3, 4]. Seed germination and early seedling stages are the most sensitive to salinity during plant life-cycle [5]. In plants, salt-stress responsive genes can be classified into three groups, which are the genes involved in sensing and signalling of the stress, transcriptional regulators such as *DRF1*, and salt-stress related genes including *LOX*, *MT2*, and *BASI* [6, 7].

Seed germination is a critical stage in establishing of a plant, and controlled by a number of factors comprising plant hormones such as gibberellins (GAs) and abscisic acid (ABA). GAs and ABA are the main endogenous factors, which act antagonistically in the control of seed germination and dormancy. While GAs promote seed germination, ABA induces and maintains seed dormancy [8]. Seed germination is regulated by differential expression of plant genes coding specific proteins necessary for germination, and also with activity of plant hormones which is regulated by the expression of plant genes [9]. Differential expression of ABA metabolic genes (e.g. *NCED1*), and ABA signalling genes including *ABI3*, *ABI5*, and *ABA7* regulate the ABA level in plant seeds, and therefore play roles in the control of dormancy and germination [8].

Gibberellins (GAs) are diterpenoids acting as a plant hormone and have important roles in seed germination, plant growth, development processes, as well as the mitigation of abiotic stresses including salt stress [10, 11]. Among more than 130 GAs, only four of them, namely GA1, GA3, GA4, and GA7 are thought to function as bioactive hormones [12, 13]. GA3 has been reported to alleviate the adverse effects of salinity on seed germination and seedling growth in crops [14, 15]. The biosynthesis of gibberellin is well clarified in model plants and mainly involves seven major enzymes including ent-kaurene oxidase (KO), ent-copalyl diphosphate synthase (CPS), ent-kaurenoic acid oxidase (KAO), ent-kaurene synthase (KS), gibberellin 20-oxidase (GA20ox), gibberellin 3-oxidase (GA3ox), and gibberellin 2-oxidase (GA2ox) which controls the inactivation of bioactive GAs [16, 17]. DELLA proteins, known as SLENDER1 (SLN1) in barley, are repressors of GA signalling in plants; thus, they repress the growth and germination in plants [18, 19].

Barley (*Hordeum vulgare* L.) is an important field crop after maize (*Zea mays* L.), wheat (*Triticum aestivum* L.), and rice (*Oryza sativa* L.) in terms of cultivated area (47 million hectares) and production (147.4 million tonnes) [7]. Global production of barley is mainly used for animal feed (more than 70%), malting in alcoholic beverages, seed, and human food [20]. While barley is considered the most salt-tolerant species among cereals, tolerance level of barley against salt stress varies depending on genotype and stage of plant growth [21]. Barley has been reported to be more sensitive to salinity during germination and young seedling stages but exhibits an increased tolerance with age [22].

The objectives of this study were to evaluate the alleviating effects of exogenous GA3 on seed germination of barley under salt stress, and determine the morphological and physiological changes as well as to determine the transcription pattern of 3 groups of genes. Group1 genes including *HvABA7*, *HvNCED1*, *HvABI3*, *HvABI5*, and *HvSLN1* are functional in suppression of

seed germination. Group2 genes involve *HvKOl*, *HvCPSl*, *HvKAOl*, *HvKSl*, *HvGA20oxl*, *HvGA3oxl* and *HvGA2ox4*, which are functional in GA3 metabolism and induction of seed germination. Group3 contains salt-stress responsive genes (e.g., *HvMT2*, *HvBASl*, *HvDRF1*, and *HvLOXl*).

2. Materials and methods

2.1. Plant material and growth conditions

The seeds of barley (*Hordeum vulgare* L.), cv. Martı (a salt-tolerant barley cultivar) were kindly provided by Agricultural Research Institute (Edirne, Turkey). Trials in the range of 40-200 mM showed that an optimal salt concentration of 120 mM was the best for salinity stress during seed germination in barley. Besides, the preliminary studies have shown that 10 mg/L GA3 (Caisson labs-G001) dissolved in pure water was the best concentration among the tested GA3 concentrations (1, 10 and 100 mg/L) under salt stress during germination of barley seeds. Ten barley seeds were germinated in petri dishes containing filter papers soaked with 3 ml dH₂O (control), 10 mg/L (~28.9 μM) GA3, 120 mM NaCl and 120 mM NaCl +10 mg/L GA3 for 3 days in dark at 25 °C in a climate chamber (Angelantoni, Ekochl 700).

2.2. Determination of osmolality, ion leakage, and proline content

In root and shoot tissues, the osmolyte measurement and ion leakage analysis were conducted according to Uçarlı and Gürel [7]. To determine the osmolyte content, 10 root or leaf samples (25-30 mg) were kept at -20 °C overnight. The cell sap was obtained by centrifugation at 4 °C for 30 min at 15000×g. An amount of 15 μl sap was taken from each sample, and after 10-fold dilution, osmolality content was measured using the osmometer (Semi-Micro Osmometer K-7400, Germany)

For ion leakage analysis, five 2 cm-length leaves (~30 mg) and 40-45 mg roots were harvested and washed three times with distilled water (dH₂O), then kept in glass tubes containing 5 ml dH₂O in a controlled chamber at 25 °C in the dark. After 16 h, 50 μl of the liquid sample was taken from glass tubes including root or leaf samples and measured the conductivity (E1) using a conductivity meter (HORIBA Scientific, NJ, USA). The glass tubes were autoclaved at 121 °C for 15 min and then, conductivity measurement was repeated to record E2 values. Ion leakage was calculated using following formula. Ion leakage (%) = (E1/E2) × 100 [7].

Proline content was determined according to Carillo et al. [23]. 20-40 mg of shoot or root tissue were pulverized with liquid nitrogen in centrifuge tubes and then 200 μl dH₂O was added into the tubes. The samples were incubated for 20 minutes in boiling water and then centrifuged at 6000 rpm for 1 minute. Supernatant was transferred into a tube and diluted 20-fold with 70%

ethanol. 150 µl aliquot of extract was added to reaction mixture (300 µl) containing 1% (w/v) ninhydrin in acetic acid 60% (v/v), and ethanol 20% (v/v). The resulting mixture was boiled in water bath at 95° C for 20 minute. After cooling at room temperature, mixture was spun down quickly at 2500 rpm for 1 minute. 200 µl of the mixture was transferred to a microplate well and read at 520 nm. The proline content was calculated using the following formula.

Proline in µmol.g-1FW = (Absextract – blank) / slope * Volextract / Volaliquot *1/FW. Absextract is the absorbance determined with the extract, blank (expressed as absorbance) and slope (expressed as absorbance·nmol-1) are determined by linear regression, Volextract is the total volume of the extract, Volaliquot is the volume used in the assay, FW (expressed in mg) is the amount of plant material extracted. It is assumed that Absextract is within the linear range.

2.3. Sampling, total RNA isolation, and qPCR

Shoot and root samples from were harvested 3 days after imbibition. Total RNA was extracted with the TRIzol® (Invitrogen, 15596-026) according to manufacturer's instructions using frozen 100 mg roots and shoots. cDNA synthesis, using SuperScript™ First-Strand Synthesis System (Invitrogen, 11904-018), and qPCR analyses were performed using CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad) according to Uçarlı and Gürel [7] .

The transcription pattern of 3 groups of genes, functional in suppression of seed germination (group1), functional in GA3 metabolism and induction of seed germination (group2), and salt-responsive genes (group3), was investigated at 3 DAI in germinating barley seeds by qPCR (Table 1). The relative mRNA levels of the genes were normalized with respect to the internal control *HvGAPDH* and the relative expression level of each gene was analyzed by delta CT method [24]. All qPCRs were performed in three biological replicates and two technical replicates.

2.4. Statistical analysis

Statistical analyses were performed with SPSS21 (IBM). Tests for significant differences were conducted using one-way analysis of variance (ANOVA) with least significance difference (LSD) tests at the 0.05 level of confidence.

Table 1: Primer sequences used in qPCR analyses

Gene	Accession No	Group No	Primer Sequences	Amplicon Size (bp)
<i>HvABA7</i>	X69817.1	1	5' GAACAAGCTTGGCAGTGACA 3' 5' CTTAGACGCCGGATCTCTGT 3'	228
<i>HvNCED1</i>	AB239297.1	1	5' CTCATCCCTCCCATCTTCTC 3' 5' GCCGCTAACTGTTTCCTCTTC 3'	126
<i>HvABI3</i>	AK369681.1	1	5' TGGTGATACCGAAGCAGAAC 3' 5' CGCATCGTCCTCGAAGTT3'	101
<i>HvABI5</i>	AY150676	1	5' GGTGAAGAAACAGACCGAGATAC 3' 5' CTTGTAAGCACTGCCTCTT 3'	96
<i>HvSLN1</i>	AF460219.1	1	5' CTACGAGTCCTGCCCTACC 3' 5' CCCTGCTTGATGCCGAAGTC 3'	114
<i>HvGA2ox4</i>	AY551432	2	5' TCCTAGCCAGCCAGCAACT 3' 5' GGCATGGACAGGACACAGA 3'	205
<i>HvGA3ox1</i>	AY551430	2	5' CCTCATAGATCAGCGCCAATA 3' 5' GGATGGAGAGACAGCAATGT 3'	110
<i>HvGA20ox1</i>	AY551428	2	5' CCAACTCGAACGTTGTTGTTATT 3' 5' CATGCATCGCTACCACTCTAC 3'	84
<i>HvCPS1</i>	AY551435	2	5' CCACCATTTCCTCCTTGGATAC 3' 5' GGACCAAACAACCAATCCAATTG 3'	110
<i>HvKSI</i>	AY551436	2	5' CATGCAAGGAGCTGTTCTGGAAGA 3' 5' GGATCAAAGGTTCACTGCCGCTTC 3'	115
<i>HvKO1</i>	AY551434	2	5' GACCCACATTCTCCTCTGAAC 3' 5' GAATCCACATCCTCGCCTAAA 3'	92
<i>HvKAO1</i>	AY326277	2	5' CCACCATTTCCTCCTTGGATAC 3' 5' GGAGGAGAGTCTGGTGATCTT 3'	118
<i>HvDRF1</i>	AY223807	3	5' TCCTCTCGGTCAGATTTGCT 3' 5' ACAGTCACCGGGTCAACTTC 3'	227
<i>HvMT2</i>	BM816564	3	5' TCAGTCGAATCAACACATGGA 3' 5' CACGAGGACGGAATAAAGC 3'	266
<i>HvLOX1</i>	U83904.1	3	5' AGCAGTGAAAGCGAGGAGAG 3' 5' AAGTCGTTGAGGTCCAGCAC 3'	142
<i>HvBAS1</i>	Z34917.1	3	5' CGTCACCAAATCGATCTCAA 3' 5' TCCACACTACGGCCAATACC 3'	141
<i>HvGAPDH</i>	AK251456	-	5' TGTCATGCCATGACTGCAA 3' 5' CCAGTGCTGCTTGAATGATG 3'	105

3. Results and Discussion

3.1. Physiological and morphological effects of NaCl and GA3 treatments

Seed germination is initiated by water imbibition and comprised with many different metabolic, cellular, and physiological processes [25]. Besides, the germination process can be

affected positively or negatively due to the environmental conditions including water, nutrition, temperature, and soil properties. Soil salinity negatively affects the germination process by retarding germination, lowering germination rate, and even preventing germination depending on salt concentration in soil. Germination and early seedling growth are two critical stages for plant establishment. Germination is the most sensitive plant growth stage to salinity stress and plants get tolerance against salt by age [3].

The effects of exogenous GA3 (10 mg/L), salt (120 mM NaCl), and their interactions during seed germination and post-germination of barley seeds were examined in terms of morphological changes (length of root and shoot) (Fig. 1), physiological changes including ion leakage, accumulation of osmolytes, and proline content (Table 2). In the present study, it has been found that salt (120 mM NaCl) has negative impact on barley germination and early seedling growth including retarding in root and shoot growth by 35% and 68%, respectively, compared to control condition (dH₂O) at 3 DAI. On the other hand, exogenous application GA3 (10 mg/L \approx 28.9 μ M) has alleviated the detrimental effects of salt stress on root and plant development, which resulted in a significant increase in length of shoot and root by 67% (from 0.84 to 1.40 cm) and 15% (from 3.60 to 4.25 cm), respectively, compared to salt-stressed seeds at $P < 0.05$ level (Fig. 1). In another study, 10 μ M GA3 application enhanced the final germination percentage up to 70%, and 31% at 100 mM and 200 mM NaCl, respectively, and the germination velocity compared to salt stress in the oilseed halophyte *Crithmum maritimum* [26]. Application of low concentration of GA3 (5 μ M) were able to reverse the detrimental effects of salinity stress (100 mM NaCl) on 10-day-old mung bean seedlings with significant increase in shoot-root elongation as well as in biomass [27]. These results are in accordance with Abdel-Hamid and Mohamed [28] who reported that the exogenous applications of GA3 (100 μ M \approx 35mg/L) enhanced the germination of barley seeds under salt stress. Furthermore, 100 μ M GA3 significantly increased the length of root and shoot under moderate salinity (100 mM NaCl) in barley cv. Giza 126, but had no effect on growth of root and shoot under higher salt stress (300 mM NaCl). Similarly, Maggio et al. [29] have reported that exogenous GA3 (100 mg/L) treatment decreased stomatal resistance and increased plant water use at low salinity (6.8 dS m⁻¹) after 141 days of treatment in tomato, while did not mitigate the salinity effects in moderate (11.7 dS m⁻¹) and high (16.7 dS m⁻¹) salt concentrations. Salinity stress causes the delay in germination and growth by reducing water potential, which resulting slower rate of imbibition. The increased growth of root and shoot with exogenous GA3 under salt stress may be attributed to the positive influence of gibberellic acid under salinity, which encourages cell division and cell elongation, and enhances efficiency of plant water use depending on salt concentration.

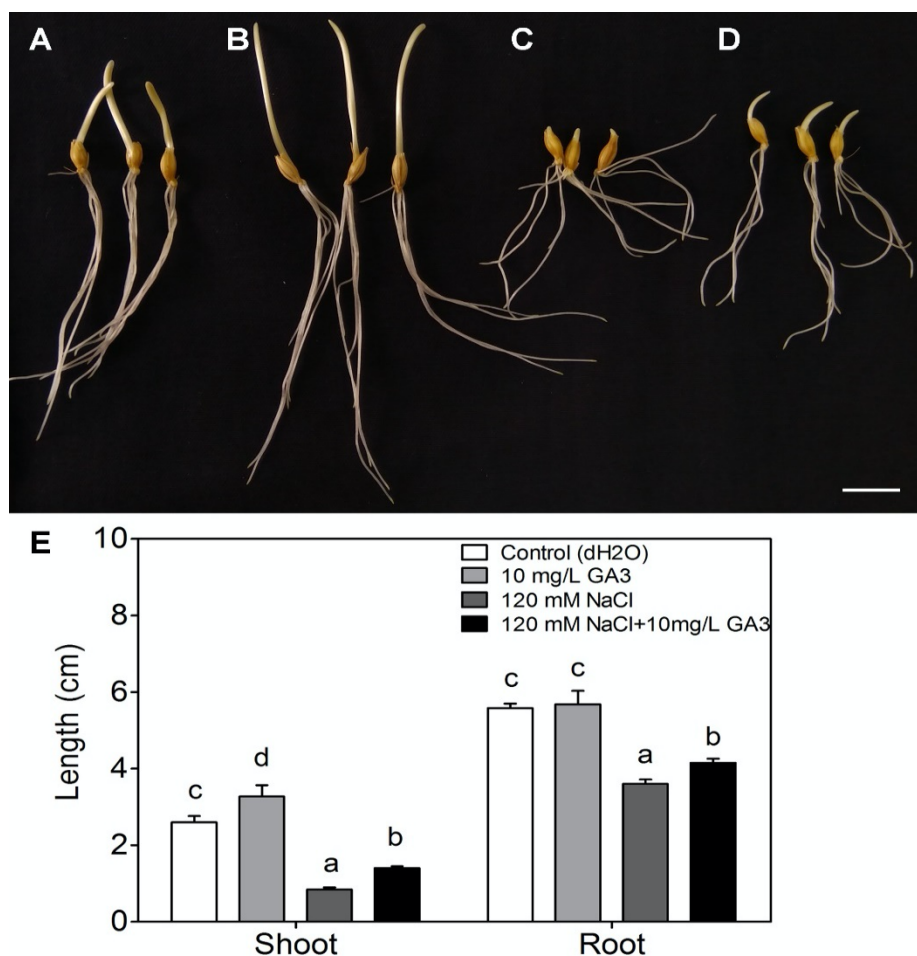


Figure 1: Phenotypes of 3-day-old barley sprouts germinated under different conditions: (A) dH₂O (Control), (B) 10 mg/L GA₃, (C) 120 mM NaCl, and (D) 120 mM NaCl+10 mg/L GA₃. The bar length is 1 cm. (E) Average shoot and root lengths of three-day-old barley sprouts from six replicates consisting of three biological and two technical replicates (n=10) with standard error of mean. Values denoted with the different letters differ significantly ($P < 0.05$, LSD post hoc test)

To cope with osmotic stress due to salinity, plants synthesize and accumulate osmolytes or compatible solutes, which are low-molecular-weight, highly soluble compounds, including proline, glycine betaine, and soluble sugars to increase osmotic potential of the cells, so to participate in osmotic adjustment [30]. In the present study, osmolyte content was significantly increased from 352 to 670 and 495 mosmol.kg⁻¹ by 120 mM NaCl treatment and 10 mg/L GA₃, respectively, at 3 DAI. When 10 mg/L GA₃ was applied with 120 mM NaCl, osmolyte content was found higher than GA₃-treated seeds and lower than salt-treated seeds. No synergetic effect was observed.

Proline plays a protective role acting as an important osmoprotectant, enzyme protectant, and free radical scavenger against salinity stress in plants [31]. 120 mM NaCl significantly

increased proline accumulation in both shoot and root tissues by 6.5 fold and 7.5 fold, respectively, compared to controls ($P < 0.05$). Exogenous 10 mg/L GA3 application did not alter the proline content in both shoot and root tissues compared with non-stressed sprouts at 3 DAI. Furthermore, combined application of 10 mg/L and 120 mM NaCl was observed to reduce proline accumulation in shoot and root by 57% and 66%, respectively, compared to seeds treated with 120 mM NaCl alone (Table 2). These results suggested that exogenous GA3 may induce the accumulation of other osmolytes instead of proline. Tuna et al. [32] have reported that foliar application of GA3 (50 mg/L) in maize seedlings under salt stress (100 mM NaCl) for 5 weeks slightly decreased in proline content compared to salt-stressed seedlings whilst higher GA3 concentration (100 mg/L) stimulated proline content leading to a big increase in similar conditions. In the present study, after soaking of seeds with 10 mg/L GA3 under salt stress for 3 days, proline content was found to have decreased by 58 and 65% in shoots and roots, respectively, compared to seeds soaking with 120 mM NaCl alone. Similarly, seedlings grown from seeds pre-soaked in GA3 (150 mg/L) was reported to exhibit less accumulation of proline than rice cultivars exposed to saline conditions (150 mg/L NaCl) for 10 days [33]. Manjili et al. [14] also have reported negative effects of GA3 (50 mg/L) application on proline content in three wheat cultivars in comparison with control and salinity conditions (3.5 and 7 dSm⁻¹). Consequently, both positive and negative interactions between GA and proline have been reported in the literature.

Ion leakage demonstrates membrane damage as a result of salt-induced oxidative damage. It has been reported that ion leakage increases under salt stress in the vegetative stage of plant [34]. In the shoot tissues, the ion leakage increased significantly from 38 to 55% after 120 mM NaCl treatment for 3 days during germination compared to control (dH₂O) at $P < 0.05$ level. On the other hand, when seeds were treated with 10 mg/L GA3 + 120 mM NaCl, ion leakage was significantly reduced from 55 to 21 % in shoot at 3 DAI ($P < 0.05$). Ion leakage was measured as 93% in roots of sprouts in response to salt stress. On the other hand, application of GA3 (10mg/L) was seen to increase membrane stability by lowering the elevated ion leakage level from 93 to 79% under salt stress. Tuna et al. [32] have reported that 50 and 100 mg/L GA3 treatments ameliorated negative effects of salt stress (100 mM NaCl) on membrane permeability by decreasing ion leakage upon GA3 concentration from 18 to 13 and 11%, respectively, in maize leaves. The significant reduction in ion leakage suggested that GA3 alleviated the harmful effects of salt stress on cell membrane, resulting in enhanced salt tolerance during germination.

Table 2: Proline content, ion leakage, and osmolality in shoot and root tissues of germinated seeds at 3 DAI under different conditions: dH₂O (C), 10 mg/L GA₃ (GA₃), 120 mM NaCl (NaCl) and 120 mM NaCl+10 mg/L GA₃ (NaCl+GA₃)

Treatment	Proline Content ($\mu\text{mol.g}^{-1}\text{FW}$)		Ion Leakage (%)		Osmolality (mosmol.kg^{-1})	
	Shoot	Root	Shoot	Root	Shoot	Root
C	3.21±0.64 a	4.10±0.77 a	37.62±3.33 b	82.86±2.86 a	352±30 a	NA
GA ₃	3.85±0.26 a	4.36±0.51 a	26.84±2.22 a	83.23±1.56 a	495±50 b	NA
NaCl	20.77±5.38 b	30.90±1.15 b	55.27±2.37 c	93.18±2.06 b	670±10 c	NA
NaCl+GA ₃	8.98±0.26 a	10.64±5.25 a	20.88±0.66 a	79.02±2.23 a	545±5 b	NA

Within each column, values denoted by same letters are not significantly different according to the LSD test at $P < 0.05$ level \pm Standard error.

3.2. Differential gene expression profiles of barley seeds after NaCl and/or GA₃ treatment

Effects of salt (120 mM NaCl), GA₃ (10 mg/L) and interaction of GA₃ and salt on the expression profiles of genes at 3 DAI in germinating barley seeds were analyzed by qPCR (Fig. 2). The genes to be analyzed were classified into 3 groups; group1 genes containing *HvABA7*, *HvNCED1*, *HvABI3*, *HvABI5*, and *HvSLN1* which are functional in suppression of seed germination, group2 genes including *HvGA2ox4*, *GA3ox1*, *GA20ox1*, *HvCPS1*, *HvKSI*, *HvKO1*, and *HvKAO1*, which are functional in GA₃ biosynthesis and induction of germination and group3 genes are functional in response to salt stress including *HvMT2*, *HvBASI*, *HvDRF1*, and *HvLOX1*.

Germination of seeds is a critical stage in a plant's life and involves multi-stage processes which are checked and regulated by differential expression of numerous genes in seed tissues [25]. GA₂₀-oxidase (GA₂₀ox) and GA₃-oxidase (GA₃ox) are major enzymes in the biosynthesis of Gas, whereas GA₂-oxidase (GA₂ox) has a role inactivation of GA [8]. In this study, application of exogenous GA₃ (10 mg/L) was found to increase the expression of *HvGA2ox4* in shoot while decreasing mRNA levels of *HvGA20ox1* and *HvGA3ox1*. It is suggested that the expression of these genes are differentiated by exogenous GA₃ to balance the endogenous GA level in cells. Furthermore, it has been showed that the active forms of gibberellins such as GA₃ could feedback regulate the expression levels of the *GA20ox* and *GA3ox* genes, which are the downstream genes in the pathway of GA biosynthesis [35]. Salt stress slightly decreased the expression of these three GA-oxidase genes compared to control conditions in shoot at 3DAI. On the other hand, only *GA2ox4* was up-regulated among these genes after 10 mg/L GA₃+120 mM NaCl treatment at 3 DAI (Fig. 2). Magome et al. [36] have reported six *GA2ox* genes, which are involved in the repression of growth under high-salinity conditions, were upregulated under high-salinity stress (150 mM NaCl) in Arabidopsis. In another study, after 150 mM NaCl treatment the

level of GA biosynthesis genes *GmGA3ox1* was also decreased in germinating soybean seeds at 9 h after sowing while expression of *GmGA2ox8*, which inactivates GA, increased compared with control [37]. In roots of germinated seeds under salt stress, exogenous GA3 was found to have no significant effect in expressions of *HvGA2ox4*, *GA20ox1*, and *HvGA3ox1* compared to salinity conditions (Fig. 2).

CPS, *KS*, *KO*, and *KAO* are the upstream genes in the GA biosynthesis pathway. *CPS* and *KS*, terpene cyclases, are located in the plastids and involved in conversion of geranylgeranyl diphosphate (GGDP) to ent-kaurene. *KO* and *KAO* are the cytochrome P450 monooxygenases, are located in the outer membrane of the plastid and the endoplasmic reticulum, respectively, catalyze ent-kaurene to GA12 aldehyde through a series of oxidation reactions [16, 35]. Salt stress slightly increased the mRNA level of *HvCPS1* and *KAOI 2* and 1.9 fold, respectively, while lowering the expression of *HvKOI* by 33 fold compared to control in shoot of germinated barley seeds at 3 DAI. 10 mg/L GA3 application was found to recover the decreased mRNA level of *HvKOI* due to the salt stress as the level of control conditions in shoot of germinating sprouts at 3 DAI. The expression patterns of *HvCPS1* and *HvKAO1* were found to be similar in both shoot and root tissues (Fig. 2). It has been reported that Gas down-regulated *GA20ox* and *GA3ox*, did not change expression of *CPS*, *KS*, *KO*, and *KAO* genes [16]. Huang et al. [35] have reported that *TaCPS*, *TaKS*, *TaKO*, and *TaKAO* genes were constitutively expressed in leaves, young spikes and stem of wheat, but their relative expression levels changed in different tissues. Furthermore, the homologs of those genes also showed the different expression pattern in wheat seedlings after GA3 treatment. In the present study, GA3 (10mg/L) altered the expression level of *HvCPS1*, *HvKSI*, *HvKAO1*, and *HvKOI* in barley sprouts at 3 DAI and those genes showed different expression levels in different tissues. These results indicate that genes involved in GA biosynthesis may play important role in germination of seeds and may function differently in tissues and different development stages of plants.

GA and ABA have antagonistic role in germination of seeds. GA enhances seed germination while ABA prevents or delays seed germination [8]. It has been reported that the reciprocal modulation of the expression of genes involved in ABA and GA metabolism and signaling plays role in coordination of changes in the ABA/GA balance [8, 38]. Salt stress has been reported to induce ABA biosynthesis by promoting expression of the *NCEDs* (9-cis-epoxycarotenoid dioxygenase) that codes a rate-limiting enzyme for ABA biosynthesis [39]. Increased accumulation of endogenous ABA content by induction of *HvNCEDI* with salt (120 mM NaCl) delayed seed germination and retarded shoot growth in barley. Addition of exogenous GA3 (10 mg/L) into root medium could help to reduce enhanced *NCEDI* level due to salinity in

shoot so recovered shoot growth in salt-stressed barley seeds. Intriguingly, salinity stress was found to result in a significant reduction in expression of *HvABA7* (10 fold) and *HvABI5* (3 fold) in shoots of germinated barley seeds at 3 DAI. These genes are participated in suppression of seed germination and generally upregulated by salinity stress. At present, only a few studies have been published on *HvABA7* gene and it has been reported that *HvABA7* was upregulated in response to salinity and drought stress in barley in contrast to this [40, 41]. Besides, application of exogenous GA3 resulted in an 18-fold increase in expression of *HvABA7* while no alteration was determined in expression of *HvABI5*. Enhanced expression of *OsNCED1* and *OsABI5* resulted in increased ABA level and response, which caused a delay in germination in rice [39]. Increased ABA content disturbs the balance between ABA and GA levels which leads to retardation in germination and growth. In the study, application of exogenous GA3 to barley seeds under salt stress was found to decrease the mRNA level of *HvNCED1* compared to salinity conditions at 3 DAI (Fig. 2).

Salinity shows the physiological effects on seed germination by differentiating the expression of some plant genes including group3 genes. It has been found that application of exogenous GA3 alone induced the expression of the salt-stress responsive genes including *HvMT2*, *HvDRF1*, *HvBAS1*, and *HvLOX1* in shoots of barley sprouts at 3 DAI. Even the expression level of *HvBAS1* and *HvDRF1* was higher in GA3-treated sprouts than in salt-treated sprouts (Fig. 2). *HvBAS1* play important role in detoxification of H₂O₂ as an antioxidant enzyme [7]. Exogenous GA3 may induce the germination and growth by enhancing the expression of the genes coding ROS scavenging enzymes including *HvBAS1* and *HvMT2*, which resulting in the reduction of the oxidative damage through accumulation of ROS.

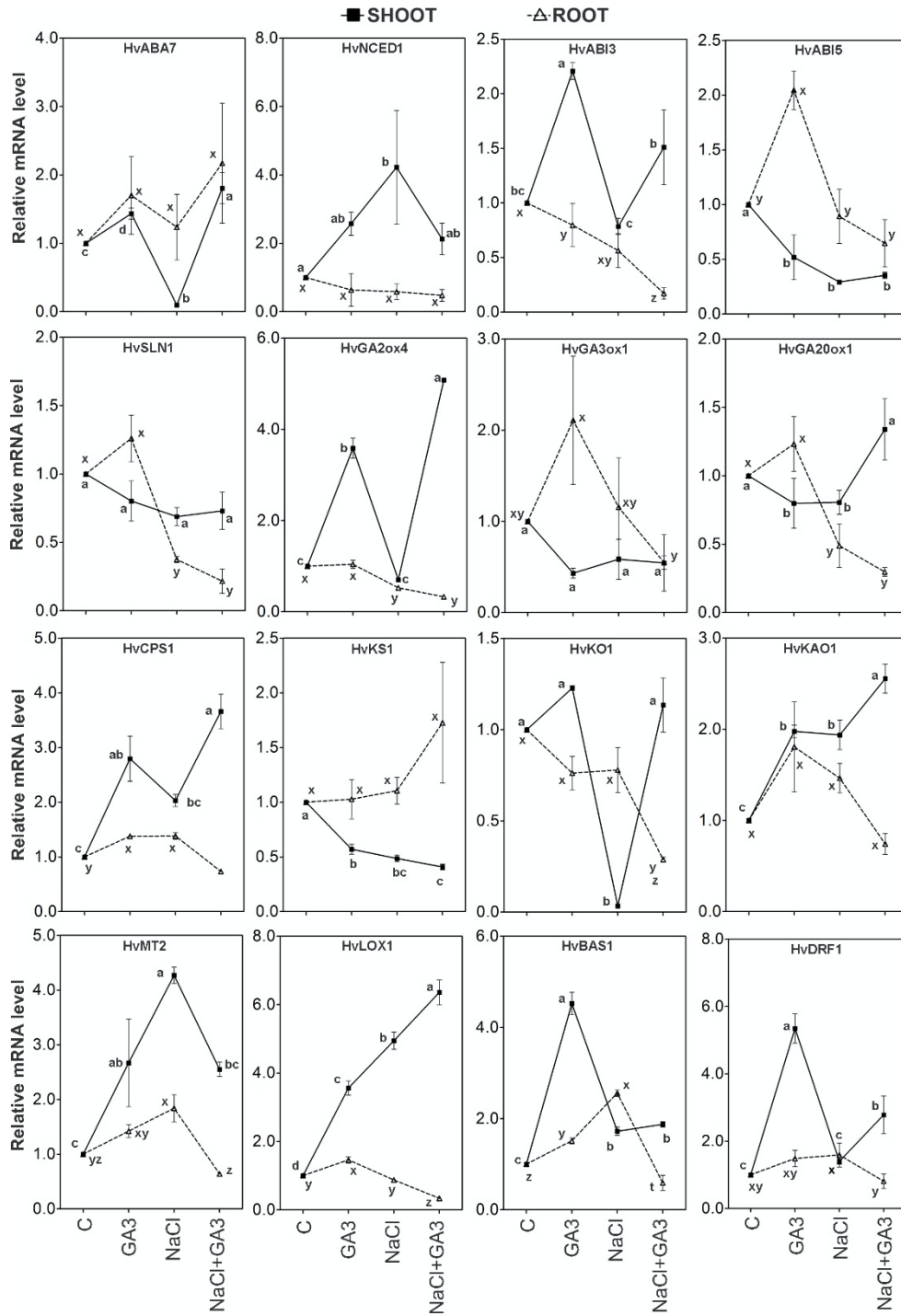


Figure 2: Gene expression profiles in shoot and root tissues of germinated seeds at 3 DAI. Barley seeds were treated with dH₂O (C), 10 mg/L GA₃ (GA₃), 120 mM NaCl (NaCl) and 120 mM NaCl+10 mg/L GA₃ (NaCl+GA₃) during germination. *HvGAPDH* was used as a reference gene. The mean values from six replicates consisting of three biological and two technical replicates with standard error. Values represented by different letters (a, b, c, d, and t, x, y, and z in shoot and root, respectively), are significantly different according to the LSD post hoc test at $P < 0.05$ level

4. Conclusion

Salinity is one of the most important abiotic stress limiting plant development and productivity in arid and semi-arid regions. Plants are most susceptible to salt stress in germination and young seedling formation stages. The exogenous application of GA3 significantly mitigated salt stress-induced delay of germination and growth inhibition in barley seeds. The results of the present study suggest that exogenous GA3 application depending on concentration could be helpful to improve germination and early seedling stages in crops under salt stress conditions. The expressions of the genes involved in GA3 metabolism and seed germination including the *HvABI5*, *HvABA7*, *HvKOl*, and *HvGA2ox4* were significantly differentiated after exogenous GA3 treatment. This study may give insight into relationship between salinity stress and genes involved in GA3 metabolism during germination and post-germination of barley seeds. However, high-throughput transcriptome analysis would reveal the mechanism underlying the role of GA3 in germination of seeds under salt stress.

Acknowledgement

This work was supported by the Scientific Research Projects Coordination Unit of Istanbul University, project no. 25516.

References

- [1] Zhu, J.K., *Abiotic stress signaling and responses in plants*, Cell 167, 313-324, 2016.
- [2] Munns, R., Tester, M. *Mechanisms of salinity tolerance*, Annual Review of Plant Biology, 59, 651-681, 2008.
- [3] Hasanuzzaman, M., Nahar, K., Fujita, M., *Plant response to salt stress and role of exogenous protectants to mitigate salt-induced damages*, In: Ahmad, P., Azooz, M.M., Prasad, M.N.V., (editors), *Ecophysiology and responses of plants under salt stress*, Springer, New York, pp. 25-87, 2013.
- [4] Kamran, M., Parveen, A., Ahmar, S., Malik, Z., Hussain, S., et al., *An overview of hazardous impacts of soil salinity in crops, tolerance mechanisms, and amelioration through selenium supplementation*, International Journal of Molecular Sciences, 21(1), 148, 2020.
- [5] Ibrahim, E.A., *Seed priming to alleviate salinity stress in germinating seeds*, Journal of Plant Physiology 192, 38-46, 2016.

- [6] Amirbakhtiar, N., Ismaili, A., Ghaffari, M.R., Nazarian F., Shobbar, Z.S., *Transcriptome response of roots to salt stress in a salinity-tolerant bread wheat cultivar*, PLoSONE, 14(3), e0213305, 2019.
- [7] Uçarlı, C., Gürel, F., *Differential physiological and molecular responses of three-leaf stage barley (*Hordeum vulgare* L.) under salt stress within hours*, Plant Biotechnology Reports, 14, 89-97, 2020.
- [8] Tuan, P.A., Kumar, R., Rehal, P.K., Toora, P.K., Ayele, B.T., *Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals*, Frontiers in Plant Science, 9, 668, 2018.
- [9] Miransari, M., Smith, D.L., *Plant hormones and seed germination*, Environmental and Experimental Botany, 99, 110–121, 2014.
- [10] Egamberdieva, D., Wirth, S.J., Alqarawi, A.A., Abd-Allah, E.F., Hashem, A., *Phytohormones and beneficial microbes: essential components for plants to balance stress and fitness*, Frontiers in Microbiology, 8, 1-14, 2017.
- [11] Liu, X., Hou, X., *Antagonistic regulation of ABA and GA in metabolism and signaling pathways*, Frontiers in Plant Science, 9:251, 2018.
- [12] Hedden, P., Sponsel, V., *A century of gibberellin research*, Journal of Plant Growth Regulation, 34, 740-760, 2015.
- [13] Binenbaum, J., Weinstain, R., Shani, E., *Gibberellin localization and transport in plants*, Trends in Plant Science, 23, 410-421, 2018.
- [14] Manjili, F.A., Sedghi, M., Pessarakli, M., *Effects of phytohormones on proline content and antioxidant enzymes of various wheat cultivars under salinity stress*, Journal of Plant Nutrition 35(7),1098-11, 2012.
- [15] Zehra, A., Shaikh, F., Ansari, R., Gul, B., Khan, M.A., *Effect of ascorbic acid on seed germination of three halophytic grass species under saline conditions*, Grass and Forage Science, 68, 339-344, 2013.
- [16] Yamaguchi, S., *Gibberellin metabolism and its regulation*, Annual Review of Plant Biology, 59, 225-251, 2008.
- [17] Kai, K., Kasa, S., Sakamoto, M., Aoki, N., Watabe, G., et al., *Role of reactive oxygen species produced by NADPH oxidase in gibberellin biosynthesis during barley seed germination*, Plant Signaling and Behavior, 11: e1180492.35, 753-759, 2016.
- [18] Chandler, P.M., Marion-Poll, A., Ellis, M., Gubler, F., *Mutants at the Slender1 locus of barley cv. Himalaya. Molecular and physiological characterization*, Plant Physiology, 129:, 181-190, 2002.

- [19] Davière, J.M., Achard, P., *A pivotal role of DELLAs in regulating multiple hormone signals*, *Molecular Plant*, 9, 10-20, 2016.
- [20] Zhou, M., *Barley production and consumption*, In: Zhang, G., Li, C., (editors), *Genetics and improvement of barley malt quality*, Springer-Verlag, Berlin Heidelberg, pp. 1-17, 2010.
- [21] Gao, R., Duan, K., Guo, G., Du, Z., Chen, Z., et al., *Comparative transcriptional profiling of two contrasting barley genotypes under salinity stress during the seedling stage*, *International journal of Genomics*, 972852, 2013.
- [22] Al-Karaki, G.N., *Germination, sodium, and potassium concentrations of barley seeds as influenced by salinity*, *Journal of Plant Nutrition*, 24(3), 511-522, 2001.
- [23] Carillo, P., Mastrodonato, G., Nacca, F., Parisi, D., Verlotta, A., et al., *Nitrogen metabolism in durum wheat under salinity: accumulation of proline and glycine betaine*, *Functional Plant Biology*, 35(5), 412-426, 2008.
- [24] Livak, K.J., Schmittgen, T.D., *Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-Delta Delta C (T)} method*, *Methods*, 25, 40-408, 2001.
- [25] Potokina, E., Sreenivasulu, N., Altschmied, L., Michalek, W., Graner, A., *Differential gene expression during seed germination in barley (*Hordeum vulgare* L.)*, *Functional and Integrative Genomics*, 2, 28-39, 2002.
- [26] Atia, A., Debez, A., Barhoumi, Z., Smaoui, A., Abdelly, C., *ABA, GA3, and nitrate may control seed germination of *Crithmum maritimum* (Apiaceae) under saline conditions*, *Comptes Rendus Biologie* 332, 704-710, 2009.
- [27] Ghosh, S, Mitra, S., Paul, A., *Physiochemical studies of sodium chloride on mungbean (*Vigna radiate* L. Wilczek) and its possible recovery with spermine and gibberellic acid*, *The Scientific World Journal*, Article ID 858016, 2015.
- [28] Abdel-Hamid, A.M.E., Mohamed, H.I., *The effect of the exogenous gibberellic acid on two salt stressed barley cultivars*, *European Scientific Journal*, 10, 228-245, 2014.
- [29] Maggio, A., Barbieri, G., Raimondi, G., De Pascale, S., *Contrasting effects of GA3 treatments on tomato plants exposed to increasing salinity*, *Journal of Plant Growth Regulation*, 2, 63–72, 2010.
- [30] Manivannan, P., Jaleel, C.A., Sankar, B., Kishorekumar, A., Somasundaram, R. et al., *Growth, biochemical modifications and proline metabolism in *Helianthus annuus* L. as induced by drought stress*, *Colloids and Surfaces. B, Biointerfaces*, 59, 141-149, 2007.
- [31] Verbruggen, N., Hermans, C., *Proline accumulation in plants: a review*, *Amino Acids*, 35(4), 753-759, 2008.

[32] Tuna, A.L., Kaya, C., Dikilitas, M., Higgs, D., *The combined effects of gibberellic acid and salinity on some antioxidant enzyme activities, plant growth parameters and nutritional status in maize plants*. *Environmental and Experimental Botany*, 62(1), 1-9, 2008.

[33] Chunthaburee, S., Sanitchon, J., Pattanagul, W., Theerakulpisut, P. *Alleviation of salt stress in seedlings of black glutinous rice by seed priming with spermidine and gibberellic acid*, *Notulae botanicae Horti Agrobotanic Cluj-Napoca*, 42, 405-413, 2014.

[34] Li, Q.Y., Niu, H.B., Yin, J., Wang, M.B., Shao, H.B., et al., *Protective role of exogenous nitric oxide against oxidative-stress induced by salt stress in barley (*Hordeum vulgare*)*, *Colloids and Surfaces. B, Biointerfaces*, 65, 220-225, 2008.

[35] Huang, Y., Yang, W., Pei, Z., Guo, X., Liu, D., et al., *The genes for gibberellin biosynthesis in wheat. Functional and Integrative Genomics*, 12, 199-206, 2012.

[36] Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y., Oda, K., *The DDF1 transcriptional activator upregulates expression of a gibberellin-deactivating gene, GA2ox7, under high-salinity stress in Arabidopsis*, *Plant Journal*, 56, 613-626, 2008.

[37] Shu, K., Qi, Y., Chen, F., Meng, Y., Luo, X., et al., *Salt stress represses soybean seed germination by negatively regulating GA biosynthesis while positively mediating aba biosynthesis*, *Frontiers in Plant Science*, 8, 1372, 2017.

[38] Izydorczyk, C., Nguyen, T.N., Jo, S., Son, S., Tuan, P.A., et al., *Spatiotemporal modulation of abscisic acid and gibberellin metabolism and signaling mediates the effects of suboptimal and supraoptimal temperatures on seed germination in wheat (*Triticum aestivum* L.)*, *Plant Cell Environment*, 41, 1022-1037, 2017.

[39] Zeng, D.E., Hou, P., Xiao, F., Liu, Y., *Overexpression of Arabidopsis XERICO gene confers enhanced drought and salt stress tolerance in rice (*Oryza Sativa* L.)*, *Journal of Plant Biochemistry and Biotechnology*, 24, 56-64, 2015.

[40] Walia, H., Wilson, C., Wahid, A., Condamine, P., Cui, X., et al., *Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress*, *Functional and Integrative Genomics*, 6, 143-156, 2006.

[41] Gürel, F., Öztürk, N.Z., Yörük, E., Uçarlı, C., Poyraz, N., *Comparison of expression patterns of selected drought-responsive genes in barley (*Hordeum vulgare* L.) under shock-dehydration and slow drought treatments*. *Plant Growth Regulation*, 80, 183-193, 2016.



On The Convergence Properties of Kantorovich-Szász Type Operators Involving Tangent Polynomials

Erkan AGYUZ^{1,*}

¹*Gaziantep University, Vocational School of Technical Science, Department of Electronic and Automation, 27310, Gaziantep, Türkiye*
agyuz@gantep.edu.tr, ORCID: 0000-0003-1110-7578

Received: 29.03.2021

Accepted: 24.08.2021

Published: 31.12.2021

Abstract

In this paper, we construct a variant of the Kantorovich Szász type operators involving the Tangent polynomials and estimate the convergence properties of these operators by using the Korovkin theorem. Also, we obtain the rate of convergence by using the modulus of continuity and Peetre's K-functional.

Keywords: Tangent polynomials; Kantorovich-Szász type operators; Rate of convergence; Generating functions; Korovkin theorems.

Tanjant Polinomlarını İçeren Kantorovich-Szász Tipli Operatörlerin Yakınsama Özellikleri Üzerine

Öz

Bu çalışmada, Tanjant polinomlarını içeren Kantorovich-Szász tipli operatörlerin bir varyantı oluşturulup, Korovkin teoremi kullanılarak bu operatörün yakınsaklık özellikleri tahmin edilmiştir. Ayrıca Peetre's K- fonksiyoneli ve süreklilik modülü kullanılarak yakınsaklık hızı elde edilmiştir.



Anahtar Kelimeler: Tanjant polinomları; Kantorovich-Szász tipli operatörler; Yakınsaklık hızı; Üreteç fonksiyonu; Korovkin Teoremi.

1. Introduction

It is well known the Tangent numbers are the coefficients in Taylor expansion for *tant* [1]:

$$\tan(t) = \sum_{n=0}^{\infty} (-1)^{n+1} T_{2n+1} \frac{t^{2n+1}}{(2n+1)!}, \quad T_{2n} = 0 \tag{1}$$

C. S. Ryoos defined the Tangent polynomials with the aid of generating function by using fermionic *p*-adic invariant integral in [2]:

$$\sum_{k=0}^{\infty} T_k(x) \frac{t^k}{k!} = \left(\frac{2}{e^{2t} + 1} \right) e^{xt}, \quad |t| < \frac{\pi}{2} \tag{2}$$

In the case $x = 0$, $T_k(0) = T_k$ is called *k* th Tangent number. Expressions of these numbers through generating functions:

$$\sum_{k=0}^{\infty} T_k \frac{t^k}{k!} = \left(\frac{2}{e^{2t} + 1} \right), \quad |t| < \frac{\pi}{2} \tag{3}$$

Because of obtaining generating function for a set of functions, we derive differential recurrence relation, pure recurrence relation or calculate certain integral. Using generating functions, many fundamental properties and identities for special polynomials and numbers can be obtained [3, 4]. For $k \in \mathbb{Z}^+$, Tangent polynomials are expressed by using generating functions at the following identity:

$$T_k(x) = \sum_{l=0}^k \binom{k}{l} x^{k-l} T_l. \tag{4}$$

Nowadays, many researchers study the central moment and test function notations which are useful for examining the convergence of sequences of linear positive operators. Moreover, central moments of Kantorovich type operators are investigated with the aid of moment-generating functions [5-13].

The organization of this paper is as follows:

In section 2, we give a variant of the Kantorovich-Szász type operators involving the Tangent polynomials and estimate the convergence properties of these operators by using

Korovkin’s theorem. In section 3, we evaluate the rate of convergence by using the modulus of continuity and Peetre’s K-functional.

2. Convergence Properties

Firstly, we define our new operator involving Tangent polynomials at the following equation:

$$K_n^*(f, x) = n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t) dt, \tag{5}$$

where $T_{2k+1} > 0$ and $k \in \mathbb{N}$.

Now, we construct the moments and test functions for $r = \{0,1,2\}$ in the following lemmas:

$$e_r = t^r \text{ and } \vartheta^r(t, x) = (t - x)^r$$

We give the values of moments under $K_n^*(f, x)$ in Lemma 1 as follows:

Lemma 1. For $\forall x \in [0, \infty)$, Eqn. (1) has the following properties:

$$K_n^*(e_0, x) = 1,$$

$$K_n^*(e_1, x) = x + \frac{1 - 3e^2}{2n(1 + e^2)},$$

$$K_n^*(e_2, x) = x^2 + \frac{x(1-e^2)}{n}.$$

Proof. Take $f = e_0$ in Eqn. (5). Then we have

$$\begin{aligned} K_n^*(1, x) &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} dt \\ &= n \frac{e^2 + 1}{2} e^{-nx} \frac{2}{e^2 + 1} e^{nx} \left(\frac{k + 1}{n} - \frac{k}{n} \right) \\ &= 1. \end{aligned}$$

Taking the derivative of both sides of Eqn. (5) with respect to t , we get

$$\sum_{k=0}^{\infty} k T_k(x) \frac{t^{k-1}}{k!} = \frac{-4e^{2t}}{(e^{2t} + 1)^2} e^{xt} + x \frac{2}{e^{2t} + 1} e^{xt}. \tag{6}$$

Replacing $t = 1$ and $x = nx$ in Eqn. (6), we have

$$\sum_{k=0}^{\infty} \frac{kT_k(x)}{k!} = \frac{-4e^2}{(e^2 + 1)^2} e^{xt} + x \frac{2}{e^2 + 1} e^{nx}. \tag{7}$$

In the same manner, we obtain

$$\sum_{k=0}^{\infty} \frac{k^2 T_k(x)}{k!} = (nx)^2 \frac{2}{e^2 + 1} e^{nx} + \left(\frac{2 - 6e^2}{(e^2 + 1)^2} \right) nxe^{nx} + \frac{2e^6 - 8e^4 - 12e^2}{(e^2 + 1)^4} e^{nx}. \tag{8}$$

For $r = 1$, we have

$$\begin{aligned} K_n^*(t, x) &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} t dt \\ &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \left[\frac{t^2}{2} \right]_{\frac{k}{n}}^{\frac{k+1}{n}} \\ &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \frac{(2k + 1)}{n^2} \\ &= 2n \frac{e^2 + 1}{2n^2} e^{-nx} \sum_{k=0}^{\infty} k \frac{T_k(nx)}{k!} + n \frac{e^2 + 1}{2n^2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \\ &= x + \frac{1-3e^2}{2n(1+e^2)}. \end{aligned}$$

For $r = 2$, we obtain

$$\begin{aligned} K_n^*(t^2, x) &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} t^2 dt \\ &= n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \left[\frac{t^3}{3} \right]_{\frac{k}{n}}^{\frac{k+1}{n}} \\ &= \frac{n}{3} \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \frac{(3k^2 + 3k + 1)}{n^3} \\ &= \frac{e^2 + 1}{2n^2} e^{-nx} \sum_{k=0}^{\infty} k^2 \frac{T_k(nx)}{k!} + \frac{e^2 + 1}{2n^2} e^{-nx} \sum_{k=0}^{\infty} k \frac{T_k(nx)}{k!} \\ &\quad + \frac{1}{3} \frac{e^2 + 1}{2n^2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \\ &= x^2 + \frac{x(1 - e^2)}{n}. \end{aligned}$$

This completes the proof.

Lemma 2. For $\forall x \in [0, \infty)$, we have

$$K_n^*((t-x), x) = \frac{1-3e^2}{2n(1+e^2)}, \tag{9}$$

$$K_n^*((t-x)^2, x) = \frac{-e^4+3e^2}{n(1+e^2)} x. \tag{10}$$

Proof. By using the linearity of (K_n^*) , we have

$$\begin{aligned} K_n^*((t-x), x) &= K_n^*(t, x) - xK_n^*(1, x) = \frac{1-3e^2}{2n(1+e^2)} \\ K_n^*((t-x)^2, x) &= K_n^*(t^2, x) - 2xK_n^*(t, x) + x^2K_n^*(1, x) \\ &= x^2 + \frac{x(1-e^2)}{n} - 2x\left(x + \frac{1-3e^2}{2n(1+e^2)}\right) + x^2 \\ &= \frac{x(1-e^2)}{n} - \frac{x(1-3e^2)}{n(1+e^2)} \\ &= \frac{-e^4+3e^2}{n(1+e^2)} x. \end{aligned}$$

So the desired results are obtained.

Now we give the main theorem by using the Lemma 1 and Korovkin’s theorem in the following:

Theorem 3. Let $f \in C_E[0, \infty) = C[0, \infty) \cap E$, where

$$E := \left\{ f: x \in [0, \infty), \frac{f(x)}{1+x^2} \text{ is convergent as } x \rightarrow \infty \right\}$$

and

$$\lim_{y \rightarrow \infty} \frac{B'(y)}{B(y)} = 1 \text{ and } \lim_{y \rightarrow \infty} \frac{B''(y)}{B(y)} = 1.$$

Then the sequence (K_n^*) converges uniformly to f on $[0, \infty)$ as $n \rightarrow \infty$.

Proof. By applying the well-known Korovkin’s first theorem and Lemma 1, we obtain

$$\lim_{n \rightarrow \infty} K_n^*(e_i, x) = x^i, i = 0,1,2. \tag{11}$$

and the operator (K_n^*) converges uniformly in each compact subset of $[0, \infty)$.The proof is completed by using the property (vii) of Theorem 4.1.4 in [6].

Estimating the degree of approximation by positive linear operators involving generating functions of special polynomials has several methods. One of them is the modulus of continuity given by

$$w(f, \delta) = \sup\{f(t) - f(x), x, t \in [a, b], |t - x| \leq \delta\}, f \in C[0, \infty], \quad \delta > 0, \quad (12)$$

where $C[0, \infty]$ is called as space of uniformly continuous functions on $[0, \infty]$. Also,

$$|f(t) - f(x)| \leq w(f, \delta) \left(\frac{|t - x|}{\delta} + 1 \right)$$

is satisfied for any $\delta > 0$ and each $x \in [0, \infty)$

We give an estimate for degree of approximation of $K_n^*(f, x)$ with the help of definition of modulus of continuity in the following theorem:

Theorem 4. If $f \in C[0, \infty] \cap E$, then

$$|K_n^*(f, x) - f(x)| \leq 2w \left(f; \sqrt{\frac{-e^4 + 3e^2}{n(1 + e^2)} x} \right). \quad (13)$$

Proof. We have

$$\begin{aligned} |K_n^*(f, x) - f(x)| &\leq n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |f(t) - f(x)| d_t \\ &\leq n \frac{e^2 + 1}{2} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} \left(\frac{|t - x|}{\delta} + 1 \right) w(f; \delta) d_t \\ &\leq \left(1 + \frac{n e^2 + 1}{\delta} e^{-nx} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |t - x| d_t \right) w(f; \delta). \end{aligned} \quad (14)$$

By applying Cauchy-Schwarz inequality, we get

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} |t - x| d_t \leq \frac{1}{\sqrt{n}} \left(\int_{\frac{k}{n}}^{\frac{k+1}{n}} |t - x|^2 d_t \right)^{\frac{1}{2}}, \quad (15)$$

that satisfies

$$\sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |t-x| d_t \leq \frac{1}{\sqrt{n}} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \left(\int_{\frac{k}{n}}^{\frac{k+1}{n}} |t-x|^2 d_t \right)^{\frac{1}{2}}. \tag{16}$$

By using the Cauchy-Schwarz inequality for the sum on the right side of Eqn. (16), we have

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |t-x| d_t &\leq \frac{\sqrt{\frac{2}{e^2+1}} e^{nx}}{\sqrt{n}} \left(\frac{2}{n} \frac{e^{nx}}{e^2+1} K_n^*((t-x)^2; x) \right)^{\frac{1}{2}} \\ &= \frac{e^{nx}}{n} \left(\frac{2}{e^2+1} \right) (K_n^*((t-x)^2; x))^{\frac{1}{2}}. \end{aligned} \tag{17}$$

From Lemma 2, using the second-order test function of the operator (K_n^*), we have the final form of Eqn. (17) as follows:

$$\sum_{k=0}^{\infty} \frac{T_k(nx)}{k!} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |t-x| d_t \leq \frac{e^{nx}}{n} \frac{2}{e^2+1} \left(\frac{-e^4+3e^2}{n(1+e^2)} \right)^{\frac{1}{2}}.$$

Hence, the proof is completed.

Another method for estimate of rate of approximation is the second order modulus of continuity given by

$$w_2(f, h) = \sup \left\{ \left| f(t) - 2f\left(\frac{t+x}{2}\right) + f(x) \right|; x, t \in [a, b], |t-x| \leq 2h \right\}, \tag{18}$$

where $f \in C[a, b], h > 0$.

The classical Peetre's K - functional is defined by

$$K(f, \delta) = \inf_{g \in C_B^2[0, \infty)} \{ \|f - g\| + \delta \|g''\| \}, \tag{19}$$

where $f \in C_B[0, \infty)$ and $C_B^2[0, \infty) = \{g \in C_B[0, \infty): g', g'' \in C_B[0, \infty)\}$.

Theorem 5. For the $K_n^*(f, x)$ operators, we have

$$|K_n^*(f, x) - f(x)| \leq \psi \|f\|_{C_B^2}, \tag{20}$$

where $\psi := \frac{-e^4+3e^2}{n(1+e^2)} x + \frac{1-3e^2}{2n(1+e^2)}$.

Proof. By applying linearity properties of $K_n^*(f, x)$ and Taylor expansion of f , we have

$$K_n^*(f, x) - f(x) = f(x)'K_n^*(s - x, x) + \frac{1}{2}f(\rho)''K_n^*((s - x)^2, x), \rho \in (x, s). \tag{21}$$

From Lemma 2, we get

$$|K_n^*(f, x) - f(x)| \leq \frac{-e^4 + 3e^2}{n(1 + e^2)} x \|f'\|_{C_B} + \frac{1}{2} \frac{1 - 3e^2}{2n(1 + e^2)} \|f''\|_{C_B}. \tag{22}$$

Therefore, the desired result is obtained.

Theorem 6. Let $f \in C_B[0, \infty)$. Then

$$|K_n^*(f, x) - f(x)| \leq 2M \left\{ w_2 \left(f, \sqrt{\frac{1}{2}\psi} \right) + \min \left(1, \frac{1}{2}\psi \right) \right\} \|f\|_{C_B}, \tag{23}$$

where $M > 0$.

Proof. We rewrite $f(t) - f(x)$ as

$$f(t) - f(x) = f(t) - g(t) + g(t) - g(x) + g(x) - f(x). \tag{24}$$

From the linearity properties of $K_n^*(f, x)$, we have

$$|K_n^*(f, x) - f(x)| \leq |K_n^*(f - g, x) - f(x)| + |K_n^*(g, x) - g(x)| + |f(x) - g(x)|. \tag{25}$$

where we assume that the function $g \in C^2[0, \infty)$. By using Theorem 5, we get

$$|K_n^*(f, x) - f(x)| \leq 2\|f - g\|_{C_B} + \psi \|g\|_{C_B^2} = 2 \left[\|f - g\|_{C_B} + \psi \|g\|_{C_B^2} \right]. \tag{26}$$

From Eqn. (26), we obtain

$$|K_n^*(f, x) - f(x)| \leq 2K(f, \psi), \tag{27}$$

where $K(f, \psi)$ is Peetre's K- functional defined by Eq.(19). We obtain the desired result.

4. Conclusion

The generating function has many useful applications in several fields. For example, generating function of Bernstein polynomials provides important results to constructing Bezier curves. In literature, many properties and relations are obtained by using the generating function of Tangent polynomials.

Recently, the approximation theory is one of the important application fields of generating function of special polynomials which have useful properties for constructing linear positive operators.

In this paper, we constructed a Kantorovich-Szász type operator, $K_n^*(f, x)$, involving the generating function of Tangent polynomials. Then, we investigated some properties such as modulus of continuity, second-order modulus of continuity, and Peetre's K functional for $K_n^*(f, x)$.

Acknowledgements

I would like to thank the referees for their valuable comments and suggestions on this paper.

References

- [1] Qi, F., *Derivatives of tangent function and tangent numbers*, Applied Mathematics and Computations, 268, 844-858, 2015.
- [2] Ryoo, C.S., *A note on the tangent numbers and polynomials*, Advanced Studies in Theoretical Physics, 7(9), 447-454, 2013.
- [3] Ryoo, C.S., *On poly-tangent numbers and polynomials and the distribution of their zeros*, Journal of Applied Mathematics and Informatics, 34, 487-494, 2016.
- [4] Ryoo, C.S., *Differential equations associated with Tangent numbers*, Journal of Applied Mathematics and Informatics, 34, 487-494, 2016.
- [5] Paltanea, R., *Approximation theory using positive linear operators*, Birkhauser Boston, 2004.
- [6] Altomare, F., Campiti M. *Korovkin-type approximation theory and its applications*, Walter de Gruyter, 17, 2011.
- [7] Aktaş, R., Söylemez, D., Taşdelen, F., *Stancu type generalization of Szász-Durrmeyer operators involving Brenke-type polynomials*, Filomat, 33(3), 855-868, 2019.
- [8] Taşdelen, F., Aktaş, R., Altun, A., *A Kantorovich type of Szász operators including Brenke-type polynomials*, Abstract and Applied Analysis, 2012, 2012.
- [9] Varma, S., Taşdelen, F., *Szász type operators involving Charlier polynomials*, Mathematical and Computer Modelling, 56(5-6), 118-122, 2012.
- [10] Cekim, B., Aktaş, R., İçöz, G., *Kantorovich-Stancu type operators including Boas-Buck type polynomials*, Hacettepe Journal of Mathematics and Statistics, 48(2), 460-471, 2019.
- [11] İçöz, G., Varma, S., Sucu, S., *Approximation by operators including generalized Appell polynomials*. Filomat, 30(2), 429-440, 2016.
- [12] Atakut, Ç., Büyükyazıcı, İ., *Approximation by Kantorovich-Szász type operators based on Brenke type polynomials*, Numerical Functional Analysis and Optimization, 37(12), 1488-1502, 2016.
- [13] Tasdelen, F., Söylemez, D., Aktaş, R. *Dunkl Gamma Type Operators including Appell polynomials*, Complex Analysis and Operator Theory, 13, 3359-3371, 2019.



Fixed Points for Generalized Type Contractions in Partially Ordered Metric Spaces

Seher Sultan YEŞİLKAYA^{1,*}

¹*Kahramanmaraş Sütçü İmam University, Department of Mathematics, 46040, Kahramanmaraş, Türkiye
sultanseher20@gmail.com, ORCID: 0000-0002-1748-2398*

Received: 03.11.2020

Accepted: 10.09.2021

Published: 31.12.2021

Abstract

In this article, we define ordered weak θ -contractive and ordered Ćirić type weak θ -contractive mappings in partially ordered metric spaces. We also introduce some fixed point theorems for such mappings. These theorems extend the main theorems of many comparable results from the current literature. Finally, an example is showed to support the new theorems.

Keywords: Fixed point theorem; Partially ordered metric spaces; Regular mapping.

Kısmi Sıralı Metrik Uzaylarda Genelleştirilmiş Tip Daralmalar için Sabit Noktalar

Öz

Bu makalede, kısmi sıralı metrik uzaylarda sıralı zayıf θ -daralma ve sıralı Ćirić tipi zayıf θ -daralma dönüşümleri tanımlanmıştır. Ayrıca, bu tür dönüşümler için bazı sabit nokta teoremleri tanıtılmıştır. Bu teoremler, mevcut literatürden birçok karşılaştırılabilir sonucun ana teoremlerini genişletir. Son olarak, yeni teoremleri destekleyen bir örnek gösterilmiştir.

Anahtar Kelimeler: Sabit nokta teoremi; Kısmi sıralı metrik uzaylar; Düzenli dönüşüm.



1. Introduction and Preliminaries

Fixed point theory in metric spaces begins with the Banach contraction mapping, introduced in 1922 [1]. The existence of fixed points in partially ordered metric spaces has been investigated in [2]. Fixed points of operators in partially ordered metric spaces are very significant and have been studied by several authors [3-9].

We said that X is regular if the ordered metric spaces (X, \preceq, d) supplies the following cases: If $\{x_r\}$ is an increasing sequence in X with respect to \preceq such that $x_r \rightarrow v \in X$, then $x_r \preceq v$ for all $r \in \mathbb{N}$.

Banach contraction principle has been weakened and generalized by many researchers. For example, the notions of θ -contraction [10], F -contraction [11], Ćirić contraction [12], weak generalized contraction [4], have been established, and several generalizations of this principle are obtained.

Jleli and Samet denote by Θ the set of functions $\theta: (0, \infty) \rightarrow (1, \infty)$ satisfying the following conditions:

(Θ_1) θ is non-decreasing;

(Θ_2) for each sequence $\{t_n\} \subset (0, \infty)$, $\lim_{n \rightarrow \infty} \theta(t_n) = 1$ if and only if $\lim_{n \rightarrow \infty} t_n = 0^+$;

(Θ_3) there exist $r \in (0, 1)$ and $l \in (0, \infty]$ such that $\lim_{t \rightarrow 0^+} \frac{\theta(t)-1}{t^r} = l$.

Define by Ψ the set of functions $\Psi: [0, \infty) \rightarrow [0, \infty)$ satisfying the following conditions:

(i) ψ is non-decreasing;

(ii) for each $k > 1$, $\lim_{n \rightarrow \infty} \psi^n(k) = 0$;

(iii) $\psi(0) = 0$, and for each $k > 1$, $\psi(k) < k$.

In this study, we present new contractive mappings in partially ordered metric spaces, inspired by the papers of Samira et al. [13], Ćirić [12], Jleli and Samet [10].

2. Main Results

Now, we introduce some fixed point theorems for ordered weak θ -contractive and ordered Ćirić type weak θ -contractive mappings in partially ordered metric space. We begin this section with the definition of ordered weak θ -contractive.

Definition 1. Let (X, \preceq, d) be a partially ordered metric space. and $K: X \rightarrow X$ be a self mapping. Let

$$Q = \{(x, y) \in X \times X: x \preceq y, d(Kx, Ky) > 0\} \tag{1}$$

and $\theta \in \Theta$ and $\psi \in \Psi$. We say that K is an ordered weak θ -contractive if there exists $\beta \in (0,1)$ such that

$$\theta(d(Kx, Ky)) \leq [\theta(\psi(d(x, y)))]^\beta, \tag{2}$$

for all $(x, y) \in Q$.

Theorem 1. Let (X, \preceq, d) be a partially ordered complete metric space and $K: X \rightarrow X$ is an ordered weakly θ -contractive. Suppose that $\theta \in \Theta$, $\psi \in \Psi$ and K is non-decreasing and there exists $x_0 \in X$ such that $x_0 \preceq Kx_0$. Therefore, K has a fixed point in X provided that at least one of the following conditions holds

- (i) K is continuous,
- (ii) X is regular.

Proof. Starting from an arbitrary point $x_0 \in X$. We consider the constructive sequence $\{x_r\} \subset X$ which is defined by $x_r = Kx_{r-1} = K^r x_0$, for all $r \in \mathbb{N}$. Suppose that there exists $r_0 \in \mathbb{N}$ such that $x_{r_0} = x_{r_0+1}$, then $x_{r_0} = x_{r_0+1} = Kx_{r_0}$ and so the proof is completed.

Now assume that for all $r \in \mathbb{N}$, $x_{r+1} \neq x_r$. Since $x_0 \preceq Kx_0$ and K is non-decreasing, we get

$$x_0 \preceq x_1 \preceq x_3 \preceq \dots \preceq x_r \preceq \dots.$$

From $x_r \preceq x_{r+1}$ and $d(Kx_{r-1}, Kx_r) > 0$ for all $r \in \mathbb{N}$, we have $(x_r, x_{r+1}) \in Q$. So, using Eqn. (2) we obtain

$$\theta(d(x_r, x_{r+1})) = \theta(d(Kx_{r-1}, Kx_r)) \leq [\theta(\psi(d(x_{r-1}, x_r)))]^\beta, \tag{3}$$

for all $r \in \mathbb{N}$. Since condition (Θ_1) , we obtain

$$\begin{aligned} \theta(d(K^r x_0, K^{r+1} x_0)) &\leq [\theta(\psi(d(K^{r-1} x_0, K^r x_0)))]^\beta \\ &\leq [\theta(\psi(d(K^{r-2} x_0, K^{r-1} x_0)))]^{\beta^2} \\ &\vdots \\ &\leq [\theta(\psi(d(x_0, Kx_0)))]^{\beta^r}. \end{aligned} \tag{4}$$

Letting $r \rightarrow \infty$ in the above inequality, we get

$$\lim_{r \rightarrow \infty} \theta(d(x_r, x_{r+1})) = 1. \tag{5}$$

From (Θ_2) this implies that

$$\lim_{r \rightarrow \infty} d(x_r, x_{r+1}) = 0^+.$$

Using (Θ_3) , there exists $w \in (0,1)$ and $V \in (0, \infty]$ such that

$$\lim_{r \rightarrow \infty} \frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} = V. \tag{6}$$

Suppose that $V < \infty$. In this case, let $S = \frac{V}{2} > 0$. From the definition of the limit, there exists $r_0 \in \mathbb{N}$ such that

$$\left| \frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} - V \right| \leq S, \quad \text{for all } r \geq r_0.$$

Then we get

$$\frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} \geq V - S = S, \quad \text{for all } r \geq r_0.$$

Then for all $r \geq r_0$, we obtain

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1],$$

where $H = \frac{1}{S}$. Assume that $V = \infty$. Let $S > 0$ be an arbitrary positive number. Thus there exists $r_0 \in \mathbb{N}$ such that

$$\frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} \geq S,$$

for all $r \geq r_0$. This implies that for all $r \geq r_0$,

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1],$$

where $H = \frac{1}{S}$. Therefore, in two cases, there exists $H > 0$ and $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$,

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1].$$

By using Eqn. (4), we get

$$r(d(x_r, x_{r+1}))^w \leq Hr([\theta(\psi(d(x_0, x_1)))]^{\beta^r} - 1), \tag{7}$$

for all $r \geq r_0$. Letting $r \rightarrow \infty$ in Eqn. (7), we obtain

$$\lim_{r \rightarrow \infty} r(d(x_r, x_{r+1}))^w = 0.$$

Therefore, there exists $r_1 \in \mathbb{N}$ such that

$$d(x_r, x_{r+1}) \leq \frac{1}{r^w}, \text{ for all } r \geq r_1. \tag{8}$$

Next, we prove that $\{x_r\}$ is a Cauchy sequence in K . There exists sequences $r, p \in \mathbb{N}$ such that $p > r \geq r_1$. Then from Eqn. (8), we obtain

$$\begin{aligned} d(x_r, x_p) &\leq d(x_r, x_{r+1}) + d(x_{r+1}, x_{r+2}) + \dots + d(x_{p-1}, x_p) \\ &\leq \sum_{i=r}^{p-1} \frac{1}{i^w} \leq \sum_{i=r}^{\infty} \frac{1}{i^w}. \end{aligned}$$

By the convergence of the series $\sum_{i=r}^{\infty} \frac{1}{i^w}$, in the limit $r \rightarrow \infty$, we get $d(x_r, x_p) \rightarrow 0$. Then $\{x_r\}$ is a Cauchy sequence in (X, d) . Because (X, d) is a complete metric space, there exists $v \in X$ such that

$$\lim_{r \rightarrow \infty} x_r = v. \tag{9}$$

If K is continuous, then we have

$$v = \lim_{r \rightarrow \infty} x_{r+1} = \lim_{r \rightarrow \infty} Kx_r = K \lim_{r \rightarrow \infty} x_r = Kv.$$

Thus $v = Kv$, and v is a fixed point of K .

We assume that X is regular, so $x_r \leq v$ for all $r \in \mathbb{N}$. Two cases arise here:

Case 1. If there exists $u \in \mathbb{N}$ for $x_u = v$, so

$$Kv = Kx_u = x_{u+1} \leq v.$$

In addition, we have $x_u \leq x_{u+1}$. So, $v \leq Kv$ and thus $v = Kv$.

Case 2. Given that $x_r \neq v$ for all $r \in \mathbb{N}$ and $d(v, Kv) > 0$. From $\lim_{r \rightarrow \infty} x_r = v$, there exists $r_1 \in \mathbb{N}$ such that $d(x_{r+1}, Kv) > 0$ and $d(x_r, v) < \frac{d(v, Kv)}{2}$ for every $r \geq r_1$. In addition, $(x_r, v) \in Q$. Thus, from (Θ_1) and ψ non-decreasing we obtain

$$\begin{aligned} \theta(d(Kx_r, Kv)) &\leq [\theta(\psi(d(x_r, v)))]^\beta \\ &\leq \theta(\psi(d(x_r, v))) \\ &\leq \theta\left(\psi\left(\frac{d(v, Kv)}{2}\right)\right). \end{aligned}$$

This implies that

$$d(x_{r+1}, Kv) \leq \psi\left(\frac{d(v, Kv)}{2}\right) < \frac{d(v, Kv)}{2}.$$

Taking limit as $r \rightarrow \infty$, we obtain

$$d(v, Kv) < \frac{d(v, Kv)}{2},$$

a contraction. Hence, we get $d(v, Kv) = 0$, that is, $v = Kv$. This concludes the proof.

Definition 2. Let (X, \leq, d) be a partially ordered metric space and $K: X \rightarrow X$ be a self mapping. Let

$$Q = \{(x, y) \in X \times X: x \leq y, d(Kx, Ky) > 0\}. \tag{10}$$

Further, $\theta \in \Theta$ and $\psi \in \Psi$. We say that K is an ordered Ćirić type weak θ -contractive if there exists $\beta \in (0,1)$ such that

$$\theta(d(Kx, Ky)) \leq [\theta(\psi(P(x, y)))]^\beta, \tag{11}$$

for all $(x, y) \in Q$, where

$$P(x, y) = \max\{d(x, y), d(x, Kx), d(y, Ky), \frac{1}{2}[d(x, Ky) + d(y, Kx)]\}. \tag{12}$$

Theorem 2. Let (X, \leq, d) be a partially ordered complete metric space and $K: X \rightarrow X$ is an ordered Ćirić type weak θ -contractive. Assume that $\theta \in \Theta$, $\psi \in \Psi$ and K is non-decreasing and there exists $x_0 \in X$ such that $x_0 \leq Kx_0$. Therefore, K has a fixed point in X provided that at least one of the following conditions holds

- (i) K is continuous,
- (ii) X is regular.

Proof. Given an arbitrary point $x_0 \in X$ we consider the constructive sequence $\{x_r\} \subset X$ which is defined by $x_r = Kx_{r-1} = K^r x_0$, for all $r \in \mathbb{N}$. Assume there exists $r_0 \in \mathbb{N}$ such that $x_{r_0} = x_{r_0+1}$, then $x_{r_0} = x_{r_0+1} = Kx_{r_0}$ and so the proof is completed.

Now assume that for all $r \in \mathbb{N}$, $x_{r+1} \neq x_r$. As $x_0 \leq Kx_0$ and K is non-decreasing, we get

$$x_0 \leq x_1 \leq x_3 \leq \dots \leq x_r \leq \dots.$$

From $x_r \leq x_{r+1}$ and $d(Kx_{r-1}, Kx_r) > 0$, for all $r \in \mathbb{N}$, we have $(x_r, x_{r+1}) \in Q$. So, from Eqn. (11) we obtain

$$\begin{aligned} \theta(d(x_r, x_{r+1})) &= \theta(d(Kx_{r-1}, Kx_r)) \\ &\leq [\theta(\psi(\max\{d(x_{r-1}, x_r), d(x_{r-1}, Kx_{r-1}), d(x_r, Kx_r), \\ &\quad \frac{1}{2}[d(x_{r-1}, Kx_r) + d(x_r, Kx_{r-1})]\}))]^\beta \\ &= [\theta(\psi(\max\{d(x_{r-1}, x_r), d(x_r, x_{r+1})\}))]^\beta, \end{aligned}$$

for all $r \in \mathbb{N}$. Then suppose that

$$d(x_{r-1}, x_r) < d(x_r, x_{r+1}).$$

We get

$$\theta(d(x_r, x_{r+1})) \leq [\theta(\psi(d(x_r, x_{r+1})))]^\beta$$

a contradiction. Then we obtain

$$d(x_r, x_{r+1}) \leq d(x_{r-1}, x_r).$$

Hence we have

$$\theta(d(Kx_{r-1}, Kx_r)) \leq [\theta(\psi(d(x_{r-1}, x_r)))]^\beta.$$

From condition (Θ_1) , we get

$$\begin{aligned} \theta(d(K^r x_0, K^{r+1} x_0)) &\leq [\theta(\psi(d(K^{r-1} x_0, K^r x_0)))]^\beta \\ &\leq [\theta(\psi(d(K^{r-2} x_0, K^{r-1} x_0)))]^{\beta^2} \\ &\vdots \\ &\leq [\theta(\psi(d(x_0, Kx_0)))]^{\beta^r}. \end{aligned} \tag{13}$$

Letting $r \rightarrow \infty$ in the above inequality, we get

$$\lim_{r \rightarrow \infty} \theta(d(x_r, x_{r+1})) = 1, \tag{14}$$

From (Θ_2) this implies that

$$\lim_{r \rightarrow \infty} d(x_r, x_{r+1}) = 0^+.$$

Using (Θ_3) , there exist $w \in (0,1)$ and $V \in (0, \infty]$ such that

$$\lim_{r \rightarrow \infty} \frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} = V. \tag{15}$$

Suppose that $V < \infty$. In this case, let $S = \frac{V}{2} > 0$. From the definition of the limit, there exists $r_0 \in \mathbb{N}$ such that

$$\left| \frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} - V \right| \leq S, \quad \text{for all } r \geq r_0.$$

Thereupon, we get

$$\frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} \geq V - S = S, \quad \text{for all } r \geq r_0.$$

So, for all $r \geq r_0$, we obtain

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1],$$

where $H = \frac{1}{S}$. Assume that $V = \infty$. Let $S > 0$ be an arbitrary positive number. Thus there exists $r_0 \in \mathbb{N}$ such that

$$\frac{\theta(d(x_r, x_{r+1})) - 1}{(d(x_r, x_{r+1}))^w} \geq S,$$

for all $r \geq r_0$. This implies that for all $r \geq r_0$,

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1],$$

where $H = \frac{1}{S}$. Therefore, in two cases, there exists $H > 0$ and $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$,

$$r(d(x_r, x_{r+1}))^w \leq Hr[\theta(d(x_r, x_{r+1})) - 1].$$

Using (13), we have

$$r(d(x_r, x_{r+1}))^w \leq Hr([\theta(\psi(d(x_0, x_1)))]^{\beta^r} - 1), \tag{16}$$

for all $r \geq r_0$. Letting $r \rightarrow \infty$ in Eqn. (16), we obtain

$$\lim_{r \rightarrow \infty} r(d(x_r, x_{r+1}))^w = 0.$$

Therefore, there exists $r_1 \in \mathbb{N}$ such that

$$d(x_r, x_{r+1}) \leq \frac{1}{r^w}, \text{ for all } r \geq r_1. \tag{17}$$

Next we prove that $\{x_r\}$ is a Cauchy sequence in K . There exist sequences $r, p \in \mathbb{N}$ such that $p > r \geq r_1$. Then from Eqn. (17), we obtain

$$\begin{aligned} d(x_r, x_p) &\leq d(x_r, x_{r+1}) + d(x_{r+1}, x_{r+2}) + \dots + d(x_{p-1}, x_p) \\ &\leq \sum_{i=r}^{p-1} \frac{1}{i^w} \leq \sum_{i=r}^{\infty} \frac{1}{i^w}. \end{aligned}$$

By the convergence of the series $\sum_{i=r}^{\infty} \frac{1}{i^w}$, in the limit $r \rightarrow \infty$, we get $d(x_r, x_p) \rightarrow 0$. Then $\{x_r\}$ is a Cauchy sequence in (X, d) . Since (X, d) is a complete metric space, then there exists $v \in X$ such that

$$\lim_{r \rightarrow \infty} x_r = v. \tag{18}$$

If K is continuous, then we have

$$v = \lim_{r \rightarrow \infty} x_{r+1} = \lim_{r \rightarrow \infty} Kx_r = K \lim_{r \rightarrow \infty} x_r = Kv.$$

Thus $v = Kv$, and, v is a fixed point of K . We given that X is regular, so $x_r \leq v$ for all $r \in \mathbb{N}$. Two conditions arise here:

Case 1. If there exists $u \in \mathbb{N}$ for that $x_u = v$, so

$$Kv = Kx_u = x_{u+1} \leq v.$$

In addition, we have $x_u \leq x_{u+1}$. So, $v \leq Kv$ and thus $v = Kv$.

Case 2. Given that $x_r \neq v$ for all $r \in \mathbb{N}$ and $d(v, Kv) > 0$. From $\lim_{r \rightarrow \infty} x_r = v$, there exists $r_1 \in \mathbb{N}$ such that $d(x_{r+1}, Kv) > 0$ and $d(x_r, v) < \frac{d(v, Kv)}{2}$ for every $r \geq r_1$. In addition, $(x_r, v) \in Q$. Thus, from (Θ_1) and ψ non-decreasing we obtain

$$\begin{aligned} \theta(d(Kx_r, Kv)) &\leq [\theta(\psi(P(x_r, v)))]^\beta \\ &\leq \theta(\psi(\max\{d(x_r, v), d(x_r, Kx_r), d(v, Kv), \\ &\quad \frac{1}{2}[d(x_r, Kv) + d(v, Kx_r)]\})) \\ &\leq \theta(\psi(d(Kv, v))). \end{aligned}$$

This implies that

$$d(x_{r+1}, Kv) \leq \psi(d(Kv, v)) < d(Kv, v).$$

Taking limit as $r \rightarrow \infty$, we obtain

$$d(v, Kv) \leq d(Kv, v),$$

a contraction. Hence, we get $d(v, Kv) = 0$, that is, $v = Kv$ this concludes the proof.

Example 1. Let $X = [0,1] \cup \{2,3\}$ and $d(x, y) = |x - y|$, for all $x, y \in X$. Define an order relation \leq on X , where \leq is usual order. (X, \leq, d) is complete and define a mapping $K: X \rightarrow X$ by

$$K(x) = \begin{cases} \frac{x}{6}, & x \in [0,1] \\ x, & x \in \{2,3\} \end{cases}$$

Then, K is non-decreasing. We claim that K is an ordered weakly θ -contractive with $\theta(p) = e^{pe^p}$, $\psi(s) = \frac{s}{2}$ and $\beta = e^{-\frac{x}{3}}$. Therefore, Theorem 1 and Theorem 2 are satisfied.

3. Conclusion

Agarwal et. al., [4] defined weak generalized contractions in partially ordered metric spaces. Ćirić [12] denoted Ćirić contraction and Jleli and Samet [10] defined θ -contraction in metric spaces. We introduce new contraction mappings by combining the ideas of Agarwal et al., Ćirić, Jleli and Samet. In addition, we present an example to show that the new theorems are applicable.

Acknowledgements

The author would like to thank the referees for their careful reading of the manuscript and valuable suggestions.

References

- [1] Banach, S., *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, *Fundamenta Mathematicae*, 3(1), 133-181, 1922.
- [2] Ran, A.C., Reurings, M.C., *A fixed point theorem in partially ordered sets and some applications to matrix equations*, *Proceedings of the American Mathematical Society*, 1435-1443, 2004.
- [3] Abbas, M., Nazir, T., Radenović, S., *Common fixed points of four maps in partially ordered metric spaces*, *Applied Mathematics Letters*, 24(9), 1520-1526, 2011.
- [4] Agarwal, R.P., El-Gebeily, M.A., O'Regan, D., *Generalized contractions in partially ordered metric spaces*, *Applicable Analysis*, 87(1), 109-116, 2008.
- [5] Yeşilkaya, S.S., *Some fixed point theorems in ordered metric spaces having t -property*, *Tbilisi Mathematical Journal*, 8-2021, 219–226, 2021.
- [6] Kumam, P., Rouzkard, F., Imdad, M., Gopal, D., *Fixed point theorems on ordered metric spaces through a rational contraction*, *Abstract and Applied Analysis*, 2013, 2013.
- [7] Durmaz, G., Minak, G., Altun, I., *Fixed points of ordered F -contractions*, *Hacettepe Journal of Mathematics and Statistics*, 45(1), 15-21, 2016.
- [8] Minak, G., Altun, I., *Ordered θ -contractions and some fixed point results*, *Nonlinear Functional Analysis*, 41, 2017.
- [9] Yeşilkaya, S.S., *Presic Type Operators for a Pair Mappings*, *Turkish Journal of Mathematics and Computer Science*, 13(1), 204-210, 2021.
- [10] Jleli, M., Samet, B., *A new generalization of the Banach contraction principle*, *Journal of Inequalities and Applications*, 2014(1), 1-8, 2014.
- [11] Wardowski, D., *Fixed points of a new type of contractive mappings in complete metric Spaces*, *Fixed Point Theory and Applications*, 2012(1), 1-6, 2012.
- [12] Ćirić, L.B., *Generalized contractions and fixed-point theorems*, *Publications de l'Institut Mathématique*, 12(26), 19-26, 1971.
- [13] Razavi, S.S., Parvaneh Masiha, H., *Generalized f -contractions in partially ordered metric spaces*, *Sahand Communications in Mathematical Analysis*, 16(1), 93-104, 2019.



A New Approach to Fibonacci Tessarines with Fibonacci and Lucas Number

Components

Faik BABADAĞ^{1,*}, Merve USLU²

¹Kırıkkale University, Faculty of Arts and Science, Department of Mathematics, 71450, Kırıkkale, Türkiye
faik.babadag@kku.edu.tr, ORCID: 0000-0001-9098-838X

²Kırıkkale University, Faculty of Arts and Science, Department of Mathematics, 71450, Kırıkkale, Türkiye
merveuslu124@gmail.com, ORCID: 0000-0003-2464-8736

Received: 11.01.2021

Accepted: 10.09.2021

Available online: 31.12.2021

Abstract

In this paper, by using identities related to the tessarines, Fibonacci numbers and Lucas numbers we define Fibonacci tessarines and Lucas tessarines. We obtain Binet formulae, D'ocagnes identity and Cassini identity for these tessarines. We also give the identities of Fibonacci negatessarines and Lucas negatessarines and define new vector which are called Fibonacci tessarine vector.

Keywords: Tessarines; Fibonacci numbers; Fibonacci tessarine vector.

Fibonacci ve Lucas Sayı Bileşenleri ile Fibonacci Tessarinelere Yeni Bir Yaklaşım

Öz

Bu makalede, tessarineler, Fibonacci ve Lucas sayılarıyla ilgili özdeşlikleri kullanarak Fibonacci tessarineler ve Lucas tessarineleri tanımladık. Bu tessarineler için Binet formüllerini, D'ocagnes özdeşliğini ve Cassini özdeşliğini elde ettik. Ayrıca, negatif Fibonacci tessarineler ve negatif Lucas tessarinelerin özdeşliklerini verdik ve Fibonacci tessarine vektörü olarak yeni bir vektör tanımladık.



Anahtar Kelimeler: Tessarineler; Fibonacci sayıları; Fibonacci tessarine vektör.

1. Introduction

A tessarine can be defined as follows [1-3];

$$\mathcal{T} = t_1 + t_2 i + t_3 j + t_4 k,$$

where t_{1-4} are real numbers and $+1, i, j, k$ are governed by relations

$$i^2 = -1, j^2 = +1, ij = ji = k, jk = i, ki = -j.$$

Let \mathcal{T} and \mathcal{T}' be tessarines. The addition, subtraction and multiplication of these numbers are presented as follows

$$\mathcal{T} \mp \mathcal{T}' = (t_1 \mp t'_1) + (t_2 \mp t'_2)i + (t_3 \mp t'_3)j + (t_4 \mp t'_4)k$$

and

$$\begin{aligned} \mathcal{T} \times \mathcal{T}' &= (t_1 + t_2 i + t_3 j + t_4 k) \times (t'_1 + t'_2 i + t'_3 j + t'_4 k) \\ &= (t_1 t'_1 - t_2 t'_2 + t_3 t'_3 - t_4 t'_4) + (t_1 t'_2 + t_2 t'_1 + t_3 t'_4 + t_4 t'_3) i \\ &\quad + (t_1 t'_3 + t_3 t'_1 - t_2 t'_4 - t_4 t'_2) j + (t_1 t'_4 + t_4 t'_1 + t_3 t'_2 + t_2 t'_3) k. \end{aligned}$$

The conjugates of a tessarine are described by $\mathcal{T}^i, \mathcal{T}^j$ and \mathcal{T}^k . In that case, there are different conjugations as follows:

$$\mathcal{T}^i = t_1 - t_2 i + t_3 j - t_4 k,$$

$$\mathcal{T}^j = t_1 + t_2 i - t_3 j - t_4 k,$$

$$\mathcal{T}^k = t_1 - t_2 i - t_3 j + t_4 k.$$

Then, the following equalities are written

$$\mathcal{T} \times \mathcal{T}^i = t_1^2 + t_2^2 + t_3^2 + t_4^2 + 2j(t_1 t_3 + t_2 t_4),$$

$$\mathcal{T} \times \mathcal{T}^j = t_1^2 - t_2^2 - t_3^2 + t_4^2 + 2i(t_1 t_2 - t_3 t_4),$$

$$\mathcal{T} \times \mathcal{T}^k = t_1^2 + t_2^2 - t_3^2 - t_4^2 + 2k(t_1 t_4 - t_2 t_3).$$

Fibonacci numbers and Lucas numbers are defined in many studies. The relations between these numbers are given and computed in [4-8]. Horadam defined the generalized Fibonacci sequences [9].

In the present paper, we introduce and study the Fibonacci tessarines and Lucas tessarines by using some properties of Fibonacci and Lucas numbers, and we obtain some identities for

them. The Fibonacci numbers f_n are defined for all integers n by the second order recurrence relation

$$f_{n+2} = f_{n+1} + f_n$$

and initial conditions $f_1 = f_2 = 1$. The Lucas numbers l_n are defined for all integers n by the same second order recurrence relation as

$$l_{n+2} = l_{n+1} + l_n,$$

but initial conditions $l_1 = 1, l_2 = 3$. The Binet formulae for the Fibonacci and Lucas numbers are as follows [10]

$$f_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \text{ and } l_n = \alpha^n + \beta^n,$$

where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

These facts are well-known and can be found in most basic references, e.g. [11, 12]. In this paper, we need some of them as follows,

$$f_{n+2}f_{n-1} = f_{n+1}^2 - f_n^2, \tag{1}$$

$$f_{n-1}f_{n+1} - f_n^2 = (-1)^n, \tag{2}$$

$$f_{2n+1} = f_n^2 + f_{n+1}^2, \tag{3}$$

$$l_n = f_{n-1} + f_{n+1}, \tag{4}$$

$$l_n = f_{n+2} - f_{n-2}, \tag{5}$$

$$l_{-n} = (-1)^n l_n, \tag{6}$$

$$5f_n f_m = l_{n+m} - (-1)^m l_{n-m}, \tag{7}$$

$$f_m f_{n+1} - f_{m+1} f_n = (-1)^n f_{m-n}, \tag{8}$$

$$f_{-n} = (-1)^{n+1} f_n, \tag{9}$$

$$l_{n-1} + l_{n+1} = 5f_n. \tag{10}$$

2. Fibonacci Tessarines and Lucas Tessarines

Definition 1. The n^{th} Fibonacci tessarine \mathcal{T}_n and the n^{th} Lucas tessarine \mathcal{T}'_n are defined by, respectively,

$$\mathcal{T}_n = f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k \tag{11}$$

and

$$\mathcal{T}'_n = l_n + l_{n+1}i + l_{n+2}j + l_{n+3}k, \tag{12}$$

where f_n is the n^{th} Fibonacci number and l_n is the n^{th} Lucas number. Also i, j and k are arbitrary units which satisfy the relations;

$$i^2 = -1, j^2 = +1, k^2 = -1 \text{ and } ij = ji = k. \tag{13}$$

Starting from $n = 0$, the Fibonacci tessarines and Lucas tessarines can be written respectively as;

$$\mathcal{T}_0 = 1j + 2k; \mathcal{T}_1 = 1 + 1i + 2j + 3k; \mathcal{T}_2 = 1 + 2i + 3j + 5k, \dots$$

$$\mathcal{T}'_0 = 2 + 1i + 3j + 4k; \mathcal{T}'_1 = 1 + 3i + 4j + 7k; \mathcal{T}'_2 = 3 + 4i + 7j + 11k, \dots$$

Now, let $\mathcal{T}_n = f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k$ and $\mathcal{T}_m = f_m + f_{m+1}i + f_{m+2}j + f_{m+3}k$ be Fibonacci tessarines. Then we have

$$\mathcal{T}_n \mp \mathcal{T}_m = (f_n + f_m) + (f_{n+1} + f_{m+1})i + (f_{n+2} + f_{m+2})j + (f_{n+3} + f_{m+3})k$$

and

$$\begin{aligned} \mathcal{T}_n \times \mathcal{T}_m = & (f_n f_m - f_{n+1} f_{m+1} + f_{n+2} f_{m+2} - f_{n+3} f_{m+3}) \\ & + (f_n f_{m+1} + f_{n+1} f_m + f_{n+2} f_{m+3} + f_{n+3} f_{m+2})i \\ & + (f_n f_{m+2} + f_{n+2} f_m - f_{n+1} f_{m+3} - f_{n+3} f_{m+1})j \\ & + (f_n f_{m+3} + f_{n+3} f_m + f_{n+2} f_{m+1} + f_{n+1} f_{m+2})k. \end{aligned}$$

Definition 2. Let \mathcal{T}_n be a Fibonacci tessarine. For $n \geq 1$, there are three different conjugations with respect to i, j and k ;

$$\mathcal{T}_n^i = f_n - f_{n+1}i + f_{n+2}j - f_{n+3}k,$$

$$\mathcal{T}_n^j = f_n + f_{n+1}i - f_{n+2}j - f_{n+3}k,$$

$$\mathcal{T}_n^k = f_n - f_{n+1}i - f_{n+2}j + f_{n+3}k.$$

Now by using definition 2 and the Eqn. (11), we can obtain

$$\mathcal{T}_n \times \mathcal{T}_n^i = (3 + 2j)f_{2n+3},$$

$$\mathcal{T}_n \times \mathcal{T}_n^j = (1 - 2i)f_{2n+3},$$

$$\mathcal{T}_n \times \mathcal{T}_n^k = -l_{2n+3} + (-1)^{n+1}k.$$

3. Some Identities on Fibonacci Tessarines and Lucas Tessarines

3.1. Identities

Let $n \geq 1$ be an integer. Then, we can give the following relations between Fibonacci tessarines

$$\mathcal{T}_n - \mathcal{T}_{n+1}i + \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k = (3 + 2j)l_{n+3},$$

$$\mathcal{T}_n + \mathcal{T}_{n+1}i - \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k = (1 - 2i)l_{n+3},$$

$$\mathcal{T}_n - \mathcal{T}_{n+1}i - \mathcal{T}_{n+2}j + \mathcal{T}_{n+3}k = -5f_{n+3}.$$

Proof. Now, we will give proof of identity $\mathcal{T}_n - \mathcal{T}_{n+1}i + \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k$. We have

$$\begin{aligned} \mathcal{T}_n - \mathcal{T}_{n+1}i + \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k &= (f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k) \\ &\quad - (f_{n+1} + f_{n+2}i + f_{n+3}j + f_{n+4}k)i \\ &\quad + (f_{n+2} + f_{n+3}i + f_{n+4}j + f_{n+5}k)j \\ &\quad - (f_{n+3} + f_{n+4}i + f_{n+5}j + f_{n+6}k)k \\ &= (f_n + f_{n+2}) + (f_{n+4} + f_{n+6}) + 2j(f_{n+2} + f_{n+4}). \end{aligned}$$

Here by using the Eqn. (4) and doing necessary calculations, we obtain

$$\begin{aligned} \mathcal{T}_n - \mathcal{T}_{n+1}i + \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k &= 2l_{n+3} + l_{n+3} + 2jl_{n+3} = 3l_{n+3} + 2jl_{n+3} \\ &= (3 + 2j)l_{n+3}. \end{aligned}$$

For all n , we compute the following identities with similar method.

$$\mathcal{T}_n + \mathcal{T}_{n+1}i - \mathcal{T}_{n+2}j - \mathcal{T}_{n+3}k = (1 - 2i)l_{n+3},$$

$$\mathcal{T}_n - \mathcal{T}_{n+1}i - \mathcal{T}_{n+2}j + \mathcal{T}_{n+3}k = -5f_{n+3}.$$

3.2. Identity

Let \mathcal{T}_n^k be the conjugation of Fibonacci tessarine with respect to the imaginary unit k . Then, for $n \geq 1$, we have

$$\mathcal{T}_n \mathcal{T}_n^k + \mathcal{T}_{n-1} \mathcal{T}_{n-1}^k = -5f_{2n+2}.$$

Proof. From the Eqn. (2) and Eqn. (3), we have

$$\begin{aligned} \mathcal{T}_n \mathcal{T}_n^k + \mathcal{T}_{n-1} \mathcal{T}_{n-1}^k &= f_n^2 + f_{n+1}^2 - (f_{n+2}^2 + f_{n+3}^2) + 2k(f_{n+3}f_n - f_{n+1}f_{n+2}) \\ &\quad + f_{n-1}^2 + f_n^2 - (f_{n+1}^2 + f_{n+2}^2) + 2k(f_{n+2}f_{n-1} - f_n f_{n+1}) \\ &= f_{2n+1} - f_{2n+5} + 2k((f_{n+2} + f_{n+1})f_n - f_{n+1}(f_{n+1} + f_n)) \\ &\quad + f_{2n-1} - f_{2n+3} + 2k((f_{n+1} + f_n)f_{n-1} - f_n(f_n + f_{n-1})) \\ &= -(f_{2n+5} - f_{2n+1}) + 2k(f_{n+2}f_n - f_{n+1}^2) - (f_{2n+3} - f_{2n-1}) \\ &\quad + 2k(f_{n+1}f_{n-1} - f_n^2) \\ &= -5f_{2n+2}. \end{aligned}$$

3.3. Identities (Binet Formulae)

Let \mathcal{T}_n and \mathcal{T}'_n be Fibonacci tessarine and Lucas tessarine, respectively. For $n \geq 1$, the Binet formulae for these numbers are given as follows;

$$\mathcal{T}_n = \frac{\bar{\alpha}\alpha^n - \bar{\beta}\beta^n}{\alpha - \beta}$$

and

$$\mathcal{T}'_n = \bar{\alpha}\alpha^n + \bar{\beta}\beta^n,$$

where

$$\bar{\alpha} = 1 + i\alpha + j\alpha^2 + k\alpha^3 \text{ and } \bar{\beta} = 1 + i\beta + j\beta^2 + k\beta^3.$$

Proof. By taking $\bar{\alpha} = 1 + i\alpha + j\alpha^2 + k\alpha^3$ and $\bar{\beta} = 1 + i\beta + j\beta^2 + k\beta^3$ and using the Binet formulae for Fibonacci and Lucas numbers, we obtain

$$\mathcal{T}_n = f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k$$

$$\begin{aligned} &= \frac{\alpha^n - \beta^n}{\alpha - \beta} + \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}i + \frac{\alpha^{n+2} - \beta^{n+2}}{\alpha - \beta}j + \frac{\alpha^{n+3} - \beta^{n+3}}{\alpha - \beta}k \\ &= \frac{(\alpha^n(1 + i\alpha + j\alpha^2 + k\alpha^3) - \beta^n(1 + i\beta + j\beta^2 + k\beta^3))}{\alpha - \beta} \\ &= \frac{\bar{\alpha}\alpha^n - \bar{\beta}\beta^n}{\alpha - \beta} \end{aligned}$$

and

$$\begin{aligned} \mathcal{T}'_n &= l_n + l_{n+1}i + l_{n+2}j + l_{n+3}k \\ &= \alpha^n + \beta^n + (\alpha^{n+1} + \beta^{n+1})i + (\alpha^{n+2} + \beta^{n+2})j + (\alpha^{n+3} + \beta^{n+3})k \\ &= \alpha(1 + i\alpha + j\alpha^2 + k\alpha^3) + \beta(1 + i\beta + j\beta^2 + k\beta^3) \\ &= \bar{\alpha}\alpha^n + \bar{\beta}\beta^n. \end{aligned}$$

3.4. Identity (D’ocagnes identity)

Let $m, n \geq 0$ be integers. Then, the D’ocagnes identity for Fibonacci tessarine is given by

$$\mathcal{T}_m\mathcal{T}_{n+1} - \mathcal{T}_{m+1}\mathcal{T}_n = (-1)^{m+1} (\mathcal{T}_{m-n} - 5f_{m-n} - f_{m-n+1}i).$$

Proof. If we use Eqn. (5), Eqn. (8) and Eqn. (9), we have

$$\begin{aligned} \mathcal{T}_m\mathcal{T}_{n+1} - \mathcal{T}_{m+1}\mathcal{T}_n &= (f_m + f_{m+1}i + f_{m+2}j + f_{m+3}k)(f_{n+1} + f_{n+2}i + f_{n+3}j + f_{n+4}k) \\ &\quad - (f_{m+1} + f_{m+2}i + f_{m+3}j + f_{m+4}k)(f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k) \\ &= -4(-1)^{m+1} f_{m-n} + (-1)^{m+1} f_{m-n+2}j + (-1)^{m+1} f_{m-n+3}k \\ &= (-1)^{m+1} [(f_{m-n} + f_{m-n+1}i + f_{m-n+2}j + f_{m-n+3}k) \\ &\quad - (5f_{m-n} + f_{m-n+1}i)] \\ &= (-1)^{m+1} (\mathcal{T}_{m-n} - 5f_{m-n} - f_{m-n+1}i). \end{aligned}$$

3.5. Identity (Cassini’s identity)

Let $n \geq 0$ be an integer. The Cassini’s identity is given as;

$$\mathcal{T}_{n+1}\mathcal{T}_{n-1} - \mathcal{T}_n^2 = (-1)^n (2\mathcal{T}_1 + 2 + 2j - 3k).$$

Proof. We have

$$\begin{aligned} \mathcal{T}_{n+1}\mathcal{T}_{n-1} - \mathcal{T}_n^2 &= (f_{n+1} + f_{n+2}i + f_{n+3}j + f_{n+4}k)(f_{n-1} + f_ni + f_{n+1}j + f_{n+2}k) \end{aligned}$$

$$\begin{aligned}
 & -(f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k)^2 \\
 = & f_n^2 - f_{n-1}f_{n+1} - (f_{n+1}^2 + f_n f_{n+2}) \\
 & -(f_{n+2}^2 - f_{n+1}f_{n+3}) + (f_{n+2}f_{n+4} - f_{n+3}^2) \\
 & + i(f_{n-1}f_{n+2} - f_n f_{n+1} + f_{n+1}f_{n+4} - f_{n+2}f_{n+3}) \\
 & + j(f_{n+1}^2 - f_n f_{n+2} - f_{n+2}^2 + f_{n-1}f_{n+3} - f_{n+1}f_{n+3} + f_{n+1}f_{n+3} - f_n f_{n+2}) \\
 & + j(f_{n+1}^2 - f_n f_{n+2} - f_{n+2}^2 + f_{n-1}f_{n+3} - f_{n+1}f_{n+3} + f_{n+1}f_{n+3} - f_n f_{n+2}) \\
 & + k(f_{n-1}f_{n+4} - f_n f_{n+3})
 \end{aligned}$$

which implies

$$\begin{aligned}
 & \mathcal{T}_{n+1}\mathcal{T}_{n-1} - \mathcal{T}_n^2 \\
 = & 4(-1)^n + i((-1)^{n+1}f_{-2} + (-1)^{n+3}f_{-2}) + j(2(-1)^n \\
 & + (-1)^{n+2}f_{-3} + (-1)^n f_3) + k(-1)^{n+3}f_{-4} \\
 = & 4(-1)^n + 2i(-1)^n f_2 + j(2(-1)^n + (-1)^n f_3 + (-1)^n f_3) + k(-1)^n f_4 \\
 = & (-1)^n(4 + 2i + 6j + 3k).
 \end{aligned}$$

Also by adding and subtracting the term $3k$, we complete the proof.

3.6. Identities (Fibonacci negatessarine and Lucas negatessarine)

Let \mathcal{T}_n and \mathcal{T}'_n be Fibonacci tessarine and Lucas tessarine, respectively. The identities of Fibonacci negatessarine and Lucas negatessarine are defined as;

$$\mathcal{T}_{-n} = (-1)^{n+1}\mathcal{T}_n + (-1)^n l_n(i + j + 2k)$$

and

$$\mathcal{T}'_{-n} = (-1)^n \mathcal{T}'_n + (-1)^{n+1} + 5f_n(i + j + 2k).$$

Proof. Now, we will give proof of identity \mathcal{T}_{-n} . We have

$$\begin{aligned}
 \mathcal{T}_{-n} &= f_{-n} + f_{-n+1}i + f_{-n+2}j + f_{-n+3}k \\
 &= f_{-n} + f_{-(n-1)}i + f_{-(n-2)}j + f_{-(n-3)}k \\
 &= (-1)^{n+1}f_n + (-1)^n f_{n-1}i + (-1)^{n+1}f_{n-2}j + (-1)^n f_{n-3}k \\
 &= (-1)^{n+1}(f_n + f_{n+1}i + f_{n+2}j + f_{n+3}k) - (-1)^{n+1}f_{n+1}i - (-1)^{n+1}f_{n+2}j \\
 &\quad - (-1)^{n+1}f_{n+3}k + (-1)^n f_{n-1}i + (-1)^{n+1}f_{n-2}j + (-1)^n f_{n-3}k
 \end{aligned}$$

In this identity, taking into account the Eqn. (1), Eqn. (5) and Eqn. (9), we obtain

$$\mathcal{T}_{-n} = (-1)^{n+1}\mathcal{T}_{-n} + (-1)^n(f_{n+1} + f_{n-1})i + (-1)^n(f_{n+2} - f_{n-2})j$$

$$\begin{aligned}
 &+(-1)^n(f_{n+3} + f_{n-3})k \\
 &= (-1)^{n+1}\mathcal{T}_n + (-1)^n l_n(i + j + 2k).
 \end{aligned}$$

Now, if we use the Eqn. (6), Eqn. (7) and Eqn. (10), we have

$$\begin{aligned}
 \mathcal{T}'_{-n} &= l_{-n} + l_{-n+1}i + l_{-n+2}j + l_{-n+3}k \\
 &= l_{-n} + l_{-(n-1)}i + l_{-(n-2)}j + l_{-(n-3)}k \\
 &= (-1)^n l_n + (-1)^{n-1}l_{n-1}i + (-1)^n l_{n-2}j + (-1)^{n-1}l_{n-3}k \\
 &= (-1)^n(l_n + l_{n+1}i + l_{n+2}j + l_{n+3}k) - (-1)^n l_{n+1}i - (-1)^n l_{n+2}j \\
 &\quad - (-1)^n l_{n+3}k + (-1)^{n-1}l_{n-1}i + (-1)^n l_{n-2}j + (-1)^{n-1}l_{n-3}k \\
 &= (-1)^n \mathcal{T}'_n + (-1)^{n+1}((l_{n+1} + l_{n-1})i + (l_{n+2} - l_{n-2})j + (l_{n+3} + l_{n-3})k) \\
 &= (-1)^n \mathcal{T}'_n + (-1)^{n+1} + 5f_n i + (-1)^{n+1} + 5f_n j + (-1)^{n+1} + 10f_n k,
 \end{aligned}$$

which completes the proof.

Example 1. Let $\mathcal{T}_0, \mathcal{T}_1$ and \mathcal{T}_2 be the Fibonacci tessarines such that

$$\mathcal{T}_0 = i + j + 2k,$$

$$\mathcal{T}_1 = 1 + i + 2j + 3k,$$

$$\mathcal{T}_2 = 1 + 2i + 3j + 5k.$$

Considering the Eqn. (11) and Eqn. (13), we have

$$\begin{aligned}
 \mathcal{T}_0 \mathcal{T}_2 - \mathcal{T}_1^2 &= (i + j + 2k)(1 + 2i + 3j + 5k) - (1 + i + 2j + 3k)^2 \\
 &= -(4 + 2i + 6j + 3k) = -(2\mathcal{T}_1 + 2 + 2j - 3k)
 \end{aligned}$$

and

$$\begin{aligned}
 \mathcal{T}_1 \mathcal{T}_3 - \mathcal{T}_2^2 &= (1 + i + 2j + 3k)(2 + 3i + 5j + 8k) - (1 + 2i + 3j + 5k)^2 \\
 &= 4 + 2i + 6j + 3k = 2\mathcal{T}_1 + 2 + 2j - 3k.
 \end{aligned}$$

4. Some Applications on the Fibonacci Tessarines

Definition 3. The n^{th} Fibonacci tessarine vector $\vec{\mathcal{T}}_n$ and the n^{th} Lucas tessarine vector $\vec{\mathcal{T}}'_n$ are defined as

$$\vec{\mathcal{T}}_n = f_{n+1}i + f_{n+2}j + f_{n+3}k$$

and

$$\vec{J}'_n = l_{n+1}i + l_{n+2}j + l_{n+3}k,$$

where f_n is the n^{th} Fibonacci number and l_n is the n^{th} Lucas number. Also i, j and k are arbitrary units which satisfy the relations;

$$i^2 = -1, j^2 = +1, k^2 = -1 \text{ and } ij = ji = k.$$

Definition 4. [13] Let \vec{J}_n and \vec{J}_m be Fibonacci tessarine vectors. The dot product and the cross product of these vectors are defined by

$$\langle \vec{J}_n, \vec{J}_m \rangle = f_{n+1}f_{m+1} + f_{n+2}f_{m+2} + f_{n+3}f_{m+3}$$

and

$$\begin{aligned} \vec{J}_n \times \vec{J}_m &= \det \begin{bmatrix} i & j & k \\ f_{n+1} & f_{n+2} & f_{n+3} \\ f_{m+1} & f_{m+2} & f_{m+3} \end{bmatrix} \\ &= i(f_{n+2}f_{m+3} + f_{m+2}f_{n+3}) - j(f_{n+1}f_{m+3} + f_{m+1}f_{n+3}) \\ &\quad + k(f_{n+1}f_{m+2} + f_{m+1}f_{n+2}). \end{aligned}$$

Some examples of Fibonacci tessarine and Lucas tessarine vectors can be given easily as;

$$\vec{J}_0 = 1i + 1j + 2k, \vec{J}_1 = 1i + 2j + 3k \text{ and } \vec{J}_2 = 2i + 3j + 5k$$

and

$$\vec{J}'_0 = 1i + 3j + 4k, \vec{J}'_1 = 3i + 4j + 7k \text{ and } \vec{J}'_2 = 4i + 7j + 11k.$$

Theorem 1. Let \vec{J}_n and \vec{J}_{n+1} be Fibonacci tessarine vectors. The dot product and the cross product of these vectors are as follows

$$\langle \vec{J}_n, \vec{J}_{n+1} \rangle = \frac{4}{5}l_{2n+5} + \frac{(-1)^n}{5}$$

and

$$\vec{J}_n \times \vec{J}_{n+1} = \vec{J}'_{2n+1} - 2f_{2n+4}(i + j) + \frac{(-1)^n}{5}(i + 2j - 2k).$$

Proof. Using the Eqn. (7) and Eqn. (10), we obtain,

$$\begin{aligned} \langle \vec{J}_n, \vec{J}_{n+1} \rangle &= f_{n+1}f_{n+2} + f_{n+2}f_{n+3} + f_{n+3}f_{n+4} \\ &= \frac{1}{5}(l_{2n+3} - (-1)^{n+1}l_1 + l_{2n+5} - (-1)^{n+2}l_1 + l_{2n+7} - (-1)^{n+3}l_1) \\ &= \frac{4}{5}l_{2n+5} + \frac{(-1)^n}{5} \end{aligned}$$

and the cross product of \vec{J}_n and \vec{J}_{n+1} is given by

$$\vec{T}_n \times \vec{T}_{n+1} = i(f_{n+2}f_{n+4} + f_{n+3}^2) - j(f_{n+1}f_{n+4} + f_{n+2}f_{n+3}) + k(f_{n+1}f_{n+3} + f_{n+2}^2)$$

Finally, from Eqn. (7), we have

$$\begin{aligned} \vec{T}_n \times \vec{T}_{n+1} &= \frac{2}{5}(l_{2n+6}i - l_{2n+5}j + l_{2n+4}k) - \frac{1}{5}(-1)^n(i + 2j - 2k) \\ &= \frac{2}{5}[(l_{2n+6} + l_{2n+2} - l_{2n+2})i + (-l_{2n+5} + l_{2n+3} - l_{2n+3})j + l_{2n+4}k] \\ &\quad + \frac{(-1)^n}{5}(i + 2j - 2k) \\ &= \vec{T}'_{2n+1} - 2f_{2n+4}(i + j) + \frac{(-1)^n}{5}(i + 2j - 2k), \end{aligned}$$

where \vec{T}'_{2n+1} is a Lucas tessarine vector.

Example 2. Let \vec{T}_1 and \vec{T}_2 be Fibonacci tessarine vectors such that $\vec{T}_1 = (1, 2, 3)$,

$\vec{T}_2 = (2, 3, 5)$. The dot product and cross product of these vectors are

$$\begin{aligned} \langle \vec{T}_1, \vec{T}_2 \rangle &= f_2f_3 + f_3f_4 + f_4f_5 \\ &= \frac{4}{5}l_7 + \frac{(-1)^1}{5} = 23 \end{aligned}$$

and

$$\vec{T}_1 \times \vec{T}_2 = \vec{T}'_3 - 2if_6(i + j) - \frac{1}{5}(i - 2j - 2k).$$

Example 3. Let \vec{T}_{n-1} , \vec{T}_n and \vec{T}_{n+1} be Fibonacci tessarine vectors. For $n \geq 1$, the Cassini identity for Fibonacci tessarine vectors is given by

$$\vec{T}_{n-1} \vec{T}_{n+1} - \vec{T}_n^2 = (-1)^n \left(\vec{T}'_0 + 3 - \frac{4}{5}k \right).$$

Proof. Now, we will give proof of identities \vec{T}_{n-1} , \vec{T}_{n+1} and \vec{T}_n^2 , we obtain

$$\begin{aligned} \vec{T}_{n-1} \vec{T}_{n+1} &= (f_n i + f_{n+1} j + f_{n+2} k)(f_{n+2} i + f_{n+3} j + f_{n+4} k) \\ &= -f_n f_{n+2} + f_{n+1} f_{n+3} - f_{n+2} f_{n+4} + i(f_{n+1} f_{n+4} + f_{n+2} f_{n+3}) \\ &\quad - j(f_n f_{n+4} + f_{n+2}^2) + k(f_{n+3} f_n + f_{n+1} f_{n+2}) \end{aligned}$$

and

$$\begin{aligned} \vec{J}_n^2 &= (f_{n+1}i + f_{n+2}j + f_{n+3}k)(f_{n+1}i + f_{n+2}j + f_{n+3}k) \\ &= -f_{n+1}^2 + f_{n+2}^2 - f_{n+3}^2 + 2if_{n+2}f_{n+3} - 2jf_{n+1}f_{n+3} + 2kf_{n+1}f_{n+2}. \end{aligned}$$

From the Eqn. (3) and Eqn. (7), doing necessary calculations, we have

$$\begin{aligned} \vec{J}_{n-1} \vec{J}_{n+1} - \vec{J}_n^2 &= (-1)^n \left[\left(3 + i + 3j - \frac{4}{5}k \right) - 4k + 4k \right] \\ &= (-1)^n \left(\vec{J}'_0 + 3 - \frac{4}{5}k \right). \end{aligned}$$

5. Conclusion

In this paper, we first describe the concepts of Fibonacci tessarine and Lucas tessarine with coefficients from the Fibonacci and Lucas numbers. We introduce many identities that have an important place in the literature on tessarines.

Acknowledgement

The authors would like to thank the referees for their significant suggestions and comments, which improve the paper.

References

- [1] Cockle, J., *On certain functions resembling quaternions and on a new imaginary in algebra*, Philosophical Magazine and Journal of Science, London-Dublin-Edinburgh, 3(33), 435-439, 1848.
- [2] Cockle, J., *On a new imaginary in algebra*, Philosophical Magazine and Journal of Science, London-Dublin-Edinburgh, 3(34), 37-47, 1849.
- [3] Cockle, J., *On the symbols of algebra and on the theory of tessarines*, Philosophical magazine and Journal of Science, London-Dublin-Edinburgh, 3(34), 406-410, 1849.
- [4] Dunlap, R.A., *The golden ratio and Fibonacci numbers*, World Scientific, 35-50, 1997.
- [5] Vajda, S., *Fibonacci and Lucas numbers and the golden section*, Theory and Application (Mathematics & Its Applications), Dover Publications, Inc, Mineola, New York, 1989.
- [6] Verner, E., Hoggatt, Jr., *Fibonacci and Lucas Numbers*, Canadian Mathematical Bulletin, 3(12), 367, 1969.
- [7] Iyer, M.R., *Some results on Fibonacci quaternions*, The Fibonacci Quarterly, 7(2), 201-210, 1969.

[8] Knuth, D., *Negafibonacci numbers and hyperbolic plane*, Annual Meeting of the Mathematical Association of America, 2013.

[9] Horadam, A.F., *A generalized Fibonacci sequence*, American Mathematical Monthly, 68, 1961.

[10] Koshy, T., *Fibonacci and Lucas numbers with applications*, A Wiley-Interscience Publication, USA, 2001.

[11] Lucas, E., *Theorie des Numbers*, Paris, 1961.

[12] Iyer, M.R., *Identities involving generalized Fibonacci numbers*, The Fibonacci Quarterly, 7, 66-72, 1969.

[13] David, G.G., *Permanent of a square matrix*, European Journal of Combinatorics, 31, 1881-1891, 2010.



The Modified Exp $(-\vartheta(\sigma))$ -Expansion Function Method for Exact Solutions of the Simplified MCH Equation and the Getmanou Equation

Şeyma TULUCE DEMIRAY^{1,*}, Sevgi KASTAL²

¹*Osmaniye Korkut Ata University, Department of Mathematics, 80000, Osmaniye, Türkiye
seymatuluçe@gmail.com, ORCID: 0000-0002-8027-7290*

²*Osmaniye Korkut Ata University, Department of Mathematics, 80000, Osmaniye, Türkiye
kastalsevgi@gmail.com, ORCID: 0000-0001-9152-3023*

Received: 10.01.2021

Accepted: 10.09.2021

Published: 31.12.2021

Abstract

The present research explores, the modified exp $(-\vartheta(\sigma))$ -expansion function method (MEFM) is tested by applying it to obtain the exact travelling wave solutions for the simplified MCH equation and the Getmanou equation. Dark optical soliton solutions and dark-bright optical soliton solutions of the simplified MCH equation and the Getmanou equation are successfully constructed by using this method. We carry out all the computations and draw the 2D and 3D graph in this paper by Wolfram Mathematica 9. Besides, the graphical representation obviously shows forcefulness of this method.

Keywords: The simplified MCH equation; The Getmanou equation; MEFM; Dark optical soliton solutions; Dark-bright optical soliton solutions; Mathematica.

Basitleştirilmiş MCH Denklemi ve Getmanou Denklemine Tam Çözümleri için Modifiye Edilmiş Exp $(-\vartheta(\sigma))$ -Açılım Fonksiyon Metodu

Öz



Mevcut araştırma, Modifiye edilmiş $\exp(-\vartheta(\sigma))$ -açılım fonksiyon metodunu (MEFM), basitleştirilmiş MCH denklemi ve Getmanou denklemi için tam hareketli dalga çözümlerini elde etmek üzere uygulayarak test eder. Basitleştirilmiş MCH denkleminin ve Getmanou denkleminin dark optik soliton ve dark-bright optik soliton çözümleri bu yöntem kullanılarak başarıyla elde edilmiştir. Bu çalışmadaki tüm hesaplamalar Wolfram Mathematica 9 tarafından yapılmış ve 2D ve 3D grafikleri çizilmiştir. Ayrıca, grafiksel gösterim bu yöntemin açıkça gücünü göstermektedir.

Anahtar Kelimeler: Basitleştirilmiş MCH denklemi; Getmanou denklemi; MEFM; Dark optik soliton çözümler; Dark-bright optik soliton çözümler; Mathematica.

1. Introduction

Many phenomena problems in the real world in applied and engineering sciences are structured by nonlinear evolution equations (NLEEs). In recent years, nonlinear evolution equations (NLEEs) have become private species of the branch of partial differential equations (PDEs). Nonlinear evolution equations (NLEEs) are often used to explain a lot of physical events in the areas such as acoustic waves, hydromagnetic waves, chemistry, meteorology, engineering, thermodynamic, biology, physics, fluid mechanic, meteorology, optical fibers, heat transfer, acoustic gravity waves in mathematics. Because of this, most of methods have been developed and applied for these problems. Some of these methods include Extended simple equation method [1, 2], The Paul-Painlevé approach method [3], Multiple Exp-function method [4], ETEM [5], GKM [6]. The goal of this study, MEFM [7, 8] will be used to acquire new exact solutions of the simplified MCH equation and the Getmanou equation.

Firstly, we consider the simplified MCH equation [9, 10],

$$s_t + 2\alpha s_x - s_{xxt} + \beta s^2 s_x = 0 \quad \alpha \in \mathfrak{R}, \beta > 0 \quad (1)$$

where α and β are constants. Wazwaz [11] investigated a modified form of the Camassa-Holm equation, which was simplified from the MCH equation and the equation expressed in Eqn. (1) is called the simplified MCH equation.

Secondly, we investigate the Getmanou equation [12, 13],

$$s_{xt} + \frac{s_x s_t}{1-s^2} - s(1-s^2) = 0 \quad (2)$$

Getmanou equation possesses high nonlinearity [13]. Fan studied the single traveling wave solutions based on the complete discrimination system of the fifth-order polynomials [13]. The

trial equation method combined with complete discrimination system for polynomial has been used to solve the Getmanou equation [12].

Here, our aim is to find new exact solutions of the simplified MCH equation and the Getmanou equation by way of suggest method. In Section 2, we explain methodology. In Section 3, we apply suggest method to the simplified MCH equation and the Getmanou equation.

2. Materials and Methods

For a known nonlinear partial differential equations are given as follows:

$$K(s, s_t, s_x, s_y, s_{tt}, s_{xx}, s_{yy}, \dots) = 0, \tag{3}$$

where $s = s(x, y, z, t)$ is an obscure function.

Step 1: Getting the transmutation as

$$s(x, y, z, t) = S(\sigma), \quad \sigma = x + y + z - ct, \tag{4}$$

Eqn. (2) is turned into the following nonlinear equation:

$$L(S, S', S'', S''', \dots) = 0. \tag{5}$$

Step 2: Taking the following equation for Eqn. (5) as solution:

$$S(\sigma) = \frac{\sum_{i=0}^p A_i [\exp(-\mathcal{G}(\sigma))]^i}{\sum_{j=0}^q B_j [\exp(-\mathcal{G}(\sigma))]^j} = \frac{A_0 + A_1 \exp(-\mathcal{G}) + \dots + A_p \exp(p(-\mathcal{G}))}{B_0 + B_1 \exp(-\mathcal{G}) + \dots + B_q \exp(q(-\mathcal{G}))}, \tag{6}$$

where $A_i, B_j, (0 \leq i \leq p, 0 \leq j \leq q)$ are constants, such that $A_p \neq 0, B_q \neq 0$, and $\mathcal{G} = \mathcal{G}(\sigma)$ described as;

$$\mathcal{G}'(\sigma) = \exp(-\mathcal{G}(\sigma)) + a \exp(\mathcal{G}(\sigma)) + b. \tag{7}$$

Eqn. (7) has the following solution families:

Family 1: When $a \neq 0, b^2 - 4a > 0$,

$$\mathcal{G}(\sigma) = \ln \left(\frac{-\sqrt{b^2 - 4a}}{2a} \tanh \left(\frac{\sqrt{b^2 - 4a}}{2} (\sigma + E) \right) - \frac{b}{2a} \right). \tag{8}$$

Family 2: When $a \neq 0, b^2 - 4a < 0$,

$$\mathcal{G}(\sigma) = \ln \left(\frac{\sqrt{-b^2 + 4a}}{2a} \tanh \left(\frac{\sqrt{-b^2 + 4a}}{2} (\sigma + E) \right) - \frac{b}{2a} \right). \tag{9}$$

Family 3: When $a = 0$, $b \neq 0$, and $b^2 - 4a > 0$,

$$\mathcal{G}(\sigma) = -\ln \left(\frac{b}{\exp(b(\sigma + E)) - 1} \right). \tag{10}$$

Family 4: When $a \neq 0$, $b \neq 0$, and $b^2 - 4a = 0$,

$$\mathcal{G}(\sigma) = \ln \left(-\frac{2b(\sigma + E) + 4}{b^2(\sigma + E)} \right). \tag{11}$$

Family 5: When $a = 0$, $b = 0$, and $b^2 - 4a = 0$,

$$\mathcal{G}(\sigma) = \ln(\sigma + E). \tag{12}$$

where $A_i, B_j, (0 \leq i \leq p, 0 \leq j \leq q), E, b, a$ are constants to be obtained later.

Step 3: Setting Eqn. (6) and Eqn. (7) into Eqn. (5), a system of $e^{-\mathcal{G}(\sigma)}$ can be obtained. We solve this system by using Mathematica to identify the coefficients $A_i, B_j, (0 \leq i \leq p, 0 \leq j \leq q), E, b, a$.

3. Application of MEFM

3.1. Example: The simplified MCH equation

Getting the transformation as

$$s = s(\xi), \quad \xi = x - ct, \tag{13}$$

Eqn. (1) demean

$$(2\alpha - c)s + cs'' + \frac{\beta}{3}s^3 = 0 \tag{14}$$

By use of balance principle in Eqn. (14), we get

$$p = q + 1. \tag{15}$$

If we get $q = 1$ so $p = 2$, we have

$$S = \frac{A_0 + A_1 \exp(-\mathcal{G}) + A_2 \exp(2(-\mathcal{G}))}{B_0 + B_1 \exp(-\mathcal{G})} = \frac{\Upsilon}{\Psi}, \tag{16}$$

and

$$S' = \frac{\Upsilon'\Psi - \Psi'\Upsilon}{\Psi^2}, \tag{17}$$

$$S'' = \frac{\Upsilon''\Psi^3 - \Psi^2\Upsilon'\Psi' - (\Psi''\Upsilon + \Psi'\Upsilon')\Psi^2 + 2(\Psi')^2\Upsilon\Psi}{\Psi^4}, \tag{18}$$

$$\vdots$$

Thus, a system of $e^{-g(\sigma)}$ can be obtained. We solve this system by using Mathematica to identify the coefficients $A_i, B_j, (0 \leq i \leq p, 0 \leq j \leq q), E, b, a$.

Case 1:

$$A_0 = \frac{bA_2B_0}{2B_1}, A_1 = \frac{1}{2}A_2\left(b + \frac{2B_0}{B_1}\right), c = -\frac{\beta A_2^2}{6B_1^2}, a = \frac{1}{4}\left(2 + b^2 + \frac{24\alpha B_1^2}{\beta A_2^2}\right). \tag{19}$$

According to Eqn. (19), dark optical soliton solution for Eqn. (1) is gotten

$$s_1(x, t) = \frac{-24\alpha B_1^2 + \beta A_2^2 \left(-2 + b \sqrt{-2 - \frac{24\alpha B_1^2}{\beta A_2^2} \tanh[r(x, t)]} \right)}{2\beta A_2 B_1 \left(b + \sqrt{-2 - \frac{24\alpha B_1^2}{\beta A_2^2} \tanh[r(x, t)]} \right)}, \tag{20}$$

where

$$r(x, t) = \frac{1}{2} \left(E + x + \frac{t\beta A_2^2}{6B_1^2} \right) \sqrt{-2 - \frac{24\alpha B_1^2}{\beta A_2^2}}.$$

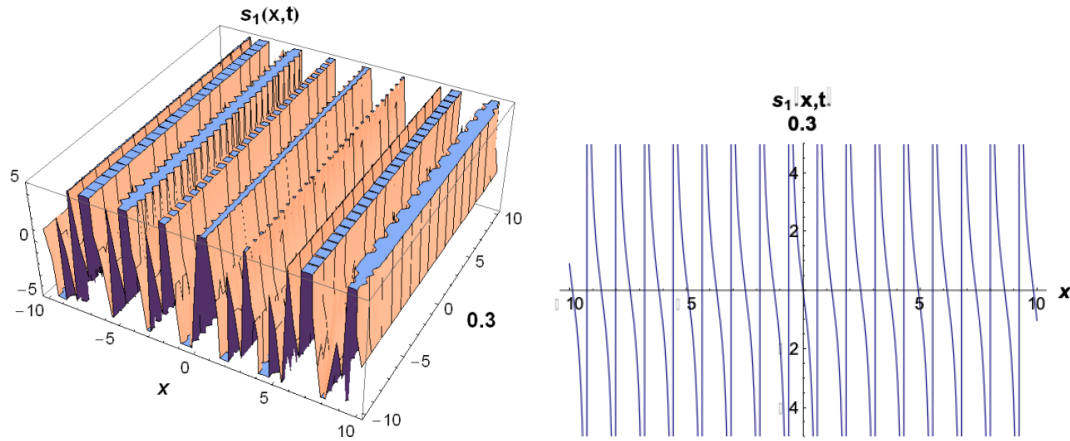


Figure 1: Three-dimensional and two-dimensional plots of imaginary values of Eqn. (20) for $A_2 = 4, B_1 = 4, E = 0.5, b = 0.3, \alpha = 3, \beta = 3, -10 < x < 10, -10 < t < 10$ and $t = 0.3$ for 2D plot

Case 2:

$$A_0 = -\frac{6\alpha B_0 B_1}{\beta A_1} + A_1 \left(\frac{B_0}{2B_1} - \frac{B_1}{8B_0} \right), A_2 = \frac{A_1 B_1}{2B_0} \tag{21}$$

$$b = \frac{2B_0}{B_1}, a = -\frac{1}{4} + B_0^2 \left(-\frac{12\alpha}{\beta A_1^2} + \frac{1}{B_1^2} \right), c = -\frac{\beta A_1^2}{24B_0^2}.$$

According to Eqn. (21), dark-bright optical soliton solution for Eqn. (1) is found

$$s_2(x,t) = -\frac{\operatorname{sech}[z(x,t)]^2 (\beta A_1^2 + 48\alpha B_0^2) (-P + \beta A_1^2 (B_0^2 - B_1^2))}{K (2B_0 + MB_1 \tanh[k(x,t)]) (P + L (B_1 + 2B_0 M \tanh[k(x,t)]))}, \tag{22}$$

where $M = \sqrt{1 + \frac{48\alpha B_0^2}{\beta A_1^2}}, P = 48\alpha B_0^2 B_1, z(x,t) = \frac{(t\beta A_1^2 + 24(E+x)B_0^2)M}{48B_0^2}$ and

$$k(x,t) = \frac{1}{2} \left(E + x + \frac{t\beta A_1^2}{24B_0^2} \right) M, \quad K = 4\beta A_1 B_0, \quad L = \beta A_1^2.$$

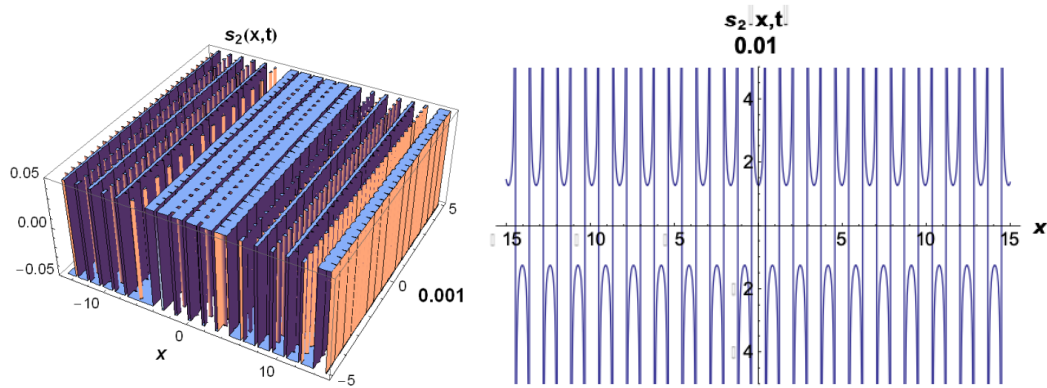


Figure 2: Three-dimensional and two-dimensional plots of imaginary values of Eqn. (22) for $A_2 = 0.2$, $B_0 = 0.3, B_1 = 0.1, E = 0.1, b = 0.3, \alpha = 1, \beta = -7, -15 < x < 15, -5 < t < 5$ and $t = 0.01$ for 2D plot

Case 3:

$$A_0 = \frac{1}{4}(-1 + b^2)A_2 - \frac{12\alpha B_0^2}{\beta b^2 A_2}, A_1 = bA_2, B_1 = \frac{2B_0}{b}, \tag{23}$$

$$c = -\frac{\beta b^2 A_2^2}{24B_0^2}, a = \frac{1}{4}\left(-1 + b^2 - \frac{48\alpha B_0^2}{\beta b^2 A_2^2}\right).$$

According to Eqn. (23), dark-bright optical soliton solution for Eqn. (1) is procured

$$s_3(x,t) = -\frac{\operatorname{sech}[h(x,t)]^2 (\beta b^2 (-1 + b^2) A_2^2 - 48\alpha B_0^2) (\beta b^2 A_2^2 + 48\alpha B_0^2)}{L (b + P \tanh[h(x,t)]) (M + \beta b^2 A_2^2 (1 + bP \tanh[h(x,t)]))}, \tag{24}$$

where $P = \sqrt{1 + \frac{48\alpha B_0^2}{\beta b^2 A_2^2}}, h(x,t) = \frac{1}{2}\left(E + x + \frac{t\beta b^2 A_2^2}{24B_0^2}\right)P, L = 4\beta b A_2 B_0, M = 48\alpha B_0^2.$

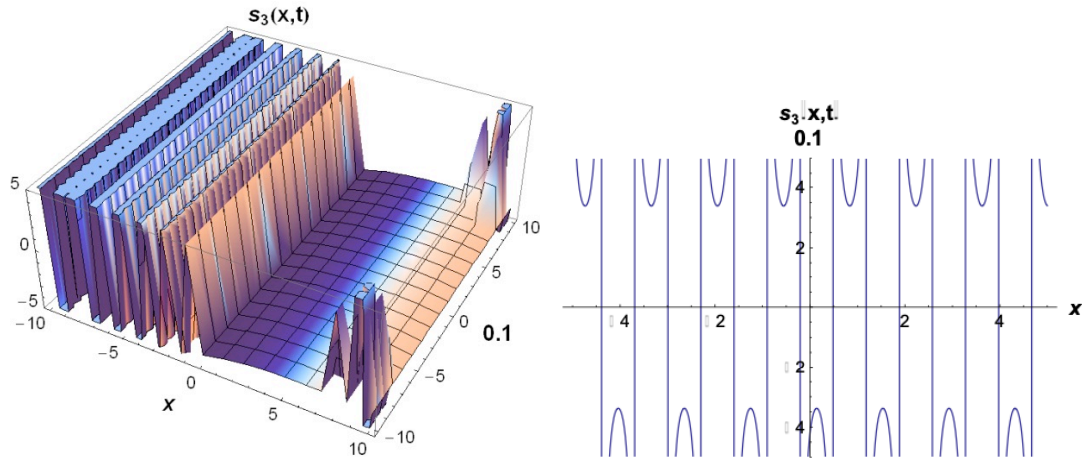


Figure 3: Three-dimensional and two-dimensional plots of real values of Eqn. (24) for $A_2 = 3, B_0 = 2, B_1 = 3, E = 0.1, b = -1, \alpha = -1, \beta = 1, -10 < x < 10, -10 < t < 10$ and $t = 0.1$ for 2D plot

Case 4:

$$A_0 = \frac{B_0 \left(A_1 - \frac{i\sqrt{6c}B_0}{\sqrt{\beta}} \right)}{B_1}, A_2 = \frac{i\sqrt{6c}B_1}{\sqrt{\beta}}, b = -\frac{\frac{i\sqrt{6c}B_0}{\sqrt{\beta}} + 6B_0}{3B_1}, \tag{25}$$

$$a = \frac{-\beta A_1^2 + 2i\sqrt{6c}\beta A_1 B_0 + 6cB_0^2 + 3(c - 2\alpha)B_1^2}{6cB_1^2}.$$

According to Eqn. (25), dark optical soliton solution for Eqn. (1) is obtained

$$s_4(x,t) = \frac{3\sqrt{6}(c - 2\alpha)B_1 + 3\sqrt{c}\sqrt{-1 + \frac{2\alpha}{c}} \left(i\sqrt{2\beta}A_1 + 2\sqrt{3c}B_0 \right) \tanh[f(x,t)]}{\sqrt{6\beta}A_1 - 3i\sqrt{c}\sqrt{\beta} \left(2B_0 - \sqrt{-2 + \frac{4\alpha}{c}}B_1 \tanh[f(x,t)] \right)}, \tag{26}$$

where $f(x,t) = (E - ct + x)\sqrt{-\frac{1}{2} + \frac{\alpha}{c}}$.

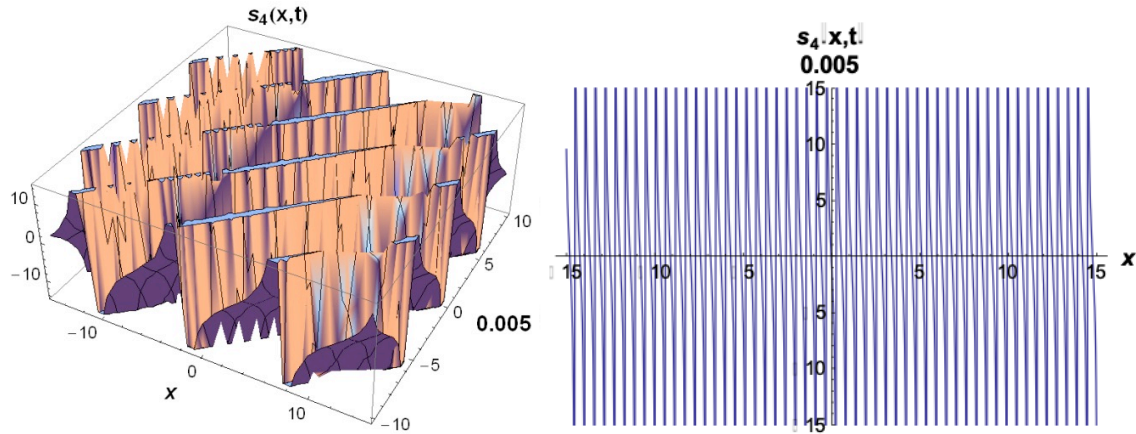


Figure 4: Three-dimensional and two-dimensional plots of real values of Eqn. (26) for $c = -0.1, \alpha = 3, \beta = -0.5, -15 < x < 15, -10 < t < 10, E = 0.3$ and $t = 0.005$ for 2D plot

Case 5:

$$A_0 = A_2 \left(-\frac{1}{4} + \frac{B_0^2}{B_1^2} \right) - \frac{3\alpha B_1^2}{\beta A_2}, A_1 = \frac{2A_2 B_0}{B_1}, \tag{27}$$

$$c = -\frac{\beta A_2^2}{6B_1^2}, b = \frac{2B_0}{B_1}, a = -\frac{1}{4} + \frac{B_0^2}{B_1^2} - \frac{3\alpha B_1^2}{\beta A_2^2}.$$

According to Eqn. (27), dark-bright optical solution for Eqn. (1) is attained

$$s_5(x,t) = \frac{\operatorname{sech}[f(x,t)]^2 (\beta A_2^2 + 12\alpha B_1^2) (12\alpha B_1^4 + \beta A_2^2 (-4B_0^2 + B_1^2))}{P(2B_0 + MB_1 \tanh[k(x,t)]) (K + \beta A_2^2 (B_1 + 2B_0 M \tanh[k(x,t)]))}, \tag{28}$$

where $M = \sqrt{1 + \frac{12\alpha B_1^2}{\beta A_2^2}}, P = 2\beta A_2 B_1, f(x,t) = \frac{(t\beta A_2^2 + 6(E+x)B_1^2)M}{12B_1^2}$ and

$$k(x,t) = \frac{1}{2} \left(E + x + \frac{t\beta A_2^2}{6B_1^2} \right) M, K = 12\alpha B_1^3.$$

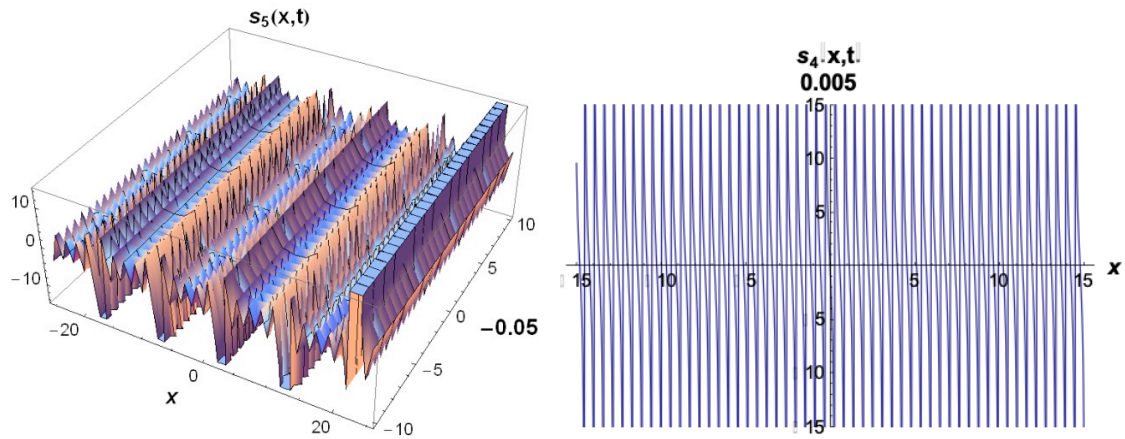


Figure 5: Three-dimensional and two-dimensional plots of real values of Eqn. (28) for $A_2 = 4, B_0 = -1, B_1 = -5, E = 0.3, \alpha = 3, \beta = -5, -25 < x < 25 -10 < t < 10$ and $t = -0.05$ for 2D plot

In Figs. 1-2, we plot two and three dimensional graphics of imaginary values of Eqn. (20) and Eqn. (22), which demonstrate the vitality of solutions with suitable parametric choices. Then, in Figs. 3-5, we draw two and three dimensional graphics of real values of Eqn. (24), Eqn. (26) and Eqn. (28), which indicate the dynamic of solutions with suitable parametric choices.

3.2. Example: The Getmanou equation

Getting the transformation as

$$s = s(\xi), \quad \xi = x - ct, \tag{29}$$

Eqn. (2) demans

$$c(s^2 - 1)s'' - c(s')^2 - s^5 + 2s^3 - s = 0. \tag{30}$$

By use of balance principle in Eqn. (30), we get

$$p = q + 1. \tag{31}$$

If we get $q = 1$ so $p = 2$, we have

$$S = \frac{A_0 + A_1 \exp(-\mathcal{G}) + A_2 \exp(2(-\mathcal{G}))}{B_0 + B_1 \exp(-\mathcal{G})} = \frac{\Upsilon}{\Psi}, \tag{32}$$

and

$$S' = \frac{\Upsilon'\Psi - \Psi'\Upsilon}{\Psi^2}, \tag{33}$$

$$S'' = \frac{\Upsilon''\Psi^3 - \Psi^2\Upsilon'\Psi' - (\Psi''\Upsilon + \Psi'\Upsilon')\Psi^2 + 2(\Psi')^2\Upsilon\Psi}{\Psi^4}, \tag{34}$$

$$\vdots$$

Thus, a system of $e^{-g(\sigma)}$ can be obtained. We solve this system by using Mathematica to identify the coefficients $A_i, B_j, (0 \leq i \leq p, 0 \leq j \leq q), E, b, a$.

Case 1:

$$A_0 = \frac{2B_0}{3} + \frac{A_2B_0^2}{B_1^2} + \frac{B_1^2}{36A_2}, A_1 = \frac{2A_2B_0}{B_1} + \frac{2B_1}{3}, c = \frac{A_2^2}{2B_1^2} \tag{35}$$

$$b = \frac{2B_0}{B_1} + \frac{5B_1}{3A_2}, a = \frac{5B_0}{3A_2} + \frac{B_0^2}{B_1^2} + \frac{B_1^2}{36A_2}$$

According to Eqn. (35), dark-bright optical soliton solution for Eqn. (2) is procured

$$s_1(x, t) = \left(\begin{aligned} & \frac{12(-1 + 8 \operatorname{sech}[f(x, t)]^2) A_2 B_0 B_1^2}{\left((B_1^2 + K(5 - 2\sqrt{6} \tanh[f(x, t)])) (K + B_1^2(5 + 2\sqrt{6} \tanh[f(x, t)])) \right)} \\ & + \frac{B_1^4(-5 + 4 \operatorname{sech}[f(x, t)]^2 - 2\sqrt{6} \tanh[f(x, t)])}{\left((B_1^2 + K(5 - 2\sqrt{6} \tanh[f(x, t)])) (K + B_1^2(5 + 2\sqrt{6} \tanh[f(x, t)])) \right)} \\ & + \frac{36A_2^2B_0^2(-5 + 4 \operatorname{sech}[f(x, t)]^2 + 2\sqrt{6} \tanh[f(x, t)])}{\left((B_1^2 + K(5 - 2\sqrt{6} \tanh[f(x, t)])) (K + B_1^2(5 + 2\sqrt{6} \tanh[f(x, t)])) \right)} \end{aligned} \right), \tag{36}$$

where $f(x, t) = \sqrt{\frac{2}{3}} \left(EE + x - \frac{tA_2^2}{2B_1^2} \right) \frac{B_1}{A_2}, K = 6A_2B_0$.

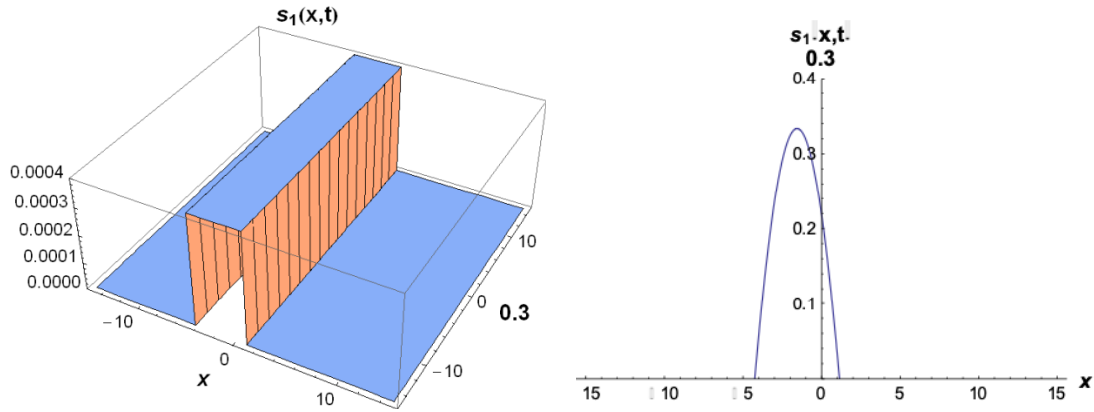


Figure 6: Three-dimensional and two-dimensional plots of solution Eqn. (36) for the values $A_2 = 5, B_0 = 0.2, B_1 = 2, E = 3, -15 < x < 15, -15 < t < 15$ and $t = 0.3$ for 2D plot

Case 2:

$$A_0 = \frac{2B_0}{3} + \frac{\sqrt{2c}B_0^2}{B_1} + \frac{B_1}{36\sqrt{2c}}, A_1 = \frac{2}{3}(3\sqrt{2c}B_0 + B_1), A_2 = \sqrt{2c}B_1 \tag{37}$$

$$b = \frac{5}{3\sqrt{2c}} + \frac{2B_0}{B_1}, a = \frac{1}{72} \left(\frac{1}{c} + \frac{12B_0\sqrt{c} + 5\sqrt{2}B_1}{\sqrt{c}B_1^2} \right)$$

According to Eqn. (37), dark-bright optical soliton solution for Eqn. (2) is found

$$s_2(x,t) = \frac{8 \operatorname{sech} [g(x,t)]^2 P - 2K + 8\sqrt{3}(72cB_0^2 - B_1^2) \tanh [g(x,t)]}{(2B_1 + M(L - 4\sqrt{3} \tanh [g(x,t)]))(M + B_1(L + 4\sqrt{3} \tanh [g(x,t)]))} \tag{38}$$

where $g(x,t) = \frac{EE - ct + x}{\sqrt{3c}}, K = 360\sqrt{2c}B_0^2 + 24\sqrt{c}B_0B_1 + 5\sqrt{2}B_1^2, L = 5\sqrt{2}$ and

$$P = 72\sqrt{2c}B_0^2 + 48\sqrt{c}B_0B_1 + \sqrt{2}B_1^2, M = 12\sqrt{c}B_0.$$

In Figs. 6-7, we plot two and three dimensional graphics of Eqn. (36) and Eqn. (38), which show the vitality of solutions with suitable parametric choices.

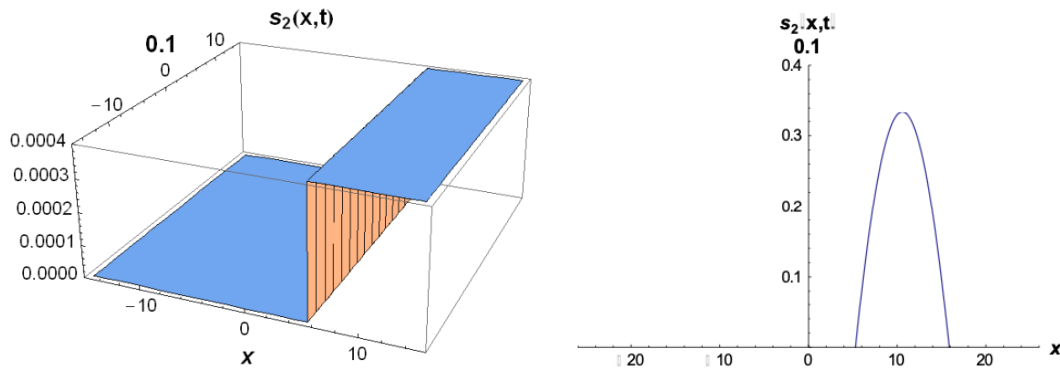


Figure 7: Three-dimensional and two-dimensional plots of solution Eqn. (38) for the values $A_2 = 0.4$, $B_0 = -8$, $B_1 = 0.015$, $E = -2.5$, $-15 < x < 15$, $-15 < t < 15$ and $t = 0.1$ for 2D plot

4. Conclusion

In this study, the simplified MCH equation and the Getmanou equation are researched by via a MEFM. After, we draw 2D and 3D graphs of dark optical soliton solutions, dark-bright optical soliton solutions of this equation by use of Mathematica. Soliton solutions are of two types as dark soliton and bright soliton. If there is a solution of type sech hyperbolic function, it is called a bright soliton solution and if there is a solution of type tanh hyperbolic function, it is called a dark soliton solution [14]. Solutions are called dark-bright solitons if they contain the sech and tanh functions at the same time.

From the obtained results, it has been deduced that MEFM is highly credible and strong in the sense that finding exact solutions. The solutions we obtained from equations are new solutions brought to the literature. The research shows that the MEFM algorithm is productive and can be used for many other NLEEs in mathematical physics.

References

- [1] Ali, A., Aly, R.S., Dianchen, L., *Dispersive solitary wave soliton solutions of (2 + 1)-dimensional Boussineq dynamical equation via extended simple equation method*, Journal of King Saud University – Science, 31, 653–658, 2019.
- [2] Lu, D., Seadawy, A.R, Ali, A., *Applications of extended simple equation method on unstable nonlinear Schrödinger equations*, Optik, 140, 136-144, 2017.
- [3] Bekir, A., Zahran, E., *Three distinct and impressive visions for the soliton solutions to the higher-order nonlinear Schrodinger equation*, Optik, 228(3), 166157, 2021.
- [4] Nisar, S.K., Ilhan, O.A., Abdulazeez, S.T., Manafian, J.M., Osman, M.S., *Novel multiple soliton solutions for some nonlinear PDEs via multiple Exp-function method*, Results in Physics, 21, 103769, 2021.
- [5] Tuluçe Demiray, S., Pandir, Y., Bulut, H., *All exact travelling wave solutions of Hirota equation and Hirota-Maccari system*, Optik, 127, 1848-1859, 2016.

- [6] Tuluçe Demiray, S., *New soliton solutions of optical pulse envelope $E(Z, \tau)$ with Beta Time Derivative*, *Optik*, 223, 165453, 2020.
- [7] Koçak, Z.F., Bulut, H., Koc, D.A., Başkonuş, H.M., *Prototype traveling wave solutions of new coupled Konno-Oono equation*, *Optik*, 127, 10786-10794, 2016.
- [8] Tuluçe Demiray, S., Kastal, S., *Dark-Bright optical soliton solutions of (3+1)-dimensional modified quantum Zakharov-Kuznetsov equation*, *Turkish Journal of Mathematics and Computer Science*, 11, 90-94, 2019.
- [9] Hossain, A.K.M.K.S., Akbar, M.A., *Solitary wave solutions of few nonlinear evolution equations*, *AIMS Mathematics*, 5, 1119-1215, 2020.
- [10] Irshad, A., Usman, M., Mohyud-Din, S.T., *Exp-function method for simplified modified Camassa-Holm equation*, *International Journal of Modern Mathematical Sciences*, 4(3), 146-155, 2012.
- [11] Wazwaz, A., *New compact and noncompact solutions for two variants of a modified Camassa-Holm equation*, *Applied Mathematics and Computations*, 163 (3), 1165-1179, 2005.
- [12] Yang, L., *Application of trial equation method for solving the Getmanou equation*, *Applied Mathematics*, 5, 1463-1473, 2014.
- [13] Fan, X.F., *Classification of all single traveling-wave solutions to Getmanou equation*, *FIZIKI A*, 19, 191-196, 2011.
- [14] Tuluçe Demiray, S., Pandır, Y., Bulut, H., *New solitary wave solutions of Maccari system*, *Ocean Engineering*, 103, 153-159, 2015.



**Solitary Wave Solutions of the (3+1)-dimensional Khokhlov–Zabolotskaya–
Kuznetsov Equation by Using the $(G'/G, 1/G)$ -Expansion Method**

Hülya DURUR¹, Serbay DURAN^{2,*}, Asif YOKUŞ³

¹*Ardahan University, Faculty of Engineering, Department of Computer Engineering, 75000, Ardahan, Türkiye*

hulyadurur@ardahan.edu.tr, ORCID: 0000-0002-9297-6873

²*Adiyaman University, Faculty of Education, Department of Mathematics and Science Education, 02040, Adiyaman, Türkiye*

sduran@adiyaman.edu.tr, ORCID: 0000-0002-3585-8061

³*Fırat University, Faculty of Science, Department of Mathematics, 23100, Elazığ, Türkiye*

asfyokus@yahoo.com, ORCID: 0000-0002-1460-8573

Received: 24.02.2021

Accepted: 04.10.2021

Published: 31.12.2021

Abstract

In this study, the (3+1)-dimensional Khokhlov-Zabolotskaya-Kuznetsov (KZK) equation, which is a mathematical model of non-absorption and dispersion in the non-linear medium, which sheds light on the sound beam phenomenon, which has a physically important place, is examined. In order to find the exact solution of this equation, an effective and reliable method, $(G'/G, 1/G)$ -expansion method, is used among analytical methods. The purpose of this method is to obtain more than one traveling wave solution classes depending on the conditions of the λ parameter. These classes are categorized into hyperbolic, trigonometric, complex trigonometric and rational forms. The graphics of the solitary waves represented by these successfully obtained solution classes are presented as 2-dimensional, 3-dimensional and contours. This article makes use of ready-made package programs for complex arithmetic operations and graphic drawings.



Keywords: $(G'/G, 1/G)$ -expansion method; (3+1)-dimensional Khokhlov–Zabolotskaya–Kuznetsov equation; Traveling wave solution.

(3+1)-Boyutlu Khokhlov–Zabolotskaya–Kuznetsov Denkleminin $(G'/G, 1/G)$ -Açılım Metodu Yardımıyla Solitary Dalga Çözümleri

Öz

Bu çalışmada, fiziksel olarak önemli bir yere sahip olan ses ışını (sound beam) olayına ışık tutan, özellikle lineer olmayan ortamda dağılım ve soğurma olmayan durumların matematiksel modeli olan (3+1)-boyutlu Khokhlov–Zabolotskaya–Kuznetsov (KZK) denklemi incelendi. Bu denklemin tam çözümünü bulmak için analitik metotlar arasında yer alan etkili ve güvenilir bir yöntem olan $(G'/G, 1/G)$ -açılım metodu kullanıldı. Bu metodun seçilme amacı λ parametresinin durumlarına bağlı olarak birden fazla yürüyen dalga çözüm sınıfları elde edilmesidir. Bu sınıflar hiperbolik, trigonometrik, kompleks trigonometrik ve rasyonel formda kategorize edilir. Başarılı bir şekilde elde edilen bu çözüm sınıflarının temsil ettiği solitary dalgaların grafikleri 2-boyutlu, 3-boyutlu ve kontur olarak sunuldu. Bu makalede karmaşık aritmetik işlemler ve grafik çizimleri için hazır paket programlardan faydalanıldı.

Anahtar Kelimeler: $(G'/G, 1/G)$ -açılım metodu; (3+1)-boyutlu Khokhlov–Zabolotskaya–Kuznetsov Denklemi; Solitary dalga çözümleri.

1. Introduction

The debates about the wave theory that started in the 18th century have been brought to a considerable level. The wave theory we are discussing today and discussed in the future can be divided into two groups, linear and nonlinear. However, nonlinear wave discussions are more valuable because life is not linear. For this reason, the traveling wave solutions of partial differential equations shed light on many events in nature, bringing mathematical models to the fore. Along with these mathematical models, many researchers have discussed the solution methods of these models. Generally, the methods that generate the solutions of nonlinear mathematical models are of the oscillating traveling wave type. In applied science, studies about perceiving the traveling wave as a signal and processing these signals have become popular today. Mathematical models, called NPDEs include quantum mechanics, plasma physics, hydrodynamic molecular biology, sheet water wave, nonlinear optics, optical fibers, chemistry, biological science, etc. as seen in various fields of nonlinear science. Investigating NPDEs provides a clearer understanding of complex events. Lately, many new mathematical models used by experts all over the world to describe real-life problems of today have attracted attention.

In this sense, some methods are trial equation method, modified simple equation method, modified extended tanh method, generalized hyperbolic-function method, sub equation method, complex method, auxiliary equation method, the homogeneous balance method, the improved Bernoulli sub-equation function method and many more methods [1-29].

We consider the following Zabolotskaya and Khokhlov (ZK) equation [30],

$$(u_t + uu_x)_x + nu_{yy} + mu_{zz} = 0. \tag{1}$$

This equation was first proposed by Zabolotskaya and Khokhlov in 1969 [31]. The physical interpretation of this equation shows the propagation of the sound beam in a non-linear medium with no dispersion or absorption [32]. This nonlinear medium in particular is not strong. This non-linear medium in particular is not strong. With the term added to Eq. (1), the following (3+1)-dimensional KZK equation is obtained [32]:

$$u_{xt} + (u_x)^2 + uu_{xx} + ru_{xxx} + nu_{yy} + mu_{zz} = 0, \tag{2}$$

where r , n and m are constant and $r \neq 0$. In addition, in Eqn. (2), which is the mathematical model of the sound beam phenomenon, the function that represents acoustic pressure and sought is $u(x, y, z, t)$. Here t represents time and $(x, y, z) \in R^3$ [33]. This equation was first proposed by Kuznetsov with the help of Eqn. (1) in 1971 [34]. The term adsorption is defined as thermo-viscous. A higher-order NPDEs have been defined by adding this term. Traveling wave solutions were investigated for Eqn. (2) by Akçagil and Aydemir in 2016 with the help of the tanh-coth method [32]. On the other hand, new exact solutions were reached by Ray with the help of Kudryashov methods for the time fractional KZK equation [35]. In 2019, analytical solutions of the (3+1) dimensional time fractional KZK equation were produced with the help of modified Riemann-Liouville derivative and (G'/G) -expansion method by Zhang et al. [36]. In addition, the effect of diffraction in these solutions was investigated. In 2021, traveling wave solutions were produced in trigonometric function and dark optical soliton solution format by applying the modified $\exp(-\Omega(\xi))$ -expansion function method for Eqn. (2) by Demiray and Kastal [37]. The main theme of this study is to obtain the traveling wave solutions of Eqn. (2) with the help of the $(G'/G, 1/G)$ - expansion method [38].

The most important reason for using this method is to produce different types of traveling wave solutions from the literature for the (3+1)-dimensional KZK equation. One of the most important advantages of this method is that it produces traveling wave solutions in three different forms. In this study, information about the methodology of the method discussed in Section 2 is given. In the Section 3, the application of the method to the Eqn. (2) and finally in the Section 4, important results are given.

2. Method

2.1. (G'/G, 1/G)-expansion method

In this section, we present analysis of the (G'/G, 1/G)-expansion method [38].

$$Z(u, u_x, u_y, u_z, u_t, u_{xx}, u_{tt}, \dots) = 0. \tag{3}$$

If $u = U(\xi) = u(x, y, z, t)$, $\xi = x + y + z - ct$ classical wave transformation is applied in Eqn. (3) while c is a constant, Eqn. (3) is converted into a nODE and this can be written as:

$$W(U, UU', U'', \dots) = 0. \tag{4}$$

Reduced complexity by integrating Eqn. (4). $G(\xi)$ function is a quadratic function ODE solution,

$$G''(\xi) + \lambda G(\xi) = \mu. \tag{5}$$

Also to ensure operational aesthetics as $\frac{G'}{G} = \phi = \phi(\xi)$ and $\psi = \psi(\xi) = \frac{1}{G(\xi)}$. Here, the derivatives of the defined functions can be written

$$\phi' = -\phi^2 + \mu\psi - \lambda, \quad \psi' = -\phi\psi. \tag{6}$$

By considering the equations given by Eqn. (6), we can present the behavior of the solution function Eqn. (5) with respect to the λ state.

i) If $\lambda < 0$

$$G(\xi) = c_1 \sinh(\sqrt{-\lambda}\xi) + c_2 \cosh(\sqrt{-\lambda}\xi) + \frac{\mu}{\lambda}, \tag{7}$$

where c_2 and c_1 are real numbers. Considering Eqn. (7);

$$\psi^2 = \frac{-\lambda}{\lambda^2\sigma + \mu^2} (\phi^2 - 2\mu\psi + \lambda), \quad \sigma = c_1^2 - c_2^2, \tag{8}$$

written in this form.

ii) If $\lambda > 0$

$$G(\xi) = c_1 \sin(\sqrt{\lambda}\xi) + c_2 \cos(\sqrt{\lambda}\xi) + \frac{\mu}{\lambda}, \tag{9}$$

where c_2 and c_1 are real numbers. Eqn. (9), there is following equation;

$$\psi^2 = \frac{\lambda}{\lambda^2\sigma - \mu^2} (\phi^2 - 2\mu\psi + \lambda), \quad \sigma = c_1^2 + c_2^2, \tag{10}$$

iii) If $\lambda = 0$

$$G(\xi) = \frac{\mu}{2}\xi^2 + c_1\xi + c_2, \tag{11}$$

where c_2 and c_1 are real numbers. Eqn. (11), there is following equation;

$$\psi^2 = \frac{1}{c_1^2 - 2\mu c_2}(\phi^2 - 2\mu\psi). \tag{12}$$

The solution of Eqn. (3) in terms of ψ and ϕ polynomials is

$$U(\xi) = \sum_{i=0}^n a_i \phi^i + \sum_{i=1}^n b_i \phi^{i-1} \psi, \tag{13}$$

where in b_i ($i = 1, \dots, n$) and a_i ($i = 0, 1, \dots, n$) are constants to calculate. n is a positive integer to be calculated according to the balance principle for Eqn. (4). The corresponding derivatives of Eqn. (13) are calculated. These derivatives are substituted in Eqn. (4). Next, the polynomial is connected to ψ and ϕ are formed. Equating the coefficients of the ψ and ϕ in the obtained polynomial to zero, a system of equations is constructed. The built equation system is solved with the help of a computer software program. The values of the calculated constants are written in their place in Eqn. (13). Solutions of Eqn. (4) are obtained. Thus, we find the solutions in relation to the hyperbolic functions for $\lambda < 0$, the trigonometric functions for $\lambda > 0$ and the rational functions for $\lambda = 0$.

3. Solutions of the (3+1)-dimensional KZK Equation via $(G'/G, 1/G)$ -expansion Method

We consider Eqn. (2). Additionally, let us consider traditional wave transform as below:

$$u = U(\xi) = u(x, y, z, t), \quad \xi = x + y + z - ct. \tag{14}$$

We write Eqn. (14) into system Eqn. (2) to attain nonlinear ODEs

$$(m + n - c)U + \frac{1}{2}U^2 + rU' = 0. \tag{15}$$

By use of balance principle in Eqn. (15), we get $n = 1$ and in Eqn. (13) the following situation is attained:

$$U(\xi) = a_0 + a_1 \phi[\xi] + b_1 \psi[\xi], \tag{16}$$

where a_0, a_1, b_1 then the constants to be determined are unknown. If Eqn. (16) is written in Eqn. (15) and the coefficients of the Eqn. (2) equal zero, we can set up the following systems of an algebraic equation

$$\begin{aligned} (\phi[\xi])^0 & : -ca_0 + ma_0 + na_0 + \frac{a_0^2}{2} - r\lambda a_1 - \frac{\lambda^2 b_1^2}{2(\mu^2 + \lambda^2 \sigma)} = 0, \\ \phi[\xi] & : -ca_1 + ma_1 + na_1 + a_0 a_1 = 0, \\ (\phi[\xi])^2 & : -ra_1 + \frac{a_1^2}{2} - \frac{\lambda b_1^2}{2(\mu^2 + \lambda^2 \sigma)} = 0, \\ \psi[\xi] & : r\mu a_1 - cb_1 + mb_1 + nb_1 + a_0 b_1 + \frac{\lambda \mu b_1^2}{\mu^2 + \lambda^2 \sigma} = 0, \\ \phi[\xi]\psi[\xi] & : -rb_1 + a_1 b_1 = 0. \end{aligned} \tag{17}$$

With the software program, we reached the solutions of the system (17) and the following situations.

If $\lambda < 0$,

Case 1.

$$a_0 = -2ir\sqrt{\lambda}, \quad a_1 = 2r, \quad b_1 = 0, \quad \mu = 0, \quad c = m + n - 2ir\sqrt{\lambda}, \quad (18)$$

where $i = \sqrt{-1}$, replacing Eqn. (18) into Eqn. (16), the following complex hyperbolic solution is attained

$$u_1(x, y, z, t) = -2ir\sqrt{\lambda} + \frac{(2r(c_2\sqrt{-\lambda} \cosh[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{-\lambda}] + c_1\sqrt{-\lambda} \sinh[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{-\lambda}]))}{(c_1 \cosh[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{-\lambda}] + c_2 \sinh[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{-\lambda}])}. \quad (19)$$

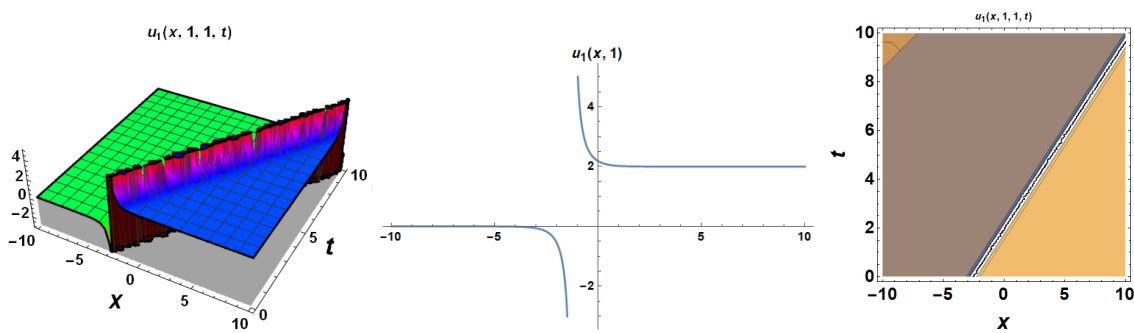


Figure 1: 3D, 2D and contour graphs for $c_2 = 2$, $c_1 = 1$, $\lambda = -1$, $r = 0.5$, $m = 0.2$, $n = 0.1$, $y = 1$, $z = 1$ of Eqn. (19)

There is $\sqrt{\lambda}\sqrt{-\lambda}$ in u that we have presented as a solution. Since $\lambda < 0$, we have presented the solution consists only of the real part.

Case 2.

$$a_0 = ir\sqrt{\lambda}, \quad a_1 = r, \quad b_1 = \frac{\sqrt{-r^2\mu^2 - r^2\lambda^2\sigma}}{\sqrt{\lambda}}, \quad c = m + n + ir\sqrt{\lambda}, \quad (20)$$

where $i = \sqrt{-1}$, replacing Eqn. (20) into Eqn. (16), the following hyperbolic solution is attained

$$u_2(x, y, z, t) = ir\sqrt{\lambda} + \frac{\sqrt{-(-c_1^2 + c_2^2)r^2\lambda^2 - r^2\mu^2}}{\sqrt{\lambda} \left(\frac{\mu}{\lambda} + c_1 \cosh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}] + c_2 \sinh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}] \right)}$$

$$+ \frac{r(c_2\sqrt{-\lambda} \cosh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}] + c_1\sqrt{-\lambda} \sinh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}])}{\frac{\mu}{\lambda} + c_1 \cosh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}] + c_2 \sinh[(x+y+z-t(m+n+ir\sqrt{\lambda}))\sqrt{-\lambda}]}. \quad (21)$$

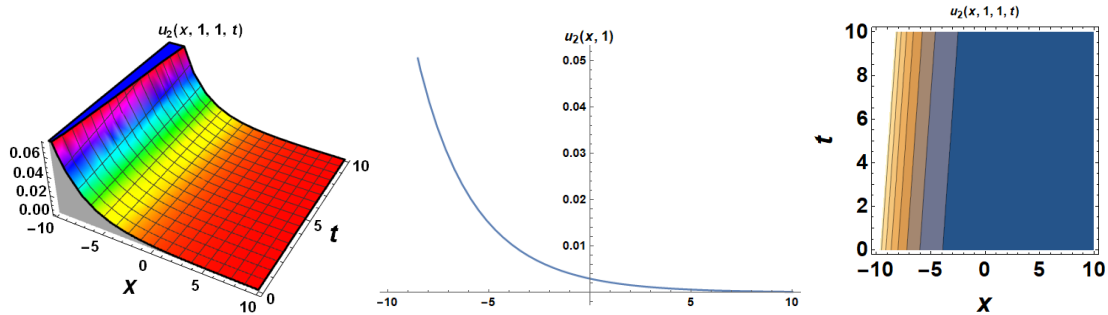


Figure 2: 3D, 2D and contour graphs for $c_2 = 2$, $c_1 = 1$, $\lambda = -0.1$, $\mu = -3$, $r = 0.5$, $m = 0.2$, $n = 0.1$, $y = 1$, $z = 1$ values of Eqn. (21)

There is $\sqrt{\lambda}\sqrt{-\lambda}$ in u that we have presented as a solution. Since $\lambda < 0$, we have presented the solution consists only of the real part.

If $\lambda > 0$,

Case 1.

$$a_0 = -2ir\sqrt{\lambda}, \quad a_1 = 2r, \quad b_1 = 0, \quad \mu = 0, \quad c = m + n - 2ir\sqrt{\lambda}, \quad (22)$$

where $i = \sqrt{-1}$, replacing Eqn. (22) in Eqn. (16), the following trigonometric solution is attained

$$u_3(x, y, z, t) = -2ir\sqrt{\lambda} + \frac{2r(c_2\sqrt{\lambda} \cos[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{\lambda}] - c_1\sqrt{\lambda} \sin[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{\lambda}])}{c_1 \cos[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{\lambda}] + c_2 \sin[(x+y+z-t(m+n-2ir\sqrt{\lambda}))\sqrt{\lambda}]}. \quad (23)$$

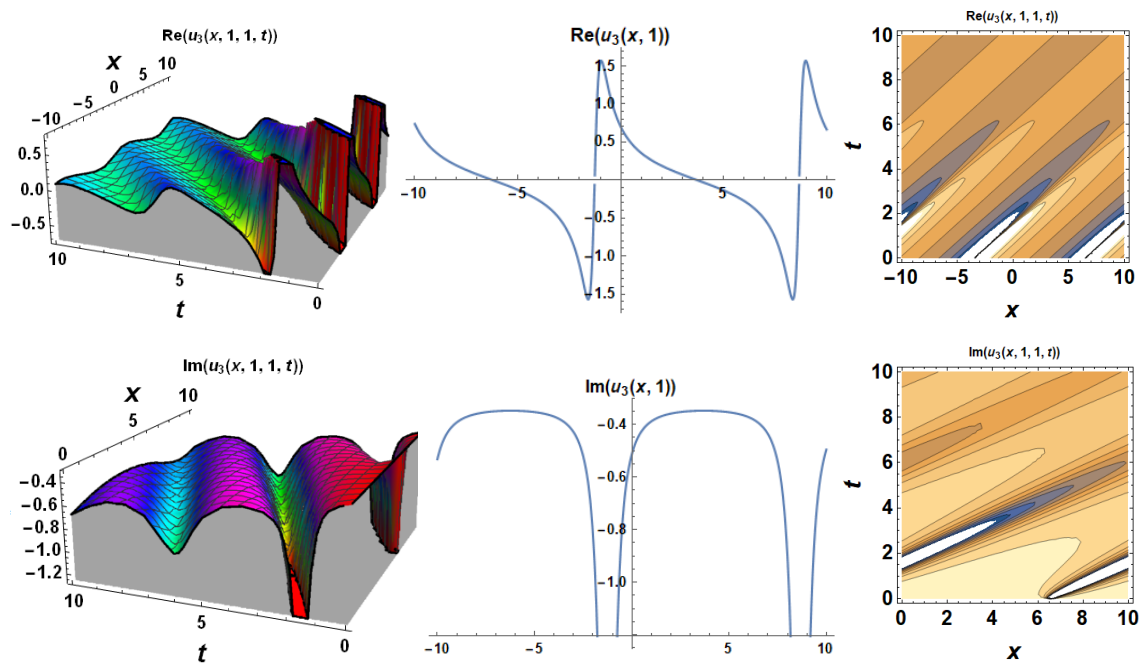


Figure 3: Real and imaginary parts of 3D, 2D and contour graphs for $c_2 = 2$, $c_1 = 1$, $\lambda = 0.1$, $r = 0.5$, $m = 1$, $n = 1.2$, $y = 1$, $z = 1$ of Eqn. (23)

If $\lambda = 0$,

Case 1.

$$a_0 = 0, \quad a_1 = 2r, \quad b_1 = 0, \quad \mu = 0, \quad c = m + n, \tag{24}$$

replacing Eqn. (24) in Eqn. (16), the following rational solution is attained

$$u_4(x, y, z, t) = \frac{2c_2r}{c_1+c_2(-m+n)t+x+y+z} \tag{25}$$

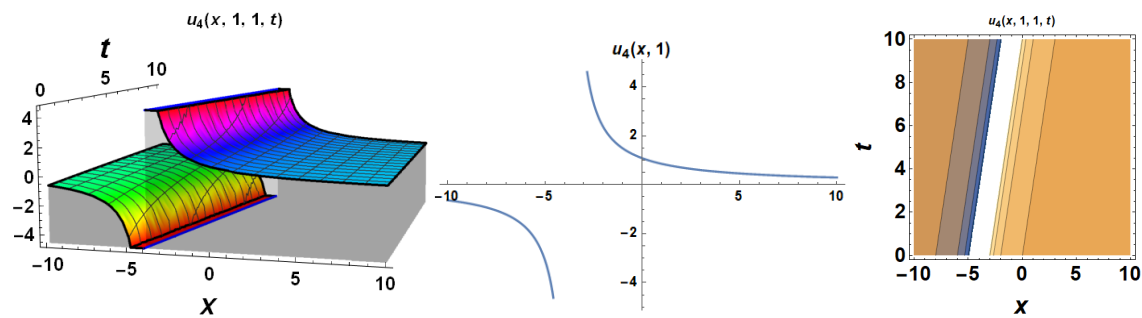


Figure 4: 3D, 2D and contour graphs for $c_2 = 0.5$, $c_1 = 1$, $\lambda = 0$, $r = 2$, $m = 0.2$, $n = 0.1$, $y = 1$, $z = 1$ of Eqn. (25)

Case 2.

$$a_0 = 0, \quad a_1 = r, \quad \mu = \frac{c_2^2r^2-b_1^2}{2c_1r^2}, \quad c = m + n, \tag{26}$$

replacing Eqn. (26) in Eqn. (16), the following rational solution is attained

$$\begin{aligned}
 &u_5(x, y, z, t) \\
 &= \frac{b_1}{c_1 + c_2(- (m + n)t + x + y + z) + \frac{(- (m + n)t + x + y + z)^2 (c_2^2 r^2 - b_1^2)}{4c_1 r^2}} \\
 &+ \frac{r \left(c_2 + \frac{(- (m + n)t + x + y + z) (c_2^2 r^2 - b_1^2)}{2c_1 r^2} \right)}{c_1 + c_2(- (m + n)t + x + y + z) + \frac{(- (m + n)t + x + y + z)^2 (c_2^2 r^2 - b_1^2)}{4c_1 r^2}}. \tag{27}
 \end{aligned}$$

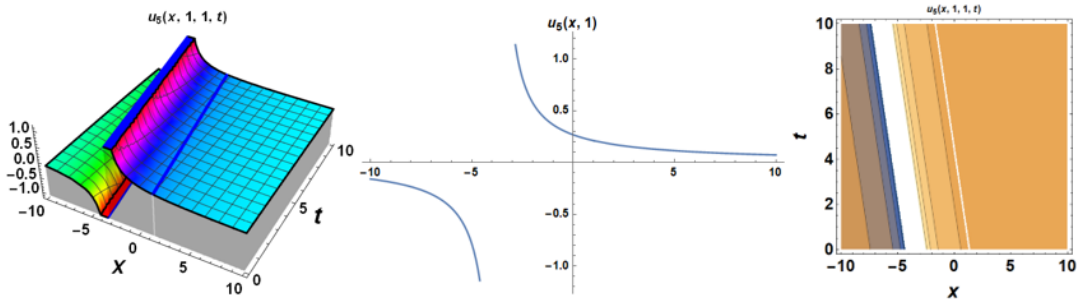


Figure 5: 3D, 2D and contour graphs for $c_2 = 0.4$, $c_1 = 1$, $b_1 = 0.5$, $\lambda = 0$, $r = 0.5$, $m = -0.2$, $n = -0.1$, $y = 1$, $z = 1$ of Eqn. (27)

Traveling wave solutions play an important role in physically transporting energy from one place to another. The traveling wave solutions obtained in this study can offer a different perspective to the acoustic theory. The graphs presented in Figs. 1-5 illustrate the wave behaviour of traveling wave solutions at any instant, which we can call a standing wave. While drawing these graphs, the y and z dimensions are considered fixed.

4. Conclusion

In this study, we have proposed hyperbolic, trigonometric, complex trigonometric and rational traveling wave solutions with the help of $(G'/G, 1/G)$ -expansion method of Eqn. (2) which is the mathematical model of the sound beam in a non-linear medium without physical dispersion and absorption. The method is generally categorized into three different classes depending on the λ parameter. The equation was checked with the help of a ready-made package program that the traveling wave solutions obtained for each class provided. In the traveling wave solutions obtained, solitary wave solutions were obtained by giving arbitrary constants to the parameters and the graphics were presented as 3D, 2D, and contour. The solution of the algebraic equation system discussed in this study, complex operations and the graphics of these solutions were obtained using a ready-made package program. It has been concluded that this method we have used is useful and reliably applicable in equations with strong nonlinearity.

References

- [1] Bulut, H., Baskonus, H.M., Pandir, Y., *The modified trial equation method for fractional wave equation and time fractional generalized Burgers equation*, Abstract and Applied Analysis, vol. 2013, Article ID 636802, 2013.
- [2] Arnous, A.H., Seadawy, A.R., Alqahtani, R.T., Biswas, A., *Optical solitons with complex Ginzburg–Landau equation by modified simple equation method*, Optik, 144, 475-480, 2017.
- [3] Xiong, M., Chen, L., Li, C., Wang, J., *Exact Solutions for $(2+1)$ -Dimensional Nonlinear Schrödinger Schrodinger Equation Based on Modified Extended tanh Method*, In the International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, 224-231, 2019.
- [4] Zayed, E.M., Shohib, R.M., *Optical solitons and other solutions to Biswas–Arshed equation using the extended simplest equation method*, Optik, 185, 626-635, 2019.
- [5] Gürefe, Y., Aktürk, T., *Investigation of Analytical Solutions of the Nonlinear Mathematical Model Representing Gas Overflowing*, Adıyaman University Journal of Science, 11(1), 182-190, 2021.
- [6] Duran, S., Karabulut, B., *Nematicons in liquid crystals with Kerr Law by sub-equation method*, Alexandria Engineering Journal, 61(2), 1695-1700, 2022.
- [7] Durur, H., Taşbozan, O., Kurt, A., Şenol, M., *New Wave Solutions of Time Fractional Kadomtsev–Petviashvili Equation Arising In the Evolution of Nonlinear Long Waves of Small Amplitude*, Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 12(2), 807-815, 2019.
- [8] Sulaiman, T.A., Yavuz, M., Bulut, H., Baskonus, H.M., *Investigation of the fractional coupled viscous Burgers' equation involving Mittag-Leffler kernel*, Physica A: Statistical Mechanics and its Applications, 527, 121126, 2019.
- [9] Tozar, A., Tasbozan, O., Kurt, A., *Analytical solutions of Cahn–Hilliard phase-field model for spinodal decomposition of a binary system*, Europhysics Letters, 130(2), 24001, 2020.
- [10] Yavuz, M., Ozdemir, N., Baskonus, H.M., *Solutions of partial differential equations using the fractional operator involving Mittag-Leffler kernel*, The European Physical Journal Plus, 133(6), 1-11, 2018.
- [11] Aktürk, T., Kubal, Ç., *Analysis of wave solutions of $(2+1)$ -dimensional Nizhnik–Novikov–Veselov equation*, Ordu Üniversitesi Bilim ve Teknoloji Dergisi, 11(1), 13-24, 2021.
- [12] Durur, H., *Different types analytic solutions of the $(1+1)$ -dimensional resonant nonlinear Schrödinger's equation using (G'/G) -expansion method*, Modern Physics Letters B, 34(03), 2050036, 2020.
- [13] Yokus, A., Durur, H., Ahmad, H., *Hyperbolic type solutions for the couple Boiti–Leon–Pempinelli system*, Facta Universitatis, Series: Mathematics and Informatics, 35(2), 523-531, 2020.
- [14] Duran, S., Yokuş, A., Durur, H., Kaya, D., *Refraction simulation of internal solitary waves for the fractional Benjamin–Ono equation in fluid Dynamics*, Modern Physics Letters B, 2150363, 2021.
- [15] Li, L., Li, E., Wang, M., *The $(G'/G, 1/G)$ -expansion method and its application to travelling wave solutions of the Zakharov equations*, Applied Mathematics-A Journal of Chinese Universities, 25, 454–462, 2010.

- [16] Duran, S., *Solitary Wave Solutions of the Coupled Konno-Oono Equation by using the Functional Variable Method and the Two Variables (G'/G , $1/G$)-Expansion Method*, Adıyaman Üniversitesi Fen Bilimleri Dergisi, 10(2), 585-594, 2020.
- [17] Yokus, A., Durur, H., Ahmad, H., Yao, S.W., *Construction of Different Types Analytic Solutions for the Zhiber-Shabat Equation*, Mathematics, 8(6), 908, 2020.
- [18] Yokuş, A., Durur, H. Duran, S. *Simulation and refraction event of complex hyperbolic type solitary wave in plasma and optical fiber for the perturbed Chen-Lee-Liu equation*, Optical and Quantum Electronics, 53, 402, 1-17, 2021.
- [19] Yavuz, M., Yokus, A., *Analytical and numerical approaches to nerve impulse model of fractional-order*, Numerical Methods for Partial Differential Equations, 36(6), 1348-1368, 2020.
- [20] Duran, S., *Exact Solutions for Time-Fractional Ramani and Jimbo-Miwa Equations by Direct Algebraic Method*, Advanced Science, Engineering and Medicine, 12(7), 982-988, 2020.
- [21] Kaya, D., Yokuş, A., Demiroğlu, U., *Comparison of exact and numerical solutions for the Sharma–Tasso–Olver equation*, In Numerical Solutions of Realistic Nonlinear Phenomena, 53-65, 2020.
- [22] Yokuş, A., Durur, H., Abro, K. A., Kaya, D., *Role of Gilson–Pickering equation for the different types of soliton solutions: a nonlinear analysis*, The European Physical Journal Plus, 135(8), 1-19, 2020.
- [23] Tozar, A., Tasbozan, O., Kurt, A., *Optical soliton solutions for the (1+1)-dimensional resonant nonlinear Schrödinger's equation arising in optical fibers*, Optical and Quantum Electronics, 53(6), 1-8, 2021.
- [24] Rezazadeh, H., Kurt, A., Tozar, A., Tasbozan, O., Mirhosseini-Alizamini, S. M., *Wave behaviors of Kundu–Mukherjee–Naskar model arising in optical fiber communication systems with complex structure*, Optical and Quantum Electronics, 53(6), 1-11, 2021.
- [25] Yokus, A., Durur, H., Ahmad, H., Thounthong, P., Zhang, Y. F., *Construction of exact traveling wave solutions of the Bogoyavlenskii equation by (G'/G , $1/G$)-expansion and ($1/G$)-expansion techniques*, Results in Physics, 103409, 2020.
- [26] Yavuz, M., Sene, N., *Approximate solutions of the model describing fluid flow using generalized p -laplace transform method and heat balance integral method*, Axioms, 9(4), 123, 2020.
- [27] Duran, S., *Breaking theory of solitary waves for the Riemann wave equation in fluid dynamics*. International Journal of Modern Physics B, 35(9), 2150130, 2021.
- [28] Saleem, S., Hussain, M.Z., Aziz, I., *A reliable algorithm to compute the approximate solution of KdV-type partial differential equations of order seven*, Plos One, 16(1), e0244027, 2021.
- [29] Duran, S., Kaya, D., *Applications of a new expansion method for finding wave solutions of nonlinear differential equations*, World Applied Sciences Journal, 18(11), 1582-1592, 2012.
- [30] Kumar, M., Kumar, R., Kumar, A., *On similarity solutions of Zabolotskaya–Khokhlov equation*, Computers & Mathematics with Applications, 68(4), 454-463, 2014.
- [31] Zabolotskaya, E.A., Khokhlov, R.V., *Quasi-plane waves, in the nonlinear acoustics of confined beams*, Soviet Physics Acoustics, 15, 35-40, 1969.

[32] Akçağlı, Ş., Aydemir, T., *New exact solutions for the Khokhlov-Zabolotskaya-Kuznetsov, the Newell-Whitehead-Segel and the Rabinovich wave equations by using a new modification of the tanh-coth method*, *Cogent Mathematics*, 3(1), 1193104, 2016.

[33] Chirkunov, Y.A., Belmetsev, N.F., *Invariant submodels and exact solutions of Khokhlov-Zabolotskaya-Kuznetsov model of nonlinear hydroacoustics with dissipation*, *International Journal of Non-Linear Mechanics*, 95, 216-223, 2017.

[34] Kuznetsov, V.P., *Equations of nonlinear acoustics*, *Soviet Physics Acoustics*, 16, 467-470, 1971.

[35] Ray, S.S., *New analytical exact solutions of time fractional KdV-KZK equation by Kudryashov methods*, *Chinese Physics B*, 25(4), 040204, 2016.

[36] Zhang, L., Ji, J., Jiang, J., Zhang, C., *The new exact analytical solutions and numerical simulation of (3+ 1)-dimensional time fractional KZK equation*, *International Journal of Computing Science and Mathematics*, 10(2), 174-192, 2019.

[37] Demiray, Ş.T., Kastal, S., *MEFM For Exact Solutions Of The (3+1) Dimensional KZK Equation and (3+1) Dimensional JM Equation*, *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 21(1), 97-105, 2021.

[38] Duran, S., *Extractions of travelling wave solutions of (2+ 1)-dimensional Boiti-Leon-Pempinelli system via (G'/G, 1/G)-expansion method*, *Optical and Quantum Electronics*, 53(6), 1-12, 2021.



Linear Convex Combination Estimators and Comparisons

Selahattin KAÇIRANLAR^{1,*}, Issam A. DAWOUD², Dünya KARAPINAR³

¹ Çukurova University, Faculty of Science and Letters, Department of Statistics, Adana, Türkiye
skacir@cu.edu.tr, ORCID: 0000-0003-0678-7935

² Al-Aqsa University, Faculty of Applied Sciences, Department of Mathematics, Gaza, Palestina
isamdawoud@gmail.com, ORCID: 0000-0002-6733-7308

³ Karadeniz Technical University, Faculty of Science, Department of Mathematics, Trabzon, Türkiye
dunyakarapinar@ktu.edu.tr, ORCID: 0000-0001-5516-4780

Received: 31.05.2021

Accepted: 05.10.2021

Published: 31.12.2021

Abstract

In this paper, we introduce two linear convex combination estimators by using known estimators such as ordinary least squares, ridge and Liu estimators and examine the predictive performance of these estimators. Furthermore, a numerical example is examined to compare these estimators under the prediction mean squared error criterion.

Keywords: Biased estimation; Ridge estimator; Linear convex combination; Liu estimator; Prediction mean square error.

Lineer Konveks Kombinasyon Tahmin Ediciler ve Karşılaştırmalar

Öz

Bu makalede, en küçük kareler, ridge ve Liu tahmin ediciler gibi bilinen tahmin edicilerle öngörü performansını karşılaştırmak için iki lineer konveks kombinasyon tahmin edicisi



tanımlanmıştır. Ayrıca, öngörü hata kareleri ortalaması kriterine göre bu tahmin edicilerin karşılaştırılmaları bir sayısal örnek ile incelenmiştir.

Anahtar Kelimeler: Yanlı tahmin; Ridge tahmin edici; Lineer konveks kombinasyon; Liu tahmin edici; Öngörü hata kareleri ortalaması.

1. Introduction

Consider the following multiple linear regression model:

$$y = X\beta + \varepsilon, \quad (1)$$

where y is an $nx1$ vector of responses, X is an $n \times p$ full column rank matrix of explanatory variables, β is a $px1$ vector of unknown parameters, and ε is an $nx1$ vector of random errors with $iid(0, \sigma^2)$.

The ordinary least squares (OLS) estimator is given by

$$\hat{\beta} = (X'X)^{-1}X'y. \quad (2)$$

In the presence of multicollinearity, the OLS estimator is unstable and gives unreliable information. As biased alternatives, ridge, Liu, and two-parameter estimators can be handled in this context.

Hoerl and Kennard [1] proposed the ordinary ridge regression (ORR) estimator which is given by

$$\hat{\beta}(k) = (X'X + kI)^{-1}X'y, \quad k \geq 0, \quad (3)$$

where k is the biasing parameter. The ORR estimator was commonly used in applied researches. For example; Askin [2] suggested several approaches for extending estimation results to forecasting with multicollinearity, Montgomery and Friedman [3] examined several biased estimation methods for forecasting and prediction with multicollinearity.

Liu [4] defined the following alternative biased estimator dealing with multicollinearity

$$\begin{aligned} \hat{\beta}(d) &= (X'X + I)^{-1}(X'y + d\hat{\beta}) \\ &= (X'X + I)^{-1}(X'X + dI)\hat{\beta}, \quad 0 < d < 1, \end{aligned} \quad (4)$$

where d is the biasing parameter. $\hat{\beta}(d)$ is called the Liu estimator by Akdeniz and Kaçiranlar [5]. Liu estimator has an advantage over the ORR estimator because it is a linear function of d and it has smaller mean square error (MSE) than the OLS estimator. Sakallıoğlu et al. [6] compared the performance of Liu estimator with the ORR and the iterative estimators using the matrix MSE

(MMSE) criterion. In the literature, Liu and Liu-type estimators were widely used in linear models.

Furthermore, Özkale and Kaçiranlar [7] introduced a new two-parameter estimator (TPE) by grafting the contraction estimator into the modified ridge estimator proposed by Swindel [8]. This estimator is given by

$$\hat{\beta}(k, d) = (X'X + kI)^{-1}(X'y + kd\hat{\beta}), \quad k \geq 0, 0 < d < 1. \quad (5)$$

$\hat{\beta}(k, d)$ is a two-parameter variation of the Liu estimator. Özkale [9] has also noted that $\hat{\beta}(k, d)$ can also be demonstrated as

$$\hat{\beta}(k, d) = d\hat{\beta} + (1 - d)\hat{\beta}(k). \quad (6)$$

The TPE is a convex combination of the OLS and the ORR estimator. It is also called the ‘affine combination type’ estimator by Özkale [9]. Using the mixed estimation method suggested by Theil [10] and Theil and Goldberger [11], we also derive $\hat{\beta}(k, d)$. Similar to the ORR and Liu estimator, $\hat{\beta}(k, d)$ was used both theoretically and practically by researchers in various fields. Özbay and Kaçiranlar [12] introduced Almon TPE based on the TPE procedure for the distributed lag models. Özbay and Kaçiranlar [13] introduced a new two-parameter-weighted mixed estimator (TPWME) by unifying the weighted mixed estimator of Schaffrin and Toutenburg [14] and the TPE. Tekeli et al. [15] introduced new algorithms using genetic algorithm (GA) for estimating the biasing parameters of TPE. Çetinkaya and Kaçiranlar [16] introduced new TPE for negative binomial regression (NBR) and Poisson regression (PR) models by unifying the TPE.

Gruber [17, 18] demonstrated that $\hat{\beta}(k, d)$ is a special case of the linear Bayes, mixed and minimax estimators. This new estimator is a general estimator which includes the OLS, the ORR, the Liu, and the contraction estimators as special cases. We have the following properties:

1. $\lim_{d \rightarrow 1} \hat{\beta}(k, d) = \hat{\beta}$ and $\lim_{k \rightarrow 0} \hat{\beta}(k, d) = \hat{\beta}$
2. $\lim_{d \rightarrow 0} \hat{\beta}(k, d) = \hat{\beta}(k)$
3. For $k = 1$, we get the Liu estimator, $\hat{\beta}(1, d) = \hat{\beta}(d)$
4. $\hat{\beta}(k, d)$ has the following alternative forms

$$\begin{aligned} \hat{\beta}(k, d) &= [I + k(X'X)^{-1}]^{-1}(\hat{\beta} - d\hat{\beta}) + d\hat{\beta} \\ &= (X'X + kI)^{-1}(X'X + kdI)\hat{\beta}. \end{aligned}$$

From this representation, it is clear that $\lim_{k \rightarrow \infty} \hat{\beta}(k, d) = d\hat{\beta}$, which is the contraction estimator [19]. In this sense, $\hat{\beta}(k, d)$ overcomes the disadvantage of the contraction estimator.

Then, Gruber [18] demonstrated that the Liu-type estimator can be given as follows:

$$\hat{\beta}_{LOB} = d\hat{\beta} + (1 - d)\hat{\beta}_b, \quad (7)$$

where d is a biasing parameter, $0 < d < 1$ and $\hat{\beta}_b$ is the linear Bayes estimator (see in details, p. 3741, Eqn. (3.7), Eqn. (3.8) for $\hat{\beta}_b$ and p. 3742, Eqn. (3.12) for $\hat{\beta}_{LOB}$).

Gruber [18] showed how the Liu-type estimator is optimal according to the Zellner's balanced loss function (ZBLF) criterion and compared the efficiency of the Liu-type estimator to the OLS estimator in terms of the MSE and the ZBLF criteria. A convex combination of two estimators can be useful when both estimators appear to be appropriate in a specific situation. Following the Liu-type estimator in Eqn. (7), we consider linear convex combination estimators taking the ORR and the Liu estimators as the special cases of $\hat{\beta}_b$. Then, the linear convex combination of the OLS estimator and the ORR estimator (LOR) can be given as follows:

$$\hat{\beta}_{LOR} = \hat{\beta}(k, d) = d\hat{\beta} + (1 - d)\hat{\beta}(k), \quad k \geq 0, 0 < d < 1. \quad (8)$$

Similarly, we can define another linear convex combination of the OLS estimator and the Liu estimator as follows:

$$\hat{\beta}_{LOL} = \hat{\beta}(d, \gamma) = \gamma\hat{\beta} + (1 - \gamma)\hat{\beta}(d), \quad (9)$$

where γ is an arbitrary scalar and $0 \leq \gamma \leq 1$. Also, $\hat{\beta}_{LOL} = \hat{\beta}(d, \gamma)$ in Eqn. (9) includes $\hat{\beta}$ and $\hat{\beta}(d)$ as special cases:

1. $\lim_{\gamma \rightarrow 1} \hat{\beta}(d, \gamma) = \hat{\beta}$,
2. $\lim_{\gamma \rightarrow 0} \hat{\beta}(d, \gamma) = \hat{\beta}(d)$.

Friedman and Montgomery [20] compared the predictive performance (PP) of the ORR, OLS and the principal component (PC) estimators according to the prediction mean square error (PMSE) criterion. Later, Özbey and Kaçiranlar [21] compared the Liu estimator with the OLS, PC and ORR estimators. Dawoud and Kaçiranlar [22] examined the PP of biased regression predictors with correlated errors. Dawoud and Kaçiranlar [23, 24] evaluated the PP of the r-k and r-d class estimators and they also focused on evaluating the PP of the Liu-type estimator which is

defined by Liu [25]. This estimator is different from Gruber's Liu-type estimator which is given in Eqn. (7). Following Özbey and Kaçiranlar [21] and Dawoud and Kaçiranlar [22], Li et al. [26] evaluated the PP of the principal component two-parameter estimator which is defined by Chang and Yang [27].

As a consequence, since $\hat{\beta}_{LOR} = \hat{\beta}(k, d)$ and $\hat{\beta}_{LOL} = \hat{\beta}(d, \gamma)$ are more general than the ORR and the Liu estimators, respectively. Therefore, the PP of the LOR and the LOL estimators are examined in the sense of the PMSE criterion. To examine the theoretical results, a numerical example study is conducted.

2. Comparisons of the Prediction Mean Squared Errors

We can obtain the PMSE of the LOR and the LOL estimators. The PMSE of a predictor \hat{y}_0 is given by

$$PMSE = E(y_0 - \hat{y}_0)^2, \quad (10)$$

where y_0 is the value to be predicted. Let J represents the PMSE. J is the sum of the variance (V) and the squared bias (B):

$$J = V + B. \quad (11)$$

The variance and the bias can be given as follows:

$$V(y_0 - \hat{y}_0) = V(y_0) + V(\hat{y}_0), \quad (12)$$

and

$$Bias = E(y_0 - \hat{y}_0). \quad (13)$$

Now, we consider the following canonical form of the model (1)

$$y = Z\alpha + \varepsilon, \quad (14)$$

where $\alpha = U'\beta$ and $Z = XU$. Then the OLS estimator of α is

$$\hat{\alpha} = (Z'Z)^{-1}Z'y = \Lambda^{-1}Z'y, \quad (15)$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ is the matrix of the eigenvalues of $Z'Z$ and for $i = 1, 2, \dots, p$ λ_i 's are in descending order. Its PMSE is given by

$$J_{OLS} = \sigma^2 \left(1 + \sum_{i=1}^p \frac{z_{0i}^2}{\lambda_i} \right), \quad (16)$$

where z_0 is the orthonormalized point for \hat{y}_0 . Since $\hat{\alpha}$ is unbiased, we have

$$J_{OLS} = V_{OLS}. \quad (17)$$

The ridge estimator of α is

$$\hat{\alpha}_k = (Z'Z + kI)^{-1}Z'y = (\Lambda + kI)^{-1}Z'y, \quad k \geq 0, \quad (18)$$

and its PMSE is

$$J_k = \sigma^2 \left(1 + \sum_{i=1}^p \frac{z_{0i}^2 \lambda_i}{a_i^2} \right) + k^2 \left(\sum_{i=1}^p \frac{z_{0i} \alpha_i}{a_i} \right)^2, \quad (19)$$

where $a_i = \lambda_i + k$. The Liu estimator of α is

$$\begin{aligned} \hat{\alpha}_d &= (Z'Z + I)^{-1}(Z'y + d\hat{\alpha}) \\ &= (\Lambda + I)^{-1}(\Lambda + dI)\hat{\alpha}, \quad 0 < d < 1, \end{aligned} \quad (20)$$

and its PMSE is

$$J_d = \sigma^2 \left(1 + \sum_{i=1}^p \frac{z_{0i}^2 c_i^2}{\lambda_i b_i^2} \right) + (1-d)^2 \left(\sum_{i=1}^p \frac{z_{0i} \alpha_i}{b_i} \right)^2, \quad (21)$$

where $b_i = \lambda_i + 1$ and $c_i = \lambda_i + d$. The LOR estimator or TPE of α is

$$\begin{aligned} \hat{\alpha}_{LOR} &= [d(Z'Z)^{-1} + (1-d)(Z'Z + kI)^{-1}]Z'y \\ &= [d\Lambda^{-1} + (1-d)(\Lambda + kI)^{-1}]Z'y, \quad k \geq 0. \end{aligned} \quad (22)$$

The variance and bias of the prediction error of the LOR estimator are given by respectively

$$\begin{aligned} V_{LOR}(y_0 - \hat{y}_0) &= V(y_0) + V_{LOR}(\hat{y}_0) \\ &= \sigma^2 + V(z_0' \hat{\alpha}_{LOR}) \\ &= \sigma^2 \left(1 + \sum_{i=1}^p \frac{[(1-d)\lambda_i + da_i]^2 z_{0i}^2}{\lambda_i a_i^2} \right), \end{aligned} \quad (23)$$

$$\begin{aligned} Bias_{LOR} &= E(y_0 - \hat{y}_0) = z_0' \alpha - z_0' E(\hat{\alpha}_{LOR}) \\ &= k(1-d) \sum_{i=1}^p \frac{z_{0i} \alpha_i}{a_i}. \end{aligned} \quad (24)$$

So, the squared bias is

$$B_{LOR} = Bias_{LOR}^2 = k^2(1-d)^2 \left(\sum_{i=1}^p \frac{z_{0i} \alpha_i}{a_i} \right)^2. \quad (25)$$

By summing up the variance and the squared bias of the LOR estimator we obtain

$$\begin{aligned} J_{LOR} &= V_{LOR} + B_{LOR} \\ &= \sigma^2 \left(1 + \sum_{i=1}^p \frac{[(1-d)\lambda_i + da_i]^2 z_{0i}^2}{\lambda_i a_i^2} \right) + k^2(1-d)^2 \left(\sum_{i=1}^p \frac{z_{0i} \alpha_i}{a_i} \right)^2. \end{aligned} \quad (26)$$

The LOL estimator of α is

$$\begin{aligned} \hat{\alpha}_{LOL} &= [\gamma I + (1 - \gamma)(Z'Z + I)^{-1}(Z'Z + dI)]\hat{\alpha} \\ &= [\gamma I + (1 - \gamma)(\Lambda + I)^{-1}(\Lambda + dI)]\Lambda^{-1}Z'y, \quad 0 < d < 1. \end{aligned} \tag{27}$$

The variance of the prediction error of the LOL estimator is

$$\begin{aligned} V_{LOL}(y_0 - \hat{y}_0) &= V(y_0) + V_{LOL}(\hat{y}_0) \\ &= \sigma^2 + V(z_0'\hat{\alpha}_{LOL}) \\ &= \sigma^2 \left(1 + \sum_{i=1}^p \frac{[\gamma b_i + (1-\gamma)c_i]^2 z_{0i}^2}{\lambda_i b_i^2} \right). \end{aligned} \tag{28}$$

Similarly, the bias, the squared bias and PMSE of the prediction error of the LOL estimator are given by respectively

$$\begin{aligned} Bias_{LOL} &= E(y_0 - \hat{y}_0) = z_0'\alpha - z_0'E(\hat{\alpha}_{LOL}) \\ &= (1 - \gamma)(1 - d) \sum_{i=1}^p \frac{z_{0i}\alpha_i}{b_i}, \end{aligned} \tag{29}$$

$$B_{LOL} = Bias_{LOL}^2 = (1 - \gamma)^2(1 - d)^2 \left(\sum_{i=1}^p \frac{z_{0i}\alpha_i}{b_i} \right)^2, \tag{30}$$

and

$$\begin{aligned} J_{LOL} &= V_{LOL} + B_{LOL} \\ &= \sigma^2 \left(1 + \sum_{i=1}^p \frac{[\gamma b_i + (1-\gamma)c_i]^2 z_{0i}^2}{\lambda_i b_i^2} \right) + (1 - \gamma)^2(1 - d)^2 \left(\sum_{i=1}^p \frac{z_{0i}\alpha_i}{b_i} \right)^2. \end{aligned} \tag{31}$$

3. Comparisons of Prediction Mean Squared Errors in Two Dimensional Space

We will study the PP of the LOR and the LOL estimators. Considering a two-dimensional space, a single prediction point (z_{01}, z_{02}) is to be predicted, the ratio z_{02}^2/z_{01}^2 can be obtained and used for a reference point in their comparisons. α_1^2 will be set to zero because non-zero values of α_1^2 increase only the intercept values for J_k , J_d , J_{LOR} and J_{LOL} but leave the curve for J_{OLS} unchanged. So, comparisons of J_{LOR} with J_{OLS} and J_k and J_{LOL} with J_{OLS} and J_d will be made.

Theorem 1.

a) If $\alpha_2^2 > \frac{\sigma^2(a_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 (1-d)^2}$, then

$$-J_{LOR} < J_{OLS} \text{ for } a_1^2 < ((1-d)\lambda_1 + da_1)^2,$$

$$-J_{LOR} < J_{OLS} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2) \text{ for } a_1^2 > ((1-d)\lambda_1 + da_1)^2.$$

b) If $\alpha_2^2 < \frac{\sigma^2(a_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 (1-d)^2}$, then

$$-J_{LOR} < J_{OLS} \text{ for } a_1^2 > ((1-d)\lambda_1 + da_1)^2,$$

$$-J_{LOR} < J_{OLS} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2) \text{ for } a_1^2 < ((1-d)\lambda_1 + da_1)^2,$$

where

$$f_1(\alpha_2^2) = \frac{\sigma^2 \left(\frac{1}{\lambda_1} - \frac{((1-d)\lambda_1 + da_1)^2}{\lambda_1 a_1^2} \right)}{\left(\frac{\sigma^2 ((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2 (1-d)^2 \alpha_2^2 \sigma^2}{a_2^2 \lambda_2} \right)}. \tag{32}$$

Proof. If the LOR estimator is better than $\hat{\alpha}$, we have $J_{LOR} < J_{OLS}$. That is,

$$\begin{aligned} \sigma^2 + \sigma^2 \left[\frac{((1-d)\lambda_1 + da_1)^2 z_{01}^2}{\lambda_1 a_1^2} + \frac{((1-d)\lambda_2 + da_2)^2 z_{02}^2}{\lambda_2 a_2^2} \right] + \frac{k^2 (1-d)^2 \alpha_2^2 z_{02}^2}{a_2^2} < \\ \sigma^2 + \sigma^2 \left(\frac{z_{01}^2}{\lambda_1} + \frac{z_{02}^2}{\lambda_2} \right). \end{aligned}$$

Rearranging this inequality, we will obtain

$$z_{02}^2 \left(\frac{\sigma^2 ((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2 (1-d)^2 \alpha_2^2}{a_2^2} - \frac{\sigma^2}{\lambda_2} \right) < z_{01}^2 \sigma^2 \left(\frac{1}{\lambda_1} - \frac{((1-d)\lambda_1 + da_1)^2}{\lambda_1 a_1^2} \right).$$

If both

$$\frac{\sigma^2 ((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2 (1-d)^2 \alpha_2^2}{a_2^2} - \frac{\sigma^2}{\lambda_2} \tag{33}$$

and

$$\sigma^2 \left(\frac{1}{\lambda_1} - \frac{((1-d)\lambda_1 + da_1)^2}{\lambda_1 a_1^2} \right) \tag{34}$$

have the same signs, the superiority condition of the LOR estimator over $\hat{\alpha}$ is

$$\frac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2). \tag{35}$$

If Eqn. (33) and Eqn. (34) have opposite signs, we have

$$\frac{z_{02}^2}{z_{01}^2} > f_1(\alpha_2^2). \tag{36}$$

If Eqn. (33) and Eqn. (34) have different signs, the right-hand side of Eqn. (36) is smaller than zero, thus, Eqn. (36) always holds. That is, in this region the LOR estimator is superior to $\hat{\alpha}$. The condition for the positiveness of Eqn. (33) can be easily written as

$$\alpha_2^2 > \frac{\sigma^2(a_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 (1-d)^2} \tag{37}$$

and the condition for the positiveness of Eqn. (34) can be given as

$$\alpha_1^2 > ((1-d)\lambda_1 + da_1)^2. \tag{38}$$

The contrary conditions are required for the negativeness of Eqn. (33) and Eqn. (34). The vertical asymptote of the hyperbola $f_1(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2(a_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 (1-d)^2}. \tag{39}$$

Corollary 1. If $d = 0$ in Theorem 1, we get Friedman and Montgomery’s [20] results.

Corollary 2. If $k = 1$ in Theorem 1, we get Özbey and Kaçıranlar’s [21] results.

Theorem 2.

a) If $\alpha_2^2 > \frac{\sigma^2(\lambda_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 [(1-d)^2 - 1]}$, then

$$-J_{LOR} < J_k \text{ for } \lambda_1^2 < ((1-d)\lambda_1 + da_1)^2,$$

$$-J_{LOR} < J_k \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2) \text{ for } \lambda_1^2 > ((1-d)\lambda_1 + da_1)^2.$$

b. If $\alpha_2^2 < \frac{\sigma^2(\lambda_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 [(1-d)^2 - 1]}$, then

$$-J_{LOR} < J_k \text{ for } \lambda_1^2 > ((1-d)\lambda_1 + da_1)^2,$$

$$-J_{LOR} < J_k \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2) \text{ for } \lambda_1^2 < ((1-d)\lambda_1 + da_1)^2.$$

where

$$f_2(\alpha_2^2) = \frac{\sigma^2 \left(\frac{\lambda_1 \cdot ((1-d)\lambda_1 + da_1)^2}{a_1^2 \lambda_1 a_1^2} \right)}{\left(\frac{\sigma^2 ((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2 (1-d)^2 \alpha_2^2 \sigma^2 \lambda_2}{a_2^2} \cdot \frac{k^2 \alpha_2^2}{a_2^2} \right)}. \tag{40}$$

Proof. Suppose LOR estimator is better than $\hat{\alpha}_k$, then, $J_{LOR} < J_k$. That is,

$$\sigma^2 + \sigma^2 \left[\frac{((1-d)\lambda_1 + da_1)^2 z_{01}^2}{\lambda_1 a_1^2} + \frac{((1-d)\lambda_2 + da_2)^2 z_{02}^2}{\lambda_2 a_2^2} \right] + \frac{k^2 (1-d)^2 \alpha_2^2 z_{02}^2}{a_2^2} <$$

$$\sigma^2 + \sigma^2 \left(\frac{\lambda_1 z_{01}^2}{a_1^2} + \frac{\lambda_2 z_{02}^2}{a_2^2} \right) + \frac{k^2 \alpha_2^2 z_{02}^2}{a_2^2}.$$

Rearranging this inequality, we get

$$z_{02}^2 \left(\frac{\sigma^2((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2(1-d)^2 \alpha_2^2}{a_2^2} - \frac{\sigma^2 \lambda_2}{a_2^2} - \frac{k^2 \alpha_2^2}{a_2^2} \right) < z_{01}^2 \sigma^2 \left(\frac{\lambda_1}{a_1^2} - \frac{((1-d)\lambda_1 + da_1)^2}{\lambda_1 a_1^2} \right).$$

If both

$$\left(\frac{\sigma^2((1-d)\lambda_2 + da_2)^2}{\lambda_2 a_2^2} + \frac{k^2(1-d)^2 \alpha_2^2}{a_2^2} - \frac{\sigma^2 \lambda_2}{a_2^2} - \frac{k^2 \alpha_2^2}{a_2^2} \right) \tag{41}$$

and

$$\sigma^2 \left(\frac{\lambda_1}{a_1^2} - \frac{((1-d)\lambda_1 + da_1)^2}{\lambda_1 a_1^2} \right) \tag{42}$$

have the same signs, we have

$$\frac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2). \tag{43}$$

If Eqn. (41) and Eqn. (42) have opposite signs, we have

$$\frac{z_{02}^2}{z_{01}^2} > f_2(\alpha_2^2). \tag{44}$$

If Eqn. (41) and Eqn. (42) have opposite signs, the right-hand side of Eqn. (44) is negative, so, Eqn. (44) always holds. The condition for the positiveness of Eqn. (41) can be written as

$$\alpha_2^2 > \frac{\sigma^2(\lambda_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 [(1-d)^2 - 1]}. \tag{45}$$

The condition for the positiveness of Eqn. (42) can be given as

$$\lambda_1^2 > ((1-d)\lambda_1 + da_1)^2. \tag{46}$$

The contrary conditions are required for the negativeness of Eqn. (41) and Eqn. (42). The vertical asymptote of the hyperbola $f_2(\alpha_2^2)$ is

$$\alpha_2^2 = \frac{\sigma^2(\lambda_2^2 - ((1-d)\lambda_2 + da_2)^2)}{\lambda_2 k^2 [(1-d)^2 - 1]}. \tag{47}$$

Theorem 3.

a) If $\alpha_2^2 > \frac{\sigma^2(b_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-\gamma)^2(1-d)^2}$, then

$$-J_{LOL} < J_{OLS} \text{ for } b_1^2 < (\gamma b_1 + (1-\gamma)c_1)^2,$$

$$-J_{LOL} < J_{OLS} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2) \text{ for } b_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2.$$

b) If $\alpha_2^2 < \frac{\sigma^2(b_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-\gamma)^2(1-d)^2}$, then

$$-J_{LOL} < J_{OLS} \text{ for } b_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2,$$

$$-J_{LOL} < J_{OLS} \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2) \text{ for } b_1^2 < (\gamma b_1 + (1-\gamma)c_1)^2,$$

where

$$f_3(\alpha_2^2) = \frac{\sigma^2 \left(\frac{1}{\lambda_1} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2} \right)}{\left(\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2 \sigma^2}{b_2^2 \lambda_2} \right)} \tag{48}$$

Proof. If the LOL estimator is superior to $\hat{\alpha}$, we have $J_{LOL} < J_{OLS}$. That is,

$$\begin{aligned} \sigma^2 + \sigma^2 \left[\frac{(\gamma b_1 + (1-\gamma)c_1)^2 z_{01}^2}{\lambda_1 b_1^2} + \frac{(\gamma b_2 + (1-\gamma)c_2)^2 z_{02}^2}{\lambda_2 b_2^2} \right] + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2 z_{02}^2}{b_2^2} \\ < \\ \sigma^2 + \sigma^2 \left(\frac{z_{01}^2}{\lambda_1} + \frac{z_{02}^2}{\lambda_2} \right). \end{aligned}$$

Rearranging this inequality, we get

$$z_{02}^2 \left(\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2}{b_2^2} - \frac{\sigma^2}{\lambda_2} \right) < z_{01}^2 \sigma^2 \left(\frac{1}{\lambda_1} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2} \right).$$

If both

$$\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2}{b_2^2} - \frac{\sigma^2}{\lambda_2} \tag{49}$$

and

$$\sigma^2 \left(\frac{1}{\lambda_1} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2} \right) \tag{50}$$

have the same signs, we have

$$\frac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2). \tag{51}$$

If Eqn. (49) and Eqn. (50) have opposite signs, we have

$$\frac{z_{02}^2}{z_{01}^2} > f_3(\alpha_2^2). \tag{52}$$

If Eqn. (49) and Eqn. (50) have opposite signs, the right-hand side of Eqn. (52) is negative, thus Eqn. (52) always holds. The condition for the positiveness of Eqn. (49) can be written as

$$\alpha_2^2 > \frac{\sigma^2(b_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-\gamma)^2(1-d)^2}. \tag{53}$$

Similarly, the condition for the positiveness of Eqn. (50) can be given as

$$b_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2. \tag{54}$$

The contrary conditions are required for the negativeness of Eqn. (49) and Eqn. (50). The vertical asymptote of the hyperbola $f_3(\alpha_2^2)$ is at the point

$$\alpha_2^2 = \frac{\sigma^2(b_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-\gamma)^2(1-d)^2}. \tag{55}$$

Corollary 3: If $\gamma = 0$ in Theorem 3, we get Özbey and Kaçıranlar’s [21] results.

Theorem 4.

a) If $\alpha_2^2 > \frac{\sigma^2(c_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-d)^2[(1-\gamma)^2 - 1]}$, then

$$-J_{LOL} < J_d \text{ for } c_1^2 < (\gamma b_1 + (1-\gamma)c_1)^2,$$

$$-J_{LOL} < J_d \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2) \text{ for } c_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2.$$

b) If $\alpha_2^2 < \frac{\sigma^2(c_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-d)^2[(1-\gamma)^2 - 1]}$, then

$$-J_{LOL} < J_d \text{ for } c_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2,$$

$$-J_{LOL} < J_d \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2) \text{ for } c_1^2 < (\gamma b_1 + (1-\gamma)c_1)^2,$$

where

$$f_4(\alpha_2^2) = \frac{\sigma^2\left(\frac{c_1^2}{\lambda_1 b_1^2} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2}\right)}{\left(\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2}{b_2^2} - \frac{\sigma^2 c_2^2}{\lambda_2 b_2^2} - \frac{(1-d)^2 \alpha_2^2}{b_2^2}\right)}. \tag{56}$$

Proof. If the LOL estimator is superior to $\hat{\alpha}_d$, we have $J_{LOL} < J_d$. That is,

$$\sigma^2 + \sigma^2 \left[\frac{(\gamma b_1 + (1-\gamma)c_1)^2 z_{01}^2}{\lambda_1 b_1^2} + \frac{(\gamma b_2 + (1-\gamma)c_2)^2 z_{02}^2}{\lambda_2 b_2^2} \right] + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2 z_{02}^2}{b_2^2} < \sigma^2 + \sigma^2 \left(\frac{c_1^2 z_{01}^2}{\lambda_1 b_1^2} + \frac{c_2^2 z_{02}^2}{\lambda_2 b_2^2} \right) + \frac{(1-d)^2 \alpha_2^2 z_{02}^2}{b_2^2}.$$

Rearranging this inequality, we get

$$z_{02}^2 \left(\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2}{b_2^2} - \frac{\sigma^2 c_2^2}{\lambda_2 b_2^2} - \frac{(1-d)^2 \alpha_2^2}{b_2^2} \right) <$$

$$z_{01}^2 \sigma^2 \left(\frac{c_1^2}{\lambda_1 b_1^2} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2} \right).$$

If both

$$\left(\frac{\sigma^2(\gamma b_2 + (1-\gamma)c_2)^2}{\lambda_2 b_2^2} + \frac{(1-\gamma)^2(1-d)^2 \alpha_2^2}{b_2^2} - \frac{\sigma^2 c_2^2}{\lambda_2 b_2^2} - \frac{(1-d)^2 \alpha_2^2}{b_2^2} \right) \quad (57)$$

and

$$\sigma^2 \left(\frac{c_1^2}{\lambda_1 b_1^2} - \frac{(\gamma b_1 + (1-\gamma)c_1)^2}{\lambda_1 b_1^2} \right) \quad (58)$$

have the same signs, we have

$$\frac{z_{02}^2}{z_{01}^2} < f_4(\alpha_2^2). \quad (59)$$

If Eqn. (55) and Eqn. (56) have opposite signs, we have

$$\frac{z_{02}^2}{z_{01}^2} > f_4(\alpha_2^2). \quad (60)$$

If Eqn. (55) and Eqn. (56) have opposite signs, the right-hand side of Eqn. (58) is negative, thus Eqn. (58) holds. The condition for the positiveness of Eqn. (55) can be given as follows

$$\alpha_2^2 > \frac{\sigma^2(c_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-d)^2[(1-\gamma)^2 - 1]}. \quad (61)$$

Similarly, the condition for the positiveness of Eqn. (56) can be given as

$$c_1^2 > (\gamma b_1 + (1-\gamma)c_1)^2. \quad (62)$$

The contrary conditions are required for the negativeness of Eqn. (55) and Eqn. (56). The vertical asymptote of the hyperbola $f_4(\alpha_2^2)$ is

$$\alpha_2^2 = \frac{\sigma^2(c_2^2 - (\gamma b_2 + (1-\gamma)c_2)^2)}{\lambda_2(1-d)^2[(1-\gamma)^2 - 1]}. \quad (63)$$

The estimation of the parameters k and d is an important issue. We have not made any attempt to estimate them. However, we refer our readers to Hoerl and Kennard [1], Kibria [28], Khalaf and Shukur [29], Muniz and Kibria [30] and Liu [4] among others.

4. Numerical Example

In this section, we will illustrate theoretical results using the example given by Friedman and Montgomery [20] (i.e., $\sigma^2 = 1$, $k = 0.1$ and $r_{12} = 0.95$) and Özbey and Kaçiranlar [21] (i.e., $d = 0.9$) as well as we let $\gamma = 0.5$.

Let us consider the LOR and the OLS estimators. From Eqn. (32), we get

$$f_1(\alpha_2^2) = \frac{0.004991}{0.004444\alpha_2^2 - 2.57778}, \quad (64)$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 580. \quad (65)$$

Because of both z_{02}^2/z_{01}^2 and α_2^2 are positive, we are interested only in the points which lie in the first quadrant. Figure 1 illustrates this situation. For values of α_2^2 smaller than 580, the LOR estimator is better than the OLS estimator. For larger values of α_2^2 , there is a trade-off between these two estimators. If the value of the ratio z_{02}^2/z_{01}^2 is smaller than the value of $f_1(\alpha_2^2)$, then the LOR estimator is superior to the OLS estimator; otherwise, the OLS estimator is better than the LOR estimator. We take different values of d as 0.1, 0.2, ..., 0.9 to determine the effect of d on the predictive performance of the LOR estimator and the OLS estimator. Table 1 shows that if d increases, the value of α_2^2 increases. That means, when α_2^2 increases, the region where the LOR estimator is uniformly superior to the OLS estimator increases.

In this part, we get the same results of the example given by Friedman and Montgomery [20] if $d = 0$. Also, we get the same results of the example given by Özbey and Kaçiranlar [21] if $k = 1$.

Let us consider the ORR and the LOR estimators. From Eqn. (40) and Eqn. (47), we get

$$f_2(\alpha_2^2) = \frac{0.04382}{0.44\alpha_2^2 - 15.2}, \quad (66)$$

and

$$\alpha_2^2 = 34.54. \quad (67)$$

Figure 2 shows this case. For values of $\alpha_2^2 < 34.54$, the LOR estimator is better than $\hat{\alpha}_k$. For great values of α_2^2 there is a trade-off between these estimators. If $(z_{02}^2/z_{01}^2) < f_2(\alpha_2^2)$, then the LOR estimator is superior to $\hat{\alpha}_k$, otherwise $\hat{\alpha}_k$ is better than the LOR estimator.

The effect of d on the PP of the LOR estimator and $\hat{\alpha}_k$ is described in Table 2. Table 2 shows that if d increases, the value of α_2^2 increases. That means, when α_2^2 increases, the region where the LOR estimator is better than $\hat{\alpha}_k$ increases.

Let us take into account the PP of the OLS and the LOL estimators. From Eqn. (48) and Eqn. (55), we have

$$f_3(\alpha_2^2) = \frac{0.017236}{0.00226\alpha_2^2 - 1.8595}, \quad (68)$$

and

$$\alpha_2^2 = 820. \tag{69}$$

Figure 3 shows this situation. For values of $\alpha_2^2 < 820$, the LOL estimator is uniformly superior to the $\hat{\alpha}$. If $(z_{02}^2/z_{01}^2) < f_3(\alpha_2^2)$, then the LOL estimator is better than $\hat{\alpha}$. Otherwise, $\hat{\alpha}$ is better than the LOL estimator.

The effect of γ on the PP of the LOL estimator and $\hat{\alpha}$ is described in Table 3. Table 3 shows that if γ increases, the value of α_2^2 increases. That means, when α_2^2 increases, the region where the LOL estimator is uniformly superior to $\hat{\alpha}$ increases.

In this part, we get the same results of the example given by Özbey and Kaçıranlar [21] if $\gamma = 0$.

Let us examine the PP of the LOL and $\hat{\alpha}_d$. From Eqn. (56) and Eqn. (63), we have

$$f_4(\alpha_2^2) = \frac{0.01694}{0.00681\alpha_2^2 - 1.7687}, \tag{70}$$

and

$$\alpha_2^2 = 260. \tag{71}$$

Figure 4 shows this case. For values of $\alpha_2^2 < 260$, the LOL estimator is superior to $\hat{\alpha}_d$. If $(z_{02}^2/z_{01}^2) < f_4(\alpha_2^2)$, then the LOL estimator is superior to $\hat{\alpha}_d$; otherwise, $\hat{\alpha}_d$ is superior to the LOL estimator.

The effect of γ on the PP of the LOL estimator and $\hat{\alpha}_d$ is described in Table 4. Table 4 shows that if γ increases, the value of α_2^2 increases. That means, when α_2^2 increases, the region where the LOL estimator is uniformly superior to $\hat{\alpha}_d$ increases.

Table 1. d and α_2^2 values for the LOR vs. the OLS

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	46.67	55.00	65.71	80.00	100.00	130.00	180.00	280.00	580.00

Table 2. d and α_2^2 values for the LOR vs. the ORR

d	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	11.58	13.33	15.29	17.50	20.00	22.86	26.15	30.00	34.54

Table 3. γ and α_2^2 values for the LOL vs. the OLS

γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	446.667	505.000	580.000	680.000	820.000	1030.000	1380.000	2080.000	4180.000

Table 4. γ and α_2^2 values for the LOL vs. the Liu

γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α_2^2	201.053	213.333	227.059	242.500	260.000	280.000	303.077	330.000	361.818

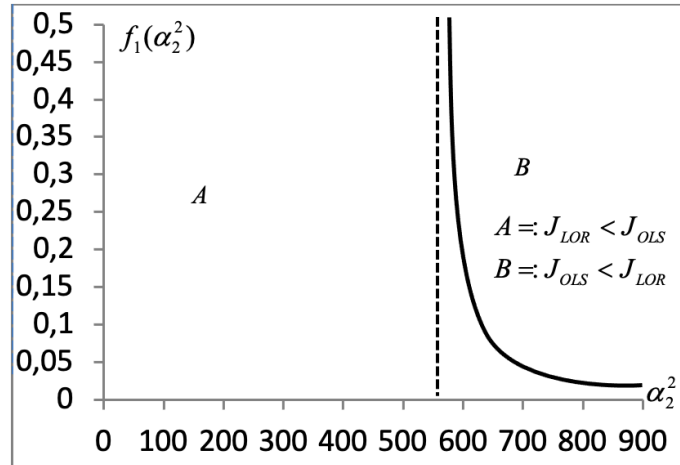


Figure 1: Comparison of the PMSE for LOR and OLS estimators

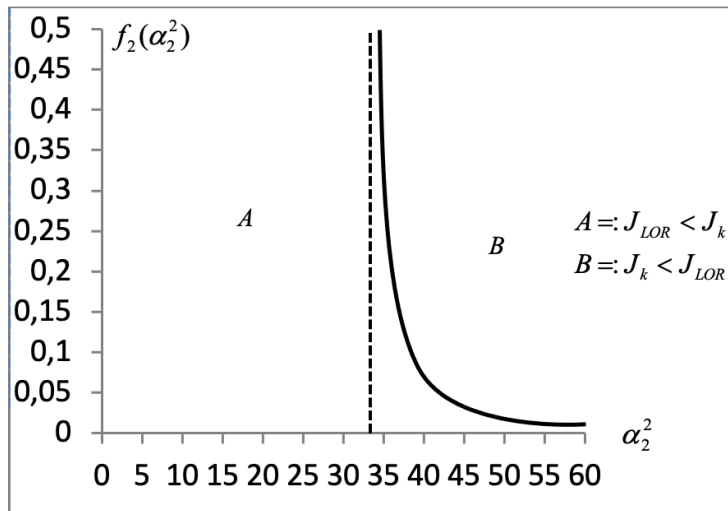


Figure 2: Comparison of the PMSE for LOR and ORR estimators

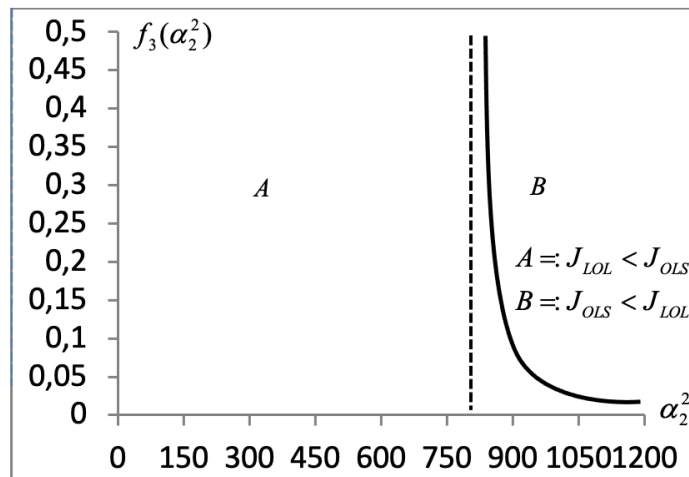


Figure 3: Comparison of the PMSE for LOL and OLS estimators

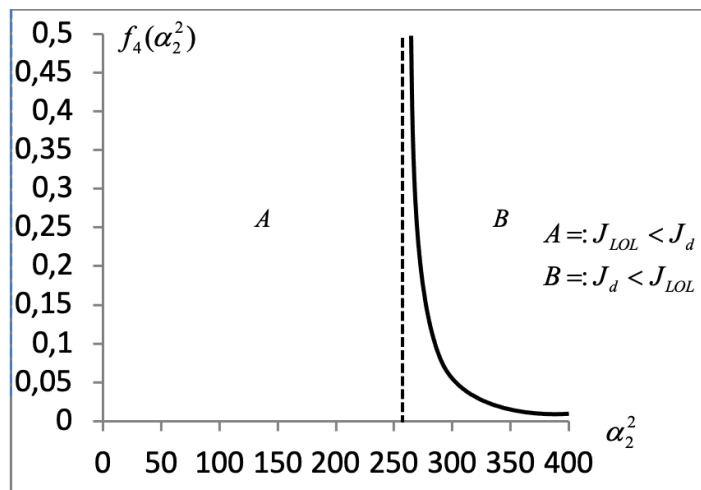


Figure 4: Comparison of the PMSE for LOL and Liu estimators

5. Conclusion

The predictive performance of the LOR estimator over the OLS and the ORR estimators is evaluated. Similarly, the predictive performance of the proposed LOL estimator over the OLS and the Liu estimators is examined in the sense of the PMSE. The comparisons of these estimators are in terms of the PMSE criterion at a specific point in the two-dimensional regressor variable spaces. In this context, the PMSE of the LOR and the LOL estimators are developed and four theorems are given. In addition, three corollaries are given here examining that the theorems given by Friedman and Montgomery [20] and Özbey and Kaçıranlar [21] are just special cases of the Theorems 1 and 3.

References

- [1] Hoerl, A.E., Kennard, R.W., *Ridge regression: biased estimation for nonorthogonal problems*, *Technometrics*, 12, 55–76, 1970.
- [2] Askın, R.G., *Multicollinearity in regression: Review and examples*, *Journal of Forecasting*, 1, 281-292, 1982.
- [3] Montgomery, D.C., Friedman, D.J., *Prediction using regression models with multicollinear predictor variables*, *IIE Transactions*, 2 (3), 73-85, 1993.
- [4] Liu, K., *A new class of biased estimate in linear regression*, *Communications in Statistics- Theory and Methods*, 22, 393-402, 1993.
- [5] Akdeniz, F., Kaçiranlar, S., *On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE*, *Communications in Statistics-Theory and Methods*, 24, 1789-1797, 1995.
- [6] Sakallıoğlu, S., Kaçiranlar, S., Akdeniz, F., *Mean squared error comparisons of some biased regression estimator*, *Communications in Statistics-Theory and Methods*, 30, 347-361, 2001.
- [7] Özkale, M.R., Kaçiranlar, S., *The restricted and unrestricted two parameter estimators*, *Communications in Statistics- Theory and Methods*, 36, 2707–2725, 2007.
- [8] Swindel, B.F., *Good ridge estimators based on prior information*, *Communications in Statistics-Theory and Methods*, A5, 1065–1075, 1976.
- [9] Özkale, M.R., *Influence measures in affine combination type regression*, *Journal of Applied Statistics*, 40, 2219-2243, 2013.
- [10] Theil, H., *On the use of incomplete prior information in regression analysis*, *Journal of American Statistical Association*, 58, 401–414, 1963.
- [11] Theil, H., Goldberger, A.S., *On pure and mixed statistical estimation in economics*, *International Economic Review*, 2, 65–78, 1961.
- [12] Özbay, N., Kaçiranlar, S., *The Almon two parameter estimator for the distributed lag models*, *Journal of Statistical Computation and Simulation*, 87, 834–843, 2017.
- [13] Özbay, N., Kaçiranlar, S., *Estimation in a linear regression model with stochastic linear restrictions: a new two parameter-weighted mixed estimator*, *Journal of Statistical Computation and Simulation*, 88, 1669–1683, 2018.
- [14] Schaffrin, B., Toutenburg, H., *Weighted mixed regression*, *Zeitschrift für Angewandte Mathematik und Mechanik*, 70, 735-738, 1990.
- [15] Tekeli, E., Kaçiranlar, S., Özbay, N., *Optimal determination of the parameters of some biased estimators using genetic algorithm*, *Journal of Statistical Computation and Simulation*, 89 (18), 3331-3353, 2019.
- [16] Çetinkaya, M.K., Kaçiranlar, S., *Improved two parameter estimators for the negative binomial and Poisson regression models*, *Journal of Statistical Computation and Simulation*, 89(14), 2645-2660, 2019.
- [17] Gruber, M.H.J., *Regression estimators: A comparative study*. 2nd ed., John Hopkins University Press, Baltimore, MD; 2010.
- [18] Gruber, M.H.J., *Liu and ridge estimators- A comparison*, *Communications in Statistics- Theory and Methods*, 41, 3739-3749, 2012.

- [19] Mayer, L.S., Willke, T.A., *On biased estimation in linear models*, *Technometrics*, 15, 497–508, 1973.
- [20] Friedman, D.J., Montgomery, D.C., *Evaluation of the predictive performance of biased regression estimators*, *Journal of Forecasting*, 4, 153-163, 1985.
- [21] Özbey, F., Kaçıranlar, S., *Evaluation of the predictive performance of the Liu estimator*, *Communications in Statistics-Theory and Methods*, 44, 1981-1993, 2015.
- [22] Dawoud, I., Kaçıranlar, S., *The predictive performance evaluation of biased regression predictors with correlated errors*, *Journal of Forecasting*, 34, 364-378, 2015.
- [23] Dawoud, I., Kaçıranlar, S., *Evaluation of the predictive performance of the r - k and r - d class estimators*, *Communications in Statistics-Theory and Methods*, 46, 4031-4050, 2017.
- [24] Dawoud, I., Kaçıranlar, S., *Evaluation of the predictive performance of the Liu-type estimator*, *Communications in Statistics–Simulation and Computation*, 46, 2800-2820, 2017.
- [25] Liu, K., *Using Liu-type estimator to combat collinearity*, *Communications in Statistics- Theory and Methods*, 32, 1009–1020, 2003.
- [26] Li, R., Li, F., Huang, J., *Evaluation of the predictive performance of the principal component two-parameter estimator*, *Concurrency and Computation Practice and Experience*, 31, e4710, 2019. <https://doi.org/10.1002/cpe.4710>.
- [27] Chang, X., Yang, H., *Combining two-parameter and principal component regression estimator*, *Statistical Papers*, 53(3), 549-562, 2012.
- [28] Kibria, B.M.G., *Performance of some new ridge regression estimators*, *Communications in Statistics–Simulation and Computation*, 32, 419-435, 2003.
- [29] Khalaf, G., Shukur, G., *Choosing ridge parameters for regression problems*, *Communications in Statistics-Theory and Methods*, 34, 1177-1182, 2005.
- [30] Muniz, G., Kibria, B.M.G., *On some ridge regression estimators: An empirical comparison*, *Communications in Statistics–Simulation and Computation*, 38, 621-630, 2009.



Identification of Damaging SNPs and Their Effects on Alzheimer's Disease-Associated PSEN1 Protein: Computational Analysis

Orcun AVSAR^{1,*}

¹Hitit University, Faculty of Arts and Sciences, Department of Molecular Biology and Genetics, 19030, Corum, Türkiye

orcunavsar@hitit.edu.tr, ORCID: 0000-0003-3556-6218

Received: 22.02.2021

Accepted: 12.10.2021

Published: 30.12.2021

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disease and pathologically characterized by the presence of neurofibrillary tangles (tau aggregation) and amyloid plaques (amyloid-beta ($A\beta$) aggregation). PSEN1 protein with 9 transmembrane helices acts as a serine protease and is one of the catalytic components of γ secretase complex, that cleaves amyloid precursor protein (APP). Furthermore, PSEN1 protein plays a significant role in the process of APP and in the generation of amyloid beta ($A\beta$). In the present study, it was aimed to estimate the probable deleterious effects of missense SNPs in *PSEN1* gene that is associated with AD on protein stability and structure by using bioinformatics tools. SIFT, PolyPhen-2, PROVEAN, PhD-SNP, and PANTHER PSEP software were used to estimate the deleterious SNPs, whereas I-Mutant 3.0 and MUpro web tools were used to determine the effects of amino acid substitution on protein stability. Additionally, the effects of wild type and mutant amino acids on protein three-dimensional structure via modeling were predicted by Project HOPE webserver. The phylogenetic conservation of amino acid residues of PSEN1 protein was analyzed by ConSurf. In total, 386 missense SNPs were found in the human PSEN1 gene from the National Center for Biotechnology Information Single Nucleotide Polymorphism (NCBI dbSNP) database and 65 SNPs of which

* Corresponding Author

DOI: 10.37094/adyujsci.884889



were determined to be deleterious or damaging. In the present study, 8 significant missense SNPs- rs63749891 (R278T), rs63750301 (P264L), rs63750353 (N135D), rs63750524 (R278S), rs63750772 (E273A), rs63751229 (P267S), rs121917807 (G266S), and rs201617677 (R157S)- were determined as high-risk pathogenic. Some differences between wild-type amino acids and mutant amino acids such as hydrophobicity, charge, size, and folding properties were determined according to the modeling findings. Our study demonstrates that high-risk pathogenic missense SNPs have the potential to alter the catalytic activity of the γ secretase complex and subsequently the amount of A β 40 and A β 42. Therefore, these missense SNPs may contribute to AD pathogenesis studies.

Keywords: Alzheimer's Disease; PSEN1; Gene; Mutation; SNP.

Zarar Verici SNP'lerin ve Alzheimer Hastalığıyla İlişkili PSEN1 Proteinine Etkilerinin Tanımlanması: Hesaplamalı Analiz

Öz

Alzheimer Hastalığı (AH), progresif nörodejeneratif hastalıktır ve patolojik olarak nörofibriler yumaklar (tau agregasyonu) ve amiloid plakların (amiloid beta ($A\beta$) agregasyonu) varlığı ile karakterize edilir. 9 transmembran heliks içeren PSEN1 proteini, aspartil proteaz olarak işlev görmektedir ve amiloid öncü proteini (APP) parçalayan γ sekretaz kompleksinin katalitik bileşenlerinden biridir. Ayrıca, PSEN1 proteini APP sürecinde ve amiloid beta ($A\beta$) oluşumunda önemli rol oynamaktadır. Bu çalışmada, AH ile ilişkili *PSEN1* genindeki missense (yanlış anlamlı) SNP'lerin protein stabilitesi ve yapısı üzerindeki olası zararlı etkilerinin biyoinformatik araçlar kullanılarak tahmin edilmesi amaçlanmıştır. Zararlı SNP'lerin tahmin edilmesinde SIFT, PolyPhen-2, PROVEAN, PhD-SNP ve PANTHER PSEP yazılımları kullanılırken, amino asit değişiminin protein stabilitesi üzerindeki etkilerini belirlemek için I-Mutant 3.0 ve MUpro web araçları kullanıldı. Ek olarak, yabani tip ve mutant amino asitlerin proteinin üç boyutlu yapısı üzerindeki etkileri ise modelleme yoluyla Project HOPE programı ile tahmin edilmiştir. PSEN1 proteininin amino asit kalıntılarının filogenetik korunumu ConSurf ile analiz edildi. NCBI dbSNP veritabanında insan *PSEN1* geninde toplam 386 missense SNP bulunduğu ve 65 SNP'nin ise zararlı veya zarar verici olduğu belirlendi. Bu çalışmada, 8 önemli missense SNP- rs63749891 (R278T), rs63750301 (P264L), rs63750353 (N135D), rs63750524 (R278S), rs63750772 (E273A), rs63751229 (P267S), rs121917807 (G266S), ve rs201617677 (R157S)- yüksekli riskli patojenik olarak belirlendi. Yabani tip ve mutant amino asitler arasındaki hidrofobiklik, yük, boyut ve katlanma özellikleri gibi bazı farklılıklar modelleme bulgularına göre belirlenmiştir. Çalışmamız, yüksek riskli patojenik missense SNP'lerin γ sekretaz kompleksinin katalitik

aktivitesini ve akabinde A β 40 ve A β 42 miktarını deęiřtirme potansiyelinin olduęunu gstermektedir. Bu nedenle, bu missense SNP'ler, AH patogenez alıřmalarına katkı saęlayabilir.

Anahtar Kelimeler: Alzheimer Hastalıęı; PSEN1; Gen; Mutasyon; SNP.

1. Introduction

Alzheimer's Disease (AD) that is the most frequent form of dementia in western populations is a multifactorial disease with a robust genetic background [1]. It is estimated that 65.7 million people worldwide will have AD in 2030 and 115.4 million in 2050. AD is pathologically characterized by the presence of neurofibrillary tangles (tau aggregation) and amyloid plaques (amyloid-beta (A β) aggregation). Consequently, these pathological hallmarks cause to disruption of synaptic transmission, neuronal cell death, and cognitive deficits [2]. It is known that the heritability of AD is approximately 60-80% and the genetic mechanisms of AD remain unclear and have to be elucidated [3].

Presenilin 1 (PSEN1, OMIM: 104311) gene that is localized on chromosome 14 with 12 exons is involved in the pathogenesis of AD. PSEN1 protein with 9 transmembrane helices acts as aspartyl protease and is one of the catalytic components of γ secretase complex, that cleaves amyloid precursor protein (APP). PSEN1 protein takes a significant role in the process of APP and in the generation of amyloid beta (A β). PSEN1 as a component of γ secretase complex is implicated in several neurobiological processes such as survival of neurons, memory, and synapse formation [4]. *PSEN1* mutations are involved in the most frequent form of inherited AD and are 100% penetrant [5].

Single nucleotide polymorphisms (SNPs) are one of the most common type of genetic variations in human genome and are used as molecular markers for genetic studies. The SNPs that are located in non-coding RNA and protein-coding genes are classified as functional and neutral. The functional SNPs affect numerous biological pathways and functions and are risky for multifactorial diseases such as Alzheimer's Disease, whereas neutral SNPs do not affect biological processes [6]. Determination of the deleterious effects of SNPs by *in silico* methods contributes to identify the SNPs associated with diseases and then to design the genotyping studies according to the findings of the *in silico* studies. In a study, the identification of the deleterious/damaging effects of missense SNPs in PSEN1 gene was investigated by molecular dynamics (MD) simulations, it has been reported that no significant association was found between the structure and function of PSEN1 protein and deleterious SNPs [7].

In the current study, we aimed to identify missense SNPs in Alzheimer's disease-associated *PSEN1* gene to investigate the probable pathogenic effects of SNPs on several properties of the amino acid residues of PSEN1 protein such as stability, charge, hydrophobicity, size, folding, structure, and evolutionary conservation by using bioinformatics tools.

2. Material and Methods

2.1. Determination of gene-gene interactions

Gene-gene functional interactions of *PSEN1* gene were analyzed by GeneMANIA database. The GeneMANIA Cytoscape app provides researchers to determine the network of gene-gene interactions. The generated network involves the genes most associated with the investigated gene or genes and functional annotations from Gene Ontology. The database contains more than 500 million interactions spanning eight organisms: *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Danio rerio*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Drosophila melanogaster* [8].

2.2. Data mining

SNPs in the human *PSEN1* gene were obtained from NCBI dbSNP database in August 2020. Those missense SNPs were extracted for further analysis. The sequence and accession number of the protein encoded by *PSEN1* gene were retrieved from NCBI dbSNP and Uniprot databases. The Single Nucleotide Polymorphism database is a variation database that involves entries submitted by private organizations and public laboratories for numerous organisms [9]. The Uniprot database contains protein sequences and related annotation in detail. Uniprot enables researchers to analyze large amounts of sequence and functional knowledge for proteins [10].

2.3. Identification of deleterious/damaging SNPs

Online publicly accessible web tools were used in order to identify deleterious/damaging SNPs. In this regard, the procedure was followed step by step as shown in Fig. 1.

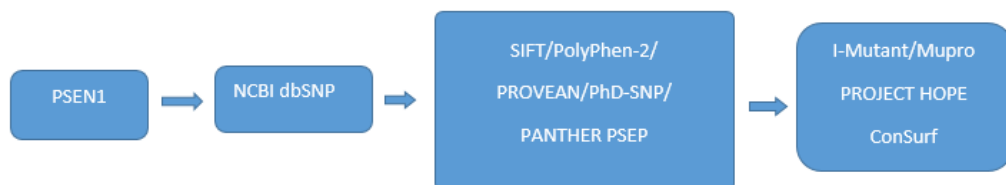


Figure 1: Online web tools used for SNP analysis (adapted from [11])

The Sorting Intolerant from Tolerant (SIFT) web server estimates the effects of substitutions of amino acids on proteins according to sequence homology and physical properties of amino acids. This algorithm enables users to characterize missense mutations [12]. The cutoff value in the SIFT software is a tolerance index of ≥ 0.05 [13]. An amino acid substitution with a value of < 0.05 is estimated as deleterious/intolerant to the human body, whereas a value of > 0.05 is estimated as tolerable [14]. PolyPhen-2 (Polymorphism Phenotyping v2) is an online web server that estimates the potential effects of amino acid substitutions on the function and stability of human proteins according to comparative and structural evolutionary assessment. The predictive findings are obtained as deleterious or intolerant [15]. This online web tool performs a score calculation based on the identification of a protein whose 3D-structure is known and its substitution site. PolyPhen score (PSIC) is calculated for each individual variant of sites and the differences were determined. It is known that there is a direct correlation between the score differences of variants and functional impacts of a specific amino acid substitution [13]. Protein Variation Effect Analyzer (PROVEAN) web server is used for the estimation of the functional effects of deletions, insertions, and amino acid substitutions. The cutoff score was adjusted to -2.5 in PROVEAN web server for high accuracy [16]. Predictor of human Deleterious Single Nucleotide Polymorphisms (PhD-SNP) predicts the effect of a mutation as a neutral or disease-related (pathogenic) with a reliability index score [17]. Freely available web tool PANTHER-PSEP estimates missense variations which may be implicated in the pathogenesis of human diseases based on the phylogenetically conservation scores [18].

2.4. Determination of the effects of deleterious/damaging SNPs on the stability of PSEN1 protein

In the present study, I-Mutant 3.0 and MUpro programs were used in order to determine the impacts of amino acid substitution on protein stability. I-Mutant 3.0 estimations are carried out initiating either from the protein sequence or protein structure. I-Mutant 3.0 software presents the association between protein stability and free energy change value (DDG) between wild type and mutant amino acids. DDG value < 0 (kcal/mol) indicates a decrease in stability, whereas DDG value > 0 (kcal/mol) presents an increase in stability. Consequently, it provides a prediction of protein stability as decreased or increased [19]. MUpro predicts the effects of mutations on protein stability based on the tertiary structure of the protein and the protein sequence [20].

2.5. Modeling of deleterious/damaging SNPs by Project HOPE

Project HOPE is a user friendly web application that analyzes the structural impacts of a mutation of interest. Project HOPE obtains structural information from several sources such as

sequence annotations in UniProt, calculations on three-dimensional protein structure, and estimations from the RefProt database. HOPE integrates this information to analyze the effects of a specific mutation on the protein structure. Moreover, HOPE enables users to predict the effects of wild type and mutant amino acids on protein 3D-structure via modeling [21].

2.6. Prediction of evolutionary conservation

ConSurf bioinformatics tool was used for the analysis of the phylogenetic conservation of amino acid residues of PSEN1 protein with the conservation scores in the range of 1 and 9 (1: rapidly evolving regions; 9: conserved positions) [22].

3. Results and Discussion

The gene-gene interaction network of *PSEN1* gene is shown in Figure 2. The findings of GeneMANIA tool show that the human *PSEN1* gene has 165 interactions in total with 21 genes.

A total of 22,537 variations in the human *PSEN1* gene, 386 of which were missense was found in NCBI dbSNP database. In the current study, 65 missense SNPs were determined as pathogenic by at least 4 of SIFT, PolyPhen-2, PROVEAN, PhD-SNP, and PANTHER PSEP software (Table 1).

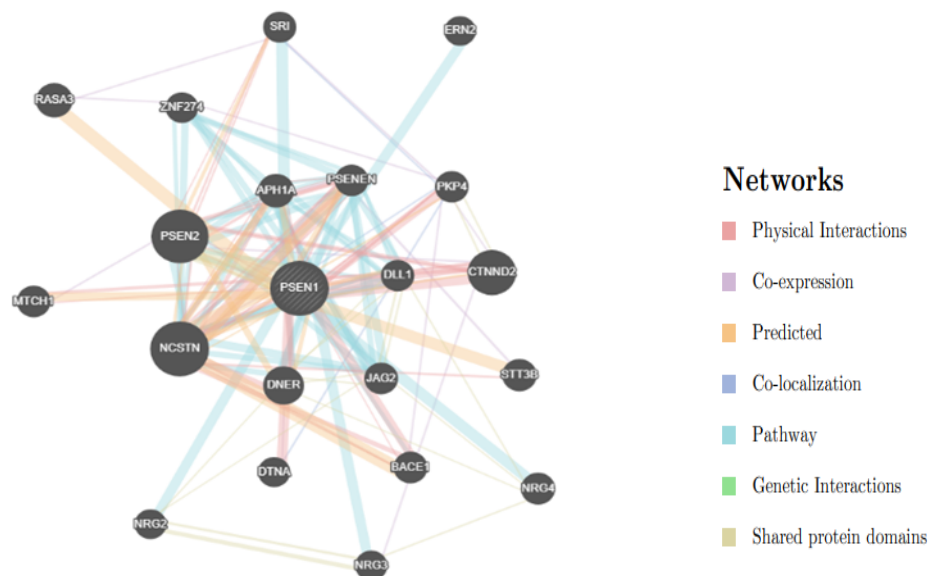


Figure 2: Gene-gene interactions of PSEN1

Protein stability affects three-dimensional structure and function of protein. Alterations in protein stability cause to protein misfolding, degradation, and aggregation [23]. The effects of 65 high-risk missense SNPs of *PSEN1* on protein stability was analyzed by I-Mutant 3.0 and MUpro software. The missense SNPs with DDG value <0 were estimated to destabilize PSEN1 protein. Moreover, the missense SNPs with DDG value <-1 were predicted to decrease significantly protein stability. These findings demonstrate that 7 of these SNPs lead to increase of protein stability (shown by at least one of two algorithmic programs), whereas 58 of these SNPs result in decrease of protein stability. 31 variants of *PSEN1* with $\Delta\Delta G$ values less than -1 kcal/mol (C410Y, L113P, F386S, L226R, Y115D, I229F, L282R, L258F, L166R, T291P, R278S, A246E, P117S, Y154N, L85P, V272A, I213F, L271V, V261F, M233T, L174R, L420R, C92S, L174M, L250S, L286V, I213T, L153V, W165C, R108W, F205L) significantly decreased the stability of PSEN1 protein (see Table 2).

Each amino acid has specific size, charge, and hydrophobicity values. Generally, wild-type and mutant residues have some differences in the regard of these properties. The findings of modeling of deleterious SNPs in *PSEN1* gene are seen in Table 3.

Phylogenetic conservation is vital for the determination of the negative consequences of mutations [23]. The amino acids located in conserved regions were highly damaging compared to the amino acids positioned in other sites. According to the ConSurf web server, R278T, P264L, N135D, R278S, E273A, P267S, G266S, P218L, and R157S were predicted as highly conserved and these residues were determined as functional residues. The findings of ConSurf are shown in Figure 3.

Table 1: Missense SNPs of *PSENI* gene estimated as pathogenic by several bioinformatics tools

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cutoff=-2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PANTHER PSEP	PANTHER PSEP Pdel
rs661	G/A	C410Y	DEL	0	Pro-damg	1.000	DEL	-9.987	Disease	7	Pro-damg	0.89
rs63749805	C/T	L113P	DEL	0.003	Pro-damg	1.000	DEL	-5.966	Disease	5	Pro-damg	0.85
rs63749824	C/T	A79V	DEL	0.018	Pro-damg	1.000	DEL	-3.627	Neutral	0	Pro-damg	0.85
rs63749836	G/A	A231T	DEL	0.046	Pro-damg	1.000	DEL	-3.647	Neutral	3	Pro-damg	0.85
rs63749860	T/C	F386S	DEL	0	Pro-damg	1.000	DEL	-7.334	Disease	5	Pro-damg	0.89
rs63749880	G/A	G209R	DEL	0	Pro-damg	1.000	DEL	-7.560	Disease	7	Pro-damg	0.89
rs63749891	G/C	R278T	DEL	0.001	Pro-damg	1.000	DEL	-5.462	Disease	2	Pro-damg	0.89
rs63749961	T/G	L226R	DEL	0	Pro-damg	1.000	DEL	-5.670	Disease	4	Pro-damg	0.86
rs63749962	T/G	Y115D	DEL	0.003	Pro-damg	1.000	DEL	-9.319	Disease	3	Pro-damg	0.86
rs63749967	G/C	V82L	DEL	0.004	Pro-damg	0.998	DEL	-2.722	Neutral	3	Pro-damg	0.85
rs63749970	A/T	I229F	DEL	0.003	Pro-damg	1.000	DEL	-3.680	Disease	4	Pro-damg	0.74
rs63749987	C/T	L219F	DEL	0.003	Pro-damg	1.000	DEL	-3.780	Disease	0	Pro-damg	0.85
rs63750050	T/G	L282R	DEL	0.001	Pro-damg	0.998	DEL	-4.560	Disease	5	Benign	0.5
rs63750053	G/T	G209V	DEL	0	Pro-damg	1.000	DEL	-8.505	Disease	8	Pro-damg	0.89
rs63750248	G/C	L258F	DEL	0.002	Pro-damg	1.000	DEL	-3.663	Neutral	3	Pro-damg	0.85
rs63750265	T/G	L166R	DEL	0	Pro-damg	0.999	DEL	-5.667	Disease	6	Pro-damg	0.85
rs63750298	A/C	T291P	DEL	0.041	Pro-damg	0.999	DEL	-3.649	Disease	5	Pro-damg	0.54
rs63750301	C/T	P264L	DEL	0.004	Pro-damg	1.000	DEL	-9.075	Disease	6	Pro-damg	0.89

Table 1: (Continued)

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cutoff=-2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PANTHER PSEP	PANTHER PSEP Pdel
rs63750306	A/C	M146L	DEL	0.015	Pro-damg	0.942	DEL	-2.838	Disease	6	Pro-damg	0.85
rs63750324	C/T	P284S	DEL	0.001	Pro-damg	1.000	DEL	-7.276	Disease	0	Pro-damg	0.89
rs63750353	A/G	N135D	DEL	0.023	Pro-damg	0.999	DEL	-4.685	Disease	7	Pro-damg	0.89
rs63750418	T/C	S169P	DEL	0.001	Pro-damg	0.997	DEL	-4.746	Disease	7	Pro-damg	0.85
rs63750444	G/A	G217D	DEL	0.001	Pro-damg	1.000	DEL	-6.448	Disease	7	Pro-damg	0.85
rs63750487	C/T	L226F	DEL	0.003	Pro-damg	1.000	DEL	-3.780	Neutral	3	Pro-damg	0.86
rs63750524	A/C	R278S	DEL	0.001	Pro-damg	1.000	DEL	-5.445	Disease	2	Pro-damg	0.89
rs63750526	C/A	A246E	DEL	0.007	Pro-damg	0.995	DEL	-3.111	Disease	5	Benign	0.5
rs63750550	C/T	P117S	DEL	0.007	Pro-damg	1.000	DEL	-7.248	Neutral	2	Pro-damg	0.85
rs63750577	C/T	S170F	DEL	0.005	Pro-damg	0.999	DEL	-5.564	Disease	6	Pro-damg	0.85
rs63750588	T/A	Y154N	DEL	0.009	Pro-damg	1.000	DEL	-8.456	Disease	5	Pro-damg	0.89
rs63750599	T/C	L85P	DEL	0.001	Pro-damg	1.000	DEL	-6.391	Disease	7	Pro-damg	0.85
rs63750601	G/T	V96F	DEL	0.001	Pro-damg	1.000	DEL	-4.643	Disease	5	Pro-damg	0.85
rs63750634	T/G	L250V	DEL	0.002	Pro-damg	1.000	DEL	-2.860	Neutral	1	Pro-damg	0.89
rs63750680	T/C	V272A	DEL	0.001	Pro-damg	0.999	DEL	-3.563	Neutral	5	Pro-damg	0.85
rs63750772	A/C	E273A	DEL	0.003	Pro-damg	0.999	DEL	-5.462	Neutral	3	Pro-damg	0.86
rs63750815	G/T	V89L	DEL	0.045	Pro-damg	0.998	DEL	-2.739	Neutral	5	Pro-damg	0.85
rs63750852	G/A	M93V	DEL	0.002	Pro-damg	0.988	DEL	-3.719	Neutral	1	Pro-damg	0.89

Table 1: (Continued)

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cutoff=-2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PANTHER PSEP	PANTHER PSEP Pdel
rs63750861	A/T	I213F	DEL	0.044	Pro-damg	1.000	DEL	-3.780	Disease	5	Pro-damg	0.85
rs63750863	C/T	P284L	DEL	0.002	Pro-damg	1.000	DEL	-9.158	Disease	5	Pro-damg	0.89
rs63750886	C/G	L271V	DEL	0.003	Pro-damg	0.999	DEL	-2.722	Neutral	6	Pro-damg	0.89
rs63750907	C/T	T147I	DEL	0.008	Pro-damg	1.000	DEL	-5.612	Disease	6	Pro-damg	0.89
rs63750964	G/T	V261F	DEL	0	Pro-damg	1.000	DEL	-4.537	Disease	4	Pro-damg	0.85
rs63751024	T/C	M233T	DEL	0.003	Pro-damg	0.998	DEL	-5.503	Disease	4	Pro-damg	0.85
rs63751025	T/G	L174R	DEL	0.001	Pro-damg	0.999	DEL	-5.552	Disease	6	Pro-damg	0.89
rs63751032	T/G	L420R	DEL	0.001	Pro-damg	0.997	DEL	-5.103	Disease	6	Pro-damg	0.85
rs63751071	T/G	M139I	DEL	0.001	Pro-damg	0.988	NEUTRAL	-1.710	Disease	6	Pro-damg	0.78
rs63751102	G/T	C263F	DEL	0.012	Pro-damg	0.987	DEL	-6.852	Disease	6	Pro-damg	0.89
rs63751141	G/C	C92S	DEL	0.008	Pro-damg	1.000	DEL	-9.131	Disease	2	Pro-damg	0.86
rs63751144	C/A	L174M	DEL	0.006	Pro-damg	0.999	NEUTRAL	-1.762	Disease	0	Pro-damg	0.89
rs63751163	T/C	L250S	DEL	0	Pro-damg	1.000	DEL	-5.720	Disease	4	Pro-damg	0.89
rs63751210	C/T	S169L	DEL	0.001	Pro-damg	0.977	DEL	-5.679	Disease	6	Pro-damg	0.85
rs63751229	C/T	P267S	DEL	0.003	Pro-damg	1.000	DEL	-7.260	Neutral	3	Pro-damg	0.89
rs63751235	C/G	L286V	DEL	0.02	Pro-damg	0.999	DEL	-2.706	Neutral	4	Pro-damg	0.89
rs63751292	A/G	Y154C	DEL	0.001	Pro-damg	1.000	DEL	-8.459	Disease	2	Pro-damg	0.89
rs63751309	T/C	I213T	DEL	0.002	Pro-damg	1.000	DEL	-4.723	Disease	3	Pro-damg	0.85
rs63751420	C/T	A260V	DEL	0	Pro-damg	1.000	DEL	-3.663	Neutral	2	Pro-damg	0.89

Table 1: (Continued)

SNP ID	Nucleotide substitution	Amino acid change	SIFT result	SIFT score	PolyPhen-2 result	PolyPhen-2 score	PROVEAN result (cutoff= -2.5)	PROVEAN score	PhD-SNP result	PhD-SNP RI	PANTHER PSEP	PANTHER Pdel
rs63751441	C/G	L153V	DEL	0.013	Pro-damg	0.999	DEL	-2.803	Neutral	2	Pro-damg	0.89
rs63751484	G/C	W165C	DEL	0	Pro-damg	1.000	DEL	-12.335	Disease	5	Pro-damg	0.89
rs121917807	G/A	G266S	DEL	0	Pro-damg	1.000	DEL	-5.445	Disease	0	Pro-damg	0.89
rs200576075	C/T	R108W	DEL	0.009	Pro-damg	1.000	DEL	-2.958	Disease	5	Pro-damg	0.57
rs267606983	G/C	G217R	DEL	0.003	Pro-damg	1.000	DEL	-7.326	Disease	5	Pro-damg	0.85
rs1042864	T/G	F205L	DEL	0.004	Pro-damg	1.000	DEL	-5.595	Neutral	1	Pro-damg	0.89
rs140064975	C/T	P218L	DEL	0.003	Pro-damg	1.000	DEL	-9.450	Disease	7	Pro-damg	0.89
rs146855665	A/G	M93V	DEL	0.003	Pro-damg	0.988	DEL	-3.719	Neutral	1	Pro-damg	0.89
rs200065583	A/G	Y195C	DEL	0.042	Pro-damg	0.999	DEL	-8.064	Disease	5	Pro-damg	0.85
rs201617677	G/T	R157S	DEL	0.007	Pro-damg	0.992	DEL	-5.445	Disease	5	Pro-damg	0.85

DEL: Deleterious; Pro-damg: Probably Damaging

Table 2: Effects of missense SNPs on PSEN1 protein stability by MUpro and I-MUTANT 3.0

SNP ID	Amino acid change	MUpro Result	MUpro DDG	I-MUTANT Result	I-MUTANT RI
rs661	C410Y	Decrease	-1.5702591	Decrease	4
rs63749805	L113P	Decrease	-1.2471141	Decrease	5
rs63749824	A79V	Increase	0.30877608	Increase	3
rs63749836	A231T	Decrease	-0.7915774	Decrease	8
rs63749860	F386S	Decrease	-2.0290316	Decrease	8
rs63749880	G209R	Decrease	-0.4073739	Decrease	6
rs63749891	R278T	Decrease	-0.82107432	Decrease	8
rs63749961	L226R	Decrease	-1.4840623	Decrease	6
rs63749962	Y115D	Decrease	-1.130926	Decrease	0
rs63749967	V82L	Decrease	-0.035332349	Decrease	8
rs63749970	I229F	Decrease	-1.0943433	Decrease	8
rs63749987	L219F	Decrease	-0.60164771	Decrease	7
rs63750050	L282R	Decrease	-1.8779089	Decrease	9
rs63750053	G209V	Decrease	-0.24172348	Decrease	4
rs63750248	L258F	Decrease	-1.1461876	Decrease	8
rs63750265	L166R	Decrease	-1.5897411	Decrease	7
rs63750298	T291P	Decrease	-1.6653398	Decrease	4
rs63750301	P264L	Decrease	-0.14082925	Decrease	8
rs63750306	M146L	Decrease	-0.35915167	Decrease	6
rs63750324	P284S	Decrease	-0.85404669	Decrease	9
rs63750353	N135D	Decrease	-0.48001694	Decrease	1
rs63750418	S169P	Increase	0.27204534	Increase	3
rs63750444	G217D	Decrease	-0.51867258	Decrease	4
rs63750487	L226F	Decrease	-0.9439387	Decrease	7
rs63750524	R278S	Decrease	-1.0422465	Decrease	9
rs63750526	A246E	Decrease	-1.1272521	Decrease	4
rs63750550	P117S	Decrease	-1.5170786	Decrease	8
rs63750577	S170F	Decrease	-0.76852242	Decrease	1
rs63750588	Y154N	Decrease	-1.1430991	Decrease	7
rs63750599	L85P	Decrease	-2.1723561	Decrease	6
rs63750601	V96F	Decrease	-0.43148039	Decrease	8
rs63750634	L250V	Decrease	-0.92485763	Decrease	8
rs63750680	V272A	Decrease	-1.5337295	Decrease	8
rs63750772	E273A	Decrease	-0.65624872	Decrease	7
rs63750815	V89L	Decrease	-0.16833404	Decrease	7
rs63750852	M93V	Decrease	-0.66951747	Decrease	5

Table 2: (Continued)

SNP ID	Amino acid change	MUpro Result	MUpro DDG	I-MUTANT Result	I-MUTANT RI
rs63750861	I213F	Decrease	-1.0867232	Decrease	8
rs63750863	P284L	Increase	0.060398754	Decrease	1
rs63750886	L271V	Decrease	-1.2309305	Increase	8
rs63750907	T147I	Increase	0.18845675	Decrease	4
rs63750964	V261F	Decrease	-1.0883674	Decrease	9
rs63751024	M233T	Decrease	-1.6717531	Decrease	7
rs63751025	L174R	Decrease	-2.1523904	Decrease	5
rs63751032	L420R	Decrease	-1.3814382	Decrease	6
rs63751071	M139I	Decrease	-0.32100029	Decrease	6
rs63751102	C263F	Decrease	-0.56900868	Decrease	3
rs63751141	C92S	Decrease	-2.0291827	Decrease	4
rs63751144	L174M	Decrease	-1.4928968	Decrease	2
rs63751163	L250S	Decrease	-1.6753384	Decrease	9
rs63751210	S169L	Decrease	-0.26655644	Decrease	0
rs63751229	P267S	Decrease	-0.77273729	Decrease	8
rs63751235	L286V	Decrease	-1.3590371	Decrease	8
rs63751292	Y154C	Decrease	-0.54194646	Decrease	4
rs63751309	I213T	Decrease	-2.0230433	Decrease	8
rs63751420	A260V	Increase	0.18155527	Increase	3
rs63751441	L153V	Decrease	-1.4723186	Decrease	7
rs63751484	W165C	Decrease	-1.3047304	Decrease	7
rs121917807	G266S	Decrease	-0.58466487	Decrease	8
rs200576075	R108W	Decrease	-1.4052264	Decrease	4
rs267606983	G217R	Decrease	-0.46268638	Decrease	5
rs1042864	F205L	Decrease	-1.315548	Decrease	5
rs140064975	P218L	Increase	0.3308844	Decrease	2
rs146855665	M93V	Decrease	-0.66951747	Decrease	5
rs200065583	Y195C	Decrease	-0.92612274	Decrease	2
rs201617677	R157S	Decrease	-0.64541036	Decrease	8

Table 3: Modeling of deleterious SNPs in *PSEN1* gene by Project HOPE

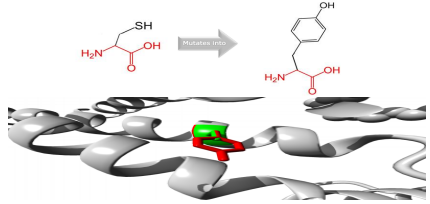
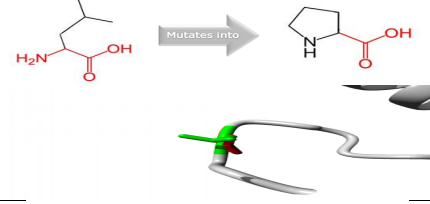
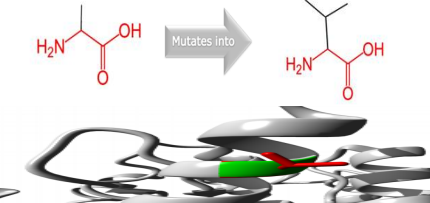
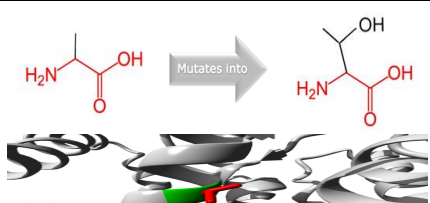
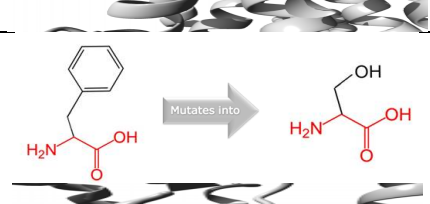
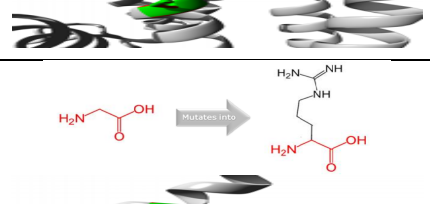
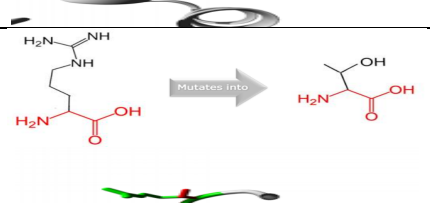
SNP ID	Modeling	Description
rs661		Conversion of cysteine into tyrosine at position 410 due to rs661 polymorphism. The mutant residue is less hydrophobic than wild-type residue
rs63749805		Conversion of leucine into proline at position 113 due to rs63749805 polymorphism. This mutation is probably damaging to the protein
rs63749824		Conversion of alanine into valine at position 79 due to rs63749824 polymorphism and it is probably damaging to the protein
rs63749836		Conversion of alanine into threonine at position 231 due to rs63749836 and it is located in transmembrane domain and affects the interactions with membrane lipids
rs63749860		Conversion of phenylalanine into serine at position 386 due to rs63749860 polymorphism and it is located close to active site and affects protein function and interactions with membrane lipids
rs63749880		Conversion of glycine into arginine at position 209 due to rs63749880 polymorphism and it is probably damaging to the protein
rs63749891		Conversion of arginine into threonine at position 278 due to rs63749891 polymorphism. This mutation is probably damaging to the protein

Table 3: (Continued)

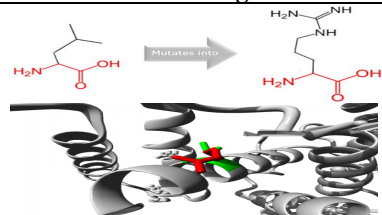
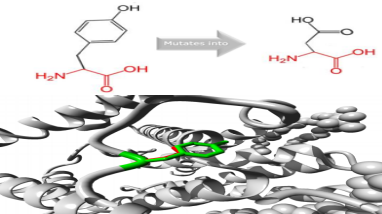
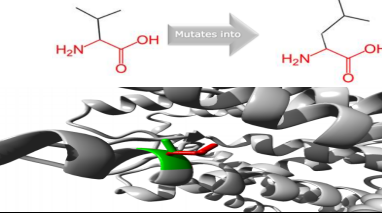
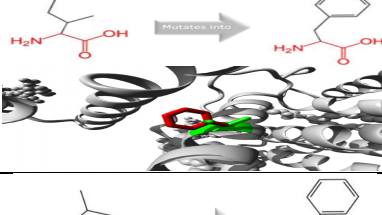
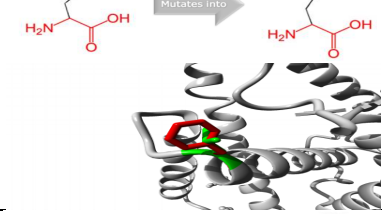
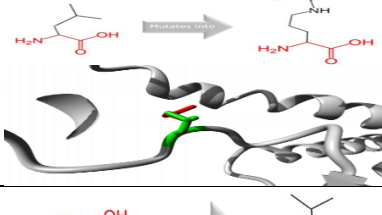
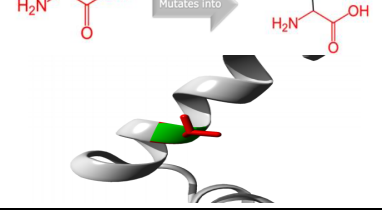
SNP ID	Modeling	Description
rs63749961		Conversion of leucine into arginine at position 226 due to rs63749961 polymorphism. This mutation is probably damaging to the protein
rs63749962		Conversion of tyrosine into aspartic acid at position 115 due to rs63749962 polymorphism. The mutant residue may cause to loss of interactions
rs63749967		Conversion of valine into leucine at position 82 due to rs63749967 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type
rs63749970		Conversion of isoleucine into phenylalanine at position 229 due to rs63749970 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type
rs63749987		Conversion of leucine into phenylalanine at position 219 due to rs63749987 polymorphism. This mutation is probably damaging to the protein
rs63750050		Conversion of leucine into arginine at position 282 due to rs63750050 polymorphism. The mutation is possibly damaging
rs63750053		Conversion of glycine into valine at position 209 due to rs63750053 polymorphism. The mutant residue is bigger and more hydrophobic than the wild-type residue

Table 3: (Continued)

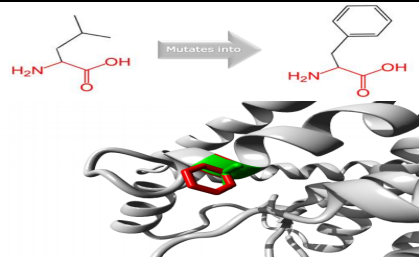
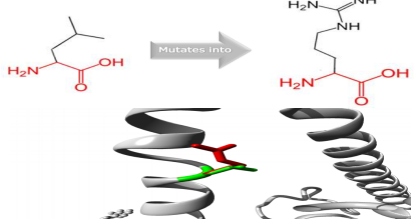
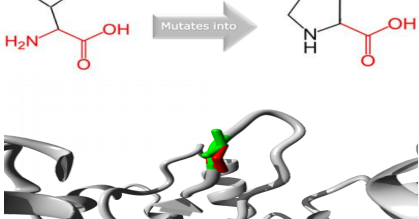
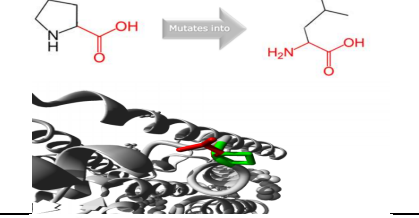
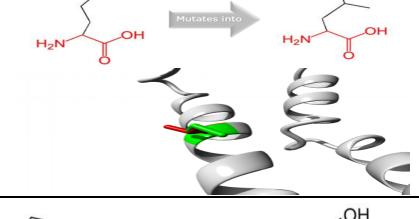
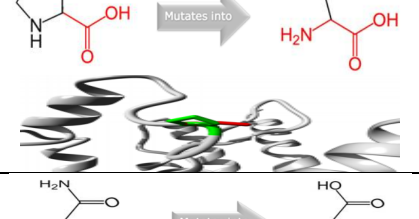
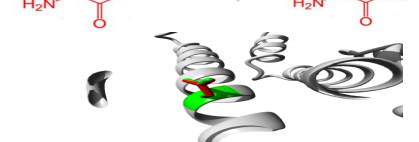
SNP ID	Modeling	Description
rs63750248		Conversion of leucine into phenylalanine at position 258 due to rs63750248 polymorphism. The mutant residue is bigger than wild-type residue and affects the interactions with lipid membrane
rs63750265		Conversion of leucine into arginine at position 166 due to rs63750265 polymorphism. The mutant residue is less hydrophobic than the wild-type residue
rs63750298		Conversion of threonine into proline at position 291 due to rs63750298 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding
rs63750301		Conversion of proline into leucine at position 264 due to rs63750301 polymorphism. This mutation might disturb the conformation
rs63750306		Conversion of methionine into leucine at position 146 due to rs63750306 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs63750324		Conversion of proline into serine at position 284 due to rs63750324 polymorphism and it is probably damaging to the protein
rs63750353		Conversion of asparagine into aspartic acid at position 135 due to rs63750353 polymorphism. The charge of mutant residue is different

Table 3: (Continued)

SNP ID	Modeling	Description
rs63750418		Conversion of serine into proline at position 169 due to rs63750418 polymorphism. The mutant residue is more hydrophobic and affects the interactions with membrane lipids
rs63750444		Conversion of glycine into aspartic acid at position 217 due to rs63750444 polymorphism. This mutation may lead to incorrect conformation and disturb the local structure
rs63750487		Conversion of leucine into phenylalanine at position 226 due to rs63750487 polymorphism and it is probably damaging to the protein
rs63750524		Conversion of arginine into serine at position 278 due to rs63750524 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding
rs63750526		Conversion of alanine into glutamic acid at position 246 due to rs63750526 polymorphism. This mutation leads to loss of hydrophobic interactions
rs63750550		Conversion of proline into serine at position 117 due to rs63750550 polymorphism. This mutation may lead to incorrect conformation and disturb the local structure
rs63750577		Conversion of serine into phenylalanine at position 170 due to rs63750577 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding

Table 3: (Continued)

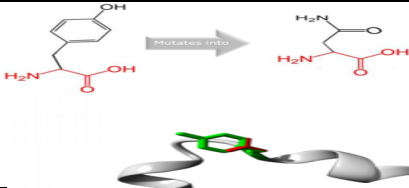
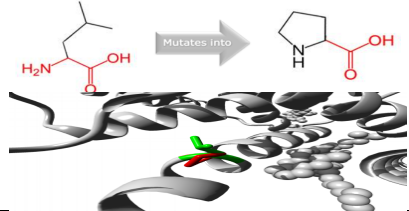
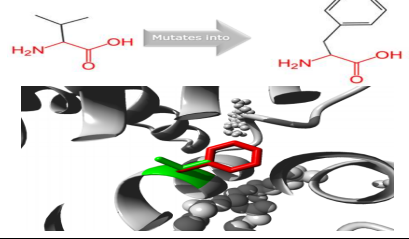
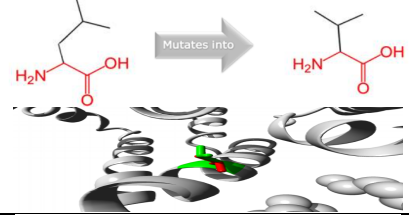
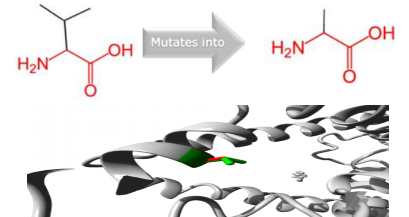
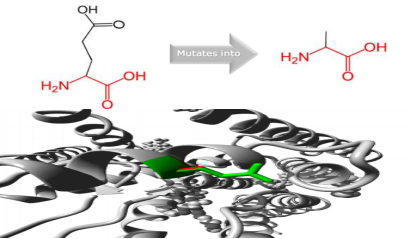
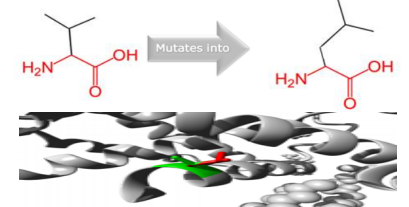
SNP ID	Modeling	Description
rs63750588		Conversion of tyrosine into asparagine at position 154 due to rs63750588 polymorphism. This mutation leads to loss of hydrophobic interactions
rs63750599		Conversion of leucine into proline at position 85 due to rs63750599 polymorphism. This mutant residue is smaller and therefore it may lead to loss of interactions
rs63750601		Conversion of valine into phenylalanine at position 96 due to rs63750601 polymorphism. This mutant residue may cause to bumps due to being bigger than wild-type
rs63750634		Conversion of leucine into valine at position 250 due to rs63750634 polymorphism. This mutant residue is smaller and it may cause to loss of interactions
rs63750680		Conversion of valine into alanine at position 272 due to rs63750680 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs63750772		Conversion of glutamic acid into alanine at position 273 due to rs63750772 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding
rs63750815		Conversion of valine into leucine at position 89 due to rs63750815 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type

Table 3: (Continued)

SNP ID	Modeling	Description
rs63750852		<p>Conversion of methionine into valine at position 93 due to rs63750852 polymorphism. This mutant residue is smaller and therefore it may lead to loss of interactions</p>
rs63750861		<p>Conversion of isoleucine into phenylalanine at position 213 due to rs63750861 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type</p>
rs63750863		<p>Conversion of proline into leucine at position 284 due to rs63750863 polymorphism and it is probably damaging to the protein</p>
rs63750886		<p>Conversion of leucine into valine at position 271 due to rs63750886 polymorphism. The mutant residue is smaller and it may cause to loss of interactions</p>
rs63750907		<p>Conversion of threonine into isoleucine at position 147 due to rs63750907 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding</p>
rs63750964		<p>Conversion of valine into phenylalanine at position 261 due to rs63750964 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type</p>
rs63751024		<p>Conversion of methionine into threonine at position 233 due to rs63751024 polymorphism. The mutant residue is smaller and it may cause to loss of interactions</p>

Table 3: (Continued).

SNP ID	Modeling	Description
rs63751025		Conversion of leucine into arginine at position 174 due to rs63751025 polymorphism. This mutation leads to loss of hydrophobic interactions
rs63751032		Conversion of leucine into arginine at position 420 due to rs63751032 polymorphism. This mutation leads to loss of hydrophobic interactions
rs63751071		Conversion of methionine into isoleucine at position 139 due to rs63751071 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs63751102		Conversion of cysteine into phenylalanine at position 263 due to rs63749836 polymorphism and it is probably damaging to the protein
rs63751141		Conversion of cysteine into serine at position 92 due to rs63751141 polymorphism and it is probably damaging to the protein
rs63751144		Conversion of leucine into methionine at position 174 due to rs63751144 polymorphism. The mutant residue may cause to bumps due to being bigger than wild-type
rs63751163		Conversion of leucine into serine at position 250 due to rs63751163 polymorphism. This mutation is probably damaging to the protein

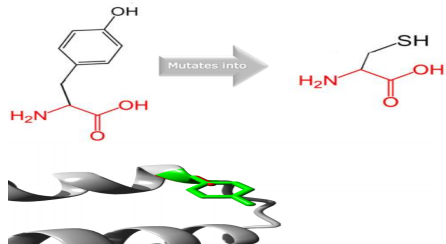
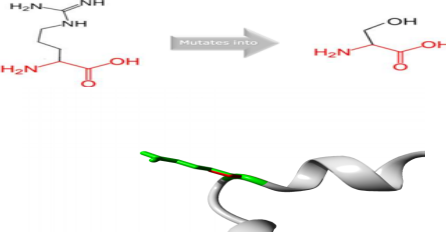
Table 3: (Continued)

SNP ID	Modeling	Description
rs63751210		Conversion of serine into leucine at position 169 due to rs63751210 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding
rs63751229		Conversion of proline into serine at position 267 due to rs63751229 polymorphism. This mutation may lead to incorrect conformation and disturb the local structure
rs63751235		Conversion of leucine into valine at position 286 due to rs63751235 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs63751292		Conversion of tyrosine into cysteine at position 154 due to rs63751292 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding
rs63751309		Conversion of isoleucine into threonine at position 213 due to rs63751309 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs63751420		Conversion of alanine into valine at position 260 due to rs63751420 polymorphism and it is probably damaging to the protein
rs63751441		Conversion of leucine into valine at position 153 due to rs63751441 polymorphism. This mutation is probably damaging to the protein

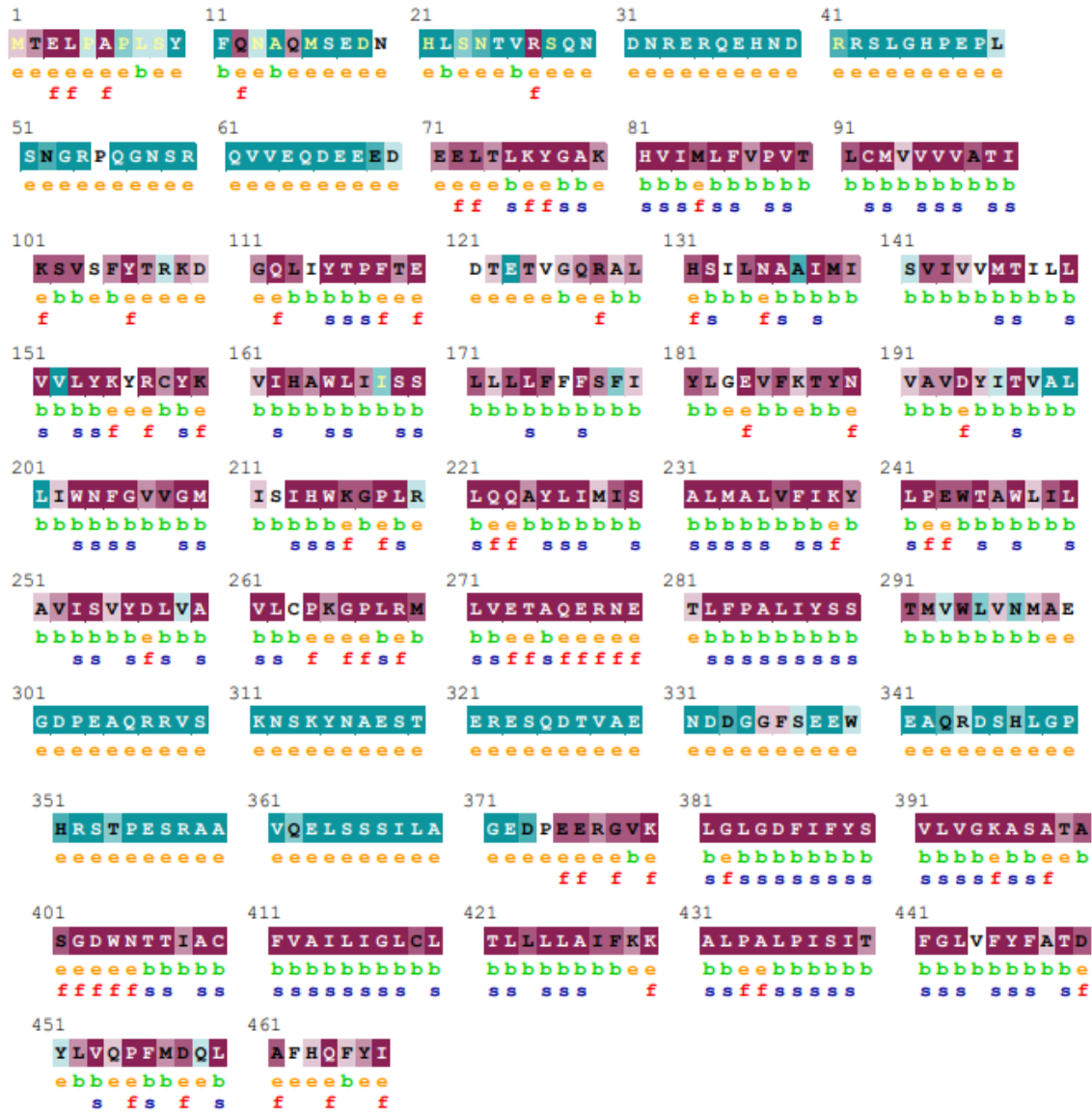
Table 3: (Continued)

SNP ID	Modeling	Description
rs63751484		Conversion of tryptophan into cysteine at position 165 due to rs63751484 polymorphism. This mutation is probably damaging to the protein
rs121917807		Conversion of glycine into serine at position 266 due to rs121917807 polymorphism. This mutation may lead to incorrect conformation and disturb the local structure
rs200576075		Conversion of arginine into tryptophan at position 108 due to rs200576075 polymorphism and it is probably damaging to the protein
rs267606983		Conversion of glycine into arginine at position 217 due to rs267606983 and it is probably damaging to the protein
rs1042864		Conversion of phenylalanine into leucine at position 205 due to rs1042864 polymorphism. The mutant residue is smaller and it may cause to loss of interactions
rs140064975		Conversion of proline into leucine at position 218 due to rs140064975 polymorphism and it is probably damaging to the protein
rs146855665		Conversion of methionine into valine at position 93 due to rs146855665 polymorphism. This mutation is probably damaging to the protein

Table 3: (Continued)

SNP ID	Modeling	Description
rs200065583		<p>Conversion of tyrosine into cysteine at position 195 due to rs200065583 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding</p>
rs201617677		<p>Conversion of arginine into serine at position 157 due to rs201617677 polymorphism. The mutant residue is more hydrophobic and may lead to loss of hydrogen bonds and misfolding</p>
<p>■ wild-type residue ■ mutant residue</p>		

ConSurf Results



The conservation scale:



- e - An exposed residue according to the neural-network algorithm.
- b - A buried residue according to the neural-network algorithm.
- f - A predicted functional residue (highly conserved and exposed).
- s - A predicted structural residue (highly conserved and buried).
- * - Insufficient data - the calculation for this site was

Figure 3: Evolutionary conservancy of PSEN1 generated by Consurf web tool

Alzheimer's Disease is a progressive neurodegenerative disease and caused by abnormal accumulation of tau protein and amyloid beta ($A\beta$) in the central nervous system [24]. Human *PSENI* is one of the causative genes in the pathogenesis of AD. On the other hand, the underpinning mechanism of that how *PSENI* mutations cause to dementia and neurodegeneration is needed to be elucidated [25].

In the current study, the target SNPs in the human *PSENI* gene with an approach based on the computer-based software tools such as PolyPhen-2, SIFT, PROVEAN, PhD-SNP, PANTHER PSEP, I-Mutant 3.0, MUpro, Project HOPE, and ConSurf were determined before experimental studies. Determination of target SNPs that have pathogenic effects on the protein structure and stability by *in silico* methods is significant for genotyping studies. In this study, the SNPs that might have functional effects on the PSEN1 protein were investigated.

It has been determined that these 386 SNPs in the human *PSENI* gene are missense and 65 polymorphisms are deleterious/damaging. The *PSENI* rs63749824, rs63750418, rs63750863, rs63750886, rs63750907, rs63751420, and rs140064975 polymorphisms resulting in A79V, S169P, P284L, L271V, T147I, A260V, and P218L amino acid substitutions, respectively, increase protein stability, whereas other 58 polymorphisms lead to decrease of protein stability.

The genes that PSEN1 gene interacted most with were NCSTN (nicastrin), PSEN2 (presenilin 2), and CTNND2 (catenin delta 2), respectively, according to the findings of the GeneMANIA software tool. Moreover, GeneMANIA results indicated that NCSTN, PSEN2, and CTNND2 genes have several significant functional roles such as beta-amyloid metabolic process, amyloid precursor protein metabolic process, and positive regulation of apoptosis.

Determination of the effects of amino acid substitutions on protein structure and function is crucial in order to elucidate the complex mechanisms of human diseases caused by single nucleotide polymorphisms [26, 27, 28]. Project HOPE software findings have provided significant information about the probable effects of missense SNPs in the human *PSENI* gene. Findings of the current study have reported that amino acid substitutions in the *PSENI* gene affect the interactions with membrane lipids, charge, and hydrophobicity, and may cause to loss of interactions, bumps due to being bigger than wild-type, loss of hydrogen bonds and misfolding, incorrect conformation and disturb the local structure. It has been demonstrated that rs63749805, rs63749824, rs63749880, rs63749891, rs63749961, rs63749987, rs63750050, rs63750324, rs63750487, rs63750863, rs63751102, rs63751141, rs63751163, rs63751420, rs63751441, rs63751484, rs200576075, rs267606983, rs140064975, rs146855665 polymorphisms in the *PSENI* gene may lead to probably damaging to PSEN1 protein.

In a study conducted with early-onset AD cases, three rare missense variants (G417A, G209A, and T119I) were found to be significant in the pathogenesis of AD [29]. It has been reported that PSEN1 W165C had pathogenic effect on early-onset AD [30]. In our study, PSEN1 W165C was determined as damaging. It has been demonstrated that PSEN1 P264L was pathogenic in Turkish dementia patients [31]. Similarly, PSEN1 P264L was found to be high-risk pathogenic missense SNP in the current study. Veugelen et al. reported that PSEN1 C410Y variants may be implicated in the generation of A β [32].

According to the data obtained from PolyPhen-2, SIFT, PROVEAN, PhD-SNP, PANTHER PSEP bioinformatics programs, 65 missense SNPs were determined as damaging. 8 missense SNPs were evaluated as high-risk pathogenic due to their effects on protein stability, being in highly conserved positions of the protein sequence, and causing changes in some properties of the protein, such as charge and hydrophobicity. Based on the findings of the current study in general, it was demonstrated 8 significant missense SNPs -rs63749891 (R278T), rs63750301 (P264L), rs63750353 (N135D), rs63750524 (R278S), rs63750772 (E273A), rs63751229 (P267S), rs121917807 (G266S), and rs201617677 (R157S)- were determined as high-risk pathogenic since: a) the 8 missense SNPs were predicted to be damaging by at least 4 bioinformatics tools; b) these missense SNPs may decrease the stability of PSEN1 protein; c) it was determined that these 8 missense SNPs were in highly conserved positions in the protein sequence; d) 3D-modeling showed that these SNPs were damaging to PSEN1 protein and may cause to some changes such as charge, and hydrophobicity.

Our study demonstrates that high-risk pathogenic missense SNPs may have the potential to alter catalytic activity of γ secretase complex and subsequently the amount of A β 40 ve A β 42. Therefore, these missense SNPs may contribute to AD pathogenesis studies.

4. Conclusion

The current study investigated the effects of functional SNPs associated with the human *PSEN1* gene via computational methods due to the relationship between PSEN1 and Alzheimer's Disease. In a total of 22,537 SNPs in PSEN1 gene, 386 SNPs were found to be missense. Moreover, 8 significant missense SNPs were determined as high-risk pathogenic by all of the bioinformatic tools used in this study. We suppose that the results of the present study comprise a basis for future experimental and *in silico* studies. The present study suggests that function and/or structure of PSEN1 protein might be disturbed by various missense SNPs. These missense SNPs may be mightily considered as main targets in causing diseases such as AD associated with PSEN1 malfunction and therefore will be helpful for the development of novel and effective

drugs. It is known that *in silico* analysis of missense SNPs in human genes that are associated with AD is notable for future population and candidate gene studies.

References

- [1] Reitz, C., *Genetic diagnosis and prognosis of Alzheimer's disease: challenges and opportunities*, Expert Review of Molecular Diagnostics, 15 (3), 339-348, 2015.
- [2] Liu, Z., Li, T., Li, P., et al., *The ambiguous relationship of oxidative stress, tau hyperphosphorylation, and autophagy dysfunction in Alzheimer's disease*, Oxidative Medicine and Cellular Longevity, 2015, 352723, 2015.
- [3] Goldman, J.S., Deerlin, V.M., *Alzheimer's disease and frontotemporal dementia: the current state of genetics and genetic testing since the advent of next generation sequencing*, Molecular Diagnosis & Therapy, 22 (5), 505-513, 2018.
- [4] Bagyinszky, E., Lee, H.M., Giau, V.V., et al., *PSEN1 p.Thr116Ile variant in two Korean families with young onset Alzheimer's disease*, International Journal of Molecular Sciences, 19 (9), 2604, 2018.
- [5] Sproul, A.A., Jacob, S., Pre, D., et al., *Characterization and molecular profiling of PSEN1 familial Alzheimer's disease iPSC-derived neural progenitors*, PLoS One, 9 (1), e84547, 2014.
- [6] Ramirez-Bello, J., Jimenez-Morales, M., *Functional implications of single nucleotide polymorphisms (SNPs) in protein-coding and non-coding RNA genes in multifactorial diseases*, Gaceta medica de Mexico, 153 (2), 238-250, 2017.
- [7] Jamal, S., Goyal, S., Shanker, A., Grover, A., *Computational screening and exploration of disease-associated genes in Alzheimer's disease*, Journal of Cellular Biochemistry, 118(6), 1471-1479, 2016.
- [8] Montojo, J., Zuberi, K., Rodriguez, H., Bader, G.D., Morris, Q., *GeneMANIA: Fast gene network construction and function prediction for Cytoscape [v1]*; ref status: indexed, <http://f1000r.es/3rv>, F1000Research, 3 (153), 2014.
- [9] Bhagwat, M., *Searching NCBI's dbSNP database*, Current Protocols in Bioinformatics, Chapter1, Unit 1-19, 2010.
- [10] The Uniprot Consortium, *UniProt: the universal protein knowledgebase*, Nucleic Acids Research, 45, D158-D169, 2017.
- [11] Arshad, M., Bhatti, A., John, P., *Identification and in silico analysis of functional SNPs of human TAGAP protein: A comprehensive study*, PLoS ONE, 13 (1): e0188143.
- [12] Sim, N.L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., Ng, P.C., *SIFT web server: predicting effects of amino acid substitutions on proteins*, Nucleic Acids Research, 40 (W1), W452-W457, 2012.
- [13] Osman, M.M., Khalifa, A.S., Mutasim, A.E.Y., Massaad, S.O., Gasemelseed, M.M., Abdagader, M.A., Ahmed, S.A., Ahmed, A.M., Altayb, H.N., Salih, M.A., *In silico Analysis of Single Nucleotide Polymorphisms (Snps) in Human FTO Gene*, JSM Bioinformatics, Genomics and Proteomics, 1 (1), 1003, 2016.
- [14] Kaur, T., Khakur, T., Singh, J., Kamboj, S.S., Kaur, M., *Identification of functional SNPs in human LGALS3 gene by in silico analyses*, Egyptian Journal of Medical Human Genetics, 18 (4), 321-328, 2017.

- [15] Adzhubei, I., Jordan, D.M., Sunyaev, S.R., *Predicting functional effect of human UNIT 7.20 missense mutations using PolyPhen-2*, *Current Protocols in Human Genetics*, 76 (1), 7.20.1-7.20.41, 2013.
- [16] Choi, Y., Chan, A.P., *PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels*, *Bioinformatics*, 31(16), 2745–2747, 2015.
- [17] Capriotti, E., Fariselli, P., *PhD-SNPg: a webservice and lightweight tool for scoring single nucleotide variants*, *Nucleic Acids Research*, 45, W247–W252, 2017.
- [18] Tang, H., Thomas, P.D., *PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation*, *Bioinformatics*, 32 (14), 2230-2232, 2016.
- [19] Capriotti, E., Fariselli, P., Casadio, R., *I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure*, *Nucleic Acids Research*, 33, 306-310, 2005.
- [20] Cheng, J., Randall, A., Baldi, P., *Prediction of protein stability changes for single-site mutations using support vector machines*, *Proteins*, 62 (4), 1125-32, 2006.
- [21] Venselaar, H., Beek, T.A., Kuipers, R.K., Hekkelman, M.L., Vriend, G., *Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces*, *BMC Bioinformatics*, 11, 548, 2010.
- [22] Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., Ben-Ta, N., *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules*, *Nucleic Acids Research*, 44, W344-W350, 2016.
- [23] Hossain, S., Roy, A.S., Islam, S., *In silico analysis predicting effects of deleterious SNPs of human RASSF5 gene on its structure and functions*. *Scientific Reports*, 10, 14542, 2020.
- [24] Giau, V.V., Pyun, J.M., Suh, J., Bagyinszky, E., An, S.S.A., Kim, S.Y., *A pathogenic PSEN1 Trp165Cys mutation associated with early-onset Alzheimer's disease*, *BMC Neurology*, 19, 188, 2018.
- [25] Kelleher, R.J., Shen, J., *Presenilin-1 mutations and Alzheimer's disease*, *PNAS*, 144 (4), 629–631, 2017.
- [26] Cargill, M., Altshuler, D., Ireland, J., et al., *Characterization of single-nucleotide polymorphisms in coding regions of human genes*, *Nature Genetics*, 22 (3), 231-8, 1999.
- [27] Teng, S., Wang, L., Srivastava, A.K., Schwartz, C.E., Alexov, E., *Structural assessment of the effects of amino acid substitutions on protein stability and protein-protein interaction*, *International Journal of Computational Biology and Drug Design*, 3 (4), 334-349, 2010.
- [28] Ozkan Oktay, E., Kaman, T., Karasakal, O.F., Ulucan, K., Konuk, M., Tarhan, N., *Alzheimer hastalığı ile ilişkilendirilen APH1A genindeki zararlı SNP'lerin in silico yöntemler ile belirlenmesi*, *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 23 (2), 472-480, 2019.
- [29] Van Giau, V.V., Bagyinszky, E., Yang, Y., Youn, Y.C., Soo, S., Kim, S.Y., *Genetic analyses of early-onset Alzheimer's disease using next generation sequencing*, *Scientific Reports*, 9, 8368, 2019.
- [30] Van Giau, V., Pyun, J.M., Suh, J., Bagyinszky, E., Soo, S., Kim, S.Y., *A pathogenic PSEN1 Trp165Cys mutation associated with early-onset Alzheimer's disease*, *BMC Neurology*, 19, 188, 2019.
- [31] Lohmann, E., Guerreiro, R.J., Erginel-Unaltuna, N., Gurlunlian, N., Bilgic, B., Gurvit, H., Hanagasi, H.A., Luu, N., Emre, M., Singleton, A., *Identification of PSEN1 and PSEN2 gene*

mutations and variants in Turkish dementia patients, *Neurobiology of Aging*, 33 (8), 1850.e17-1850.e27, 2012.

[32] Veugelen, S., Saito, T., Saido, T.C., Chavez-Gutierrez, L., Stroope, B., *Familial Alzheimer's disease mutations in presenilin generate amyloidogenic Ab peptide seeds*, *Neuron*, 90 (2), 410-416, 2016.



Comparatively Investigation of Textile Dye Decolorization by a White Rot Fungus and Various Bacterial Strains

Emre BİRHANLI^{1,*}, Özfer YEŞİLADA¹, Ahmet ÇABUK², Filiz BORAN¹, Eray TATLICI¹

¹ Inonu University, Arts and Science Faculty, Department of Biology, Malatya, Turkey.

emre.birhanli@inonu.edu.tr, ORCID ID: 0000-0001-7965-4979; *ozfer.yesilada@inonu.edu.tr*, ORCID

ID: 0000-0003-0038-6575; *filiz.kuru@inonu.edu.tr*, ORCID ID: 0000-0002-8801-7987;

eraytatlici@gmail.com, ORCID ID: 0000-0003-2589-4234

² Eskisehir Osmangazi University, Arts and Science Faculty, Department of Biology, Eskisehir, Turkey

acabuk@ogu.edu.tr, ORCID ID: 0000-0002-4619-6948

Received: 19.07.2021

Accepted: 19.10.2021

Published: 31.12.2021

Abstract

The aim of this study is to comparatively investigate the decolorization of Reactive Blue 171 (RB 171) by using three different bacterial strains as *Bacillus megaterium* A1 (A1), *Gordonia* sp. MC-D1 (D1), and *Bacillus pumilus* D3 (D3) and also a white rot fungus (*Pleurotus ostreatus*). All tested bacteria were incubated with RB 171 dye at 150 mg/L concentration for 24-72 h under static or agitated (150 rpm) conditions while the fungus was incubated with 150 mg/L of the dye under the same conditions for 3-24 h. The highest bacterial decolorization values were obtained after 72 h of incubation under static conditions, and the maximum decolorization rates were detected as 84, 83 and 75% for A1, D1, and D3, respectively. However, the color of RB 171 dye was removed at 93% rate by *Pleurotus ostreatus* under static conditions after 24 h of incubation. Similar results were also obtained from the agitated studies of the tested fungal and bacterial strains except A1. The maximum



decolorization values obtained with A1, D1, and D3 at 150 rpm, 72 h were 30, 88 and 89%, respectively. The highest decolorization activities of *Pleurotus ostreatus* were 93% for both static and agitated (150 rpm) conditions. In addition, zymogram analyzes of the fungal culture fluids obtained from SBM, and SBM containing 150 mg/L RB 171 were also performed to detect the presence of laccase.

Keywords: Bacterium; Decolorization; Textile dye; White rot fungus; Zymogram.

Bir Beyaz Çürükçül Fungus ve Çeşitli Bakteriyel Suşlar ile Tekstil Boyasının Renginin Gideriminin Karşılaştırmalı Olarak Araştırılması

Öz

Bu çalışmanın amacı, üç farklı bakteriyel suş *Bacillus megaterium* A1 (A1), *Gordonia* sp. MC-D1 (D1) and *Bacillus pumilus* D3 (D3) ve ayrıca bir beyaz çürükçül fungusun (*Pleurotus ostreatus*) kullanımı ile Reaktif Mavi 171' in (RM 171) renginin gideriminin karşılaştırmalı olarak araştırılmasıdır. Test edilen tüm bakteriler 150 mg/L konsantrasyonda RM 171 boyası ile 24-72 saat boyunca statik ve çalkalamalı (150 rpm) koşullarda inkübe edilirken fungus 150 mg/L boya ile aynı koşullarda 3-24 saat boyunca inkübe edildi. En yüksek bakteriyel renk giderimleri statik koşullarda 72 saat inkübasyon sonrasında elde edildi ve maksimum renk giderim oranları A1, D1 ve D3 için sırasıyla 84, 83 ve %75 olarak saptandı. Bununla birlikte, *Pleurotus ostreatus* ile statik koşullarda, 24 saat inkübasyon sonrası RM 171 boyasının rengi %93 oranında giderildi. Benzer sonuçlar A1 hariç test edilen fungal ve bakteriyel suşların çalkalamalı çalışmalarından da elde edildi. A1, D1 ve D3 ile 150 rpm ve 72 saatte elde edilen maksimum renk giderim değerleri sırasıyla 30, 88 ve %89' dur. *Pleurotus ostreatus*' un en yüksek renk giderim aktiviteleri hem statik hem de çalkalamalı (150 rpm) koşullarda %93 idi. Ayrıca lakkaz varlığının tespiti amacıyla SBM ve 150 mg/L RB 171 içeren SBM'den elde edilen fungal kültür sıvılarının zimogram analizleri de yapıldı.

Anahtar Kelimeler: Bakteri; Renk giderimi; Tekstil boyası; Beyaz çürükçül fungus; Zimogram.

1. Introduction

Textile and dye industry wastewaters contain various textile dyes. Among these dyes; Reactive dyes are widely used in the textile industry due to their high binding capacities to cellulosic fibers. Therefore, reactive dyes are used at much higher rates than conventional dyes, and consequently there are more reactive dyes residues in textile factory wastewaters [1]. These wastewaters can be discharged to aquatic environments or soil without any or sufficient

treatment and negatively affect the aquatic life by reducing photosynthesis and also oxygen level in the water because of their intense color [2, 3]. However, the toxic and/or genotoxic effects of the dyes found in these wastewaters can adversely affect aquatic and terrestrial ecosystems, leading to serious environmental problems [4, 5]. For this reason, these wastewaters must be subjected to treatment process before being discharged to the environment. Because many synthetic dyes have antimicrobial effects, the textile dyes are generally resistant to biodegradation by the conventional biological wastewater treatment systems. Accordingly, studies on the color removal of textile wastewaters were generally carried out using physicochemical methods such as adsorption, precipitation, filtration, electrochemical oxidation in the literature [6]. But high amounts of sludge formation and insufficient degradation of dyes and removal of their degradation products are the disadvantages of physicochemical processes. However, their economic unfeasibility and environmental pollutant potential of the chemicals used are the other negative aspects of physicochemical processes [7-9]. In addition, the yield of color removal in traditional biological wastewater treatment methods is low. Thus, environmentally friendly and effective methods are needed for color removal of these dyes. In accordance with this purpose, many researchers are trying to find ecofriendly and efficient biological alternatives for dye decolorization [3, 10, 11]. Microbial decolorization may be an alternative method for color removal of these dyes, by selecting the appropriate methods and microorganisms. Bacteria and also white rot fungus can decolorize and detoxify of textile dyes [12, 13]. Laccases (E.C. 1.10.3.2) are produced by most of white rot fungi, and these enzymes are generally secreted extracellularly [14]. There are various reports in the literature that fungal laccases are responsible for the decolorization of many dyes [12, 13, 15, 16]. Accordingly, the spectrophotometric and zymogram analyzes were performed to detect if the laccase was the enzyme responsible from fungal dye decolorization. The aim of the study is to compare the decolorization potentials of Reactive Blue 171 (RB 171) dye by using the isolated bacteria like *Bacillus megaterium* A1, *Gordonia* sp. MC-D1, *Bacillus pumilus* D3, and a white rot fungus (*Pleurotus ostreatus*) in various conditions.

2. Materials and Methods

2.1. Microorganisms

In this study, three isolated bacteria and one white rot fungus were used for decolorization of Reactive Blue 171 (RB 171). Bacteria used in the studies were *Bacillus megaterium* strain A1 (A1, Accession number KC579390), *Gordonia* sp. MC-D1 (D1, Accession number KF663602), *Bacillus pumilus* D3 (D3, Accession number JX860616) while the fungus was *Pleurotus ostreatus* (*P. ostreatus*). Bacteria were transferred to fresh Nutrient Agar (NA) plates every

month and produced as pure solid culture at 30 °C. Similarly, the fungus was transferred to the plates containing fresh Sabouraud Dextrose Agar (SDA) every month and produced as pure filamentous fungal culture at 30 °C. The obtained pure bacteria and fungus cultures are kept in stock at 4 °C in a refrigerator of Inonu University Arts and Science Faculty Biology Department.

2.2. Textile dye

RB 171 (Fig. 1), an azo dye, was used as a textile dye in the studies. This dye was added to Nutrient Broth (NB) and Stock Basal Medium (SBM) at final concentration as 150 mg/L.

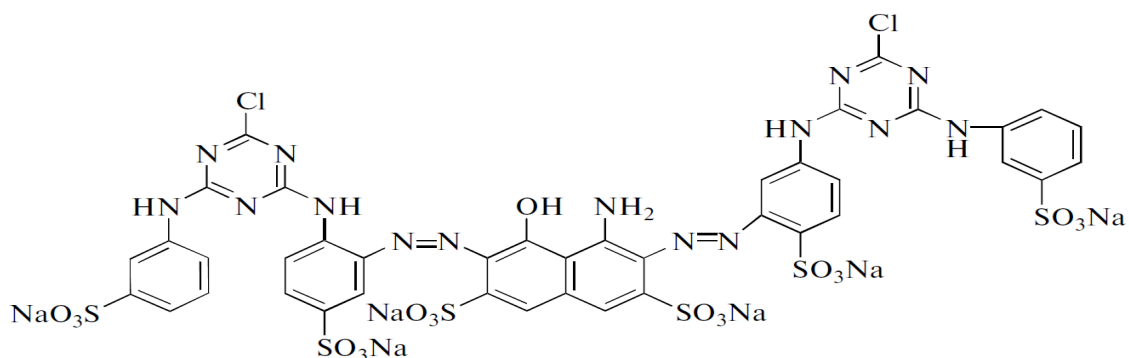


Figure 1: The chemical structure of RB 171 dye

2.3. Preparation of bacterial cultures

The samples taken from the stock solid cultures of A1, D1 and D3 produced on NA were transferred to NB media. Then these bacteria were incubated at 30 °C at 150 rpm for obtaining the liquid bacterial cultures to be used in the experiments.

2.4. Preparation of fungal cultures

The mycelia of *P. ostreatus* were incubated at 30 °C on slant Sabouraud Dextrose Agar (SDA) for 7 days. Then the mycelial suspensions were utilized for inoculum [17]. The fungus was precultured in 250 mL flasks containing 100 mL Sabouraud Dextrose Broth (SDB) at 150 rpm for 5 days and then homogenized at low speed under aseptic conditions. The homogenized mycelia were utilized as the inocula [18].

2.5. Dye decolorization studies

After preliminary preparations, 1 mL of A1, D1 and D3 liquid cultures were separately transferred to NB media containing 150 mg/L RB 171 at the final concentration. However, 2.5 mL of homogenized *P. ostreatus* mycelia were transferred into SBM containing 150 mg/L

RB 171 dye. SBM consisted of (g/L): KH_2PO_4 0.2; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 0.1; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0.05; $\text{NH}_4\text{H}_2\text{PO}_4$ 0.5; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 0.035; glucose 2; yeast extract 1. The samples were incubated at 30 °C under static and agitated (150 rpm) conditions.

Dye decolorization activities of microorganisms were determined by spectrophotometric (Shimadzu-UV-1601, UV/Visible) measurements at 24, 48 and 72th hours for the bacterial strains. However, the decolorization rates obtained with use of the fungus were detected at 3, 6, 9, 12, and 24th hours. The maximum wavelength of the RB 171 was determined as 619 nm via spectrophotometric scanning. The percents of dye decolorization activities obtained after the treatment of RB 171 dye with the microorganisms as stated above were calculated against the control groups at 619 nm. All experiments were performed at least 3 replicates.

2.6. Spectrophotometric analyzes of fungal laccase

The crude laccase activities of *Pleurotus ostreatus* culture fluids from SBM and SBM+150 mg/L RB 171 were detected by spectrophotometrically at 420 nm using ABTS [2,2-Azino-bis (3-ethylbenzthiazoline-6- sulfonic acid)] as the substrate [19].

2.7. Zymogram analysis of *Pleurotus ostreatus* culture fluids

Pleurotus ostreatus culture fluids obtained from two different media (SBM and SBM containing 150 mg/L RB 171) were used as the crude laccase enzyme sources. Native polyacrylamide gel electrophoresis (Native PAGE) were performed on native polyacrylamide gels. Laccase activities were determined on native polyacrylamide gels with ABTS as the substrate [12].

3. Results and Discussion

3.1. Bacterial dye decolorization

Among the tested bacterial species, D1 and D3 performed better color removals at agitated condition while A1 effectively decolorized RB 171 dye at static condition. This color removal clearly seen at Fig. 2 as macroscobically, and this experiment was also supported by spectrophotometric scanning (Fig. 3).

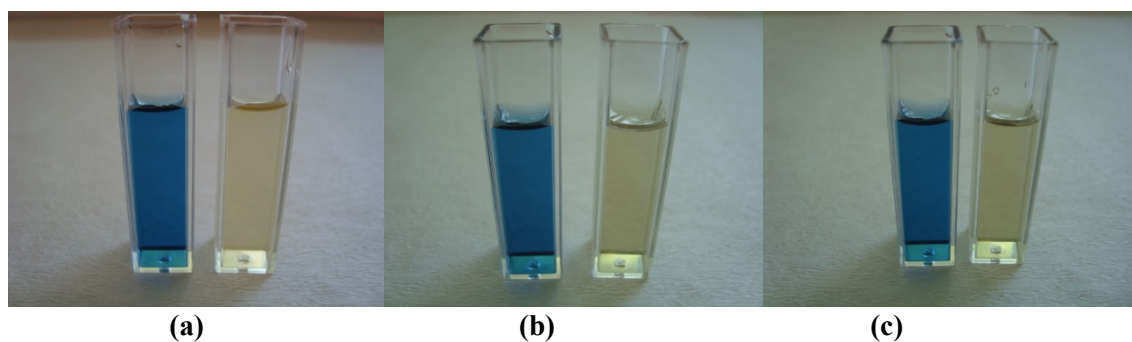


Figure 2: Macroscopic images of the decolorization of RB 171 under static condition by A1 (a) and under agitated condition by D1 (b) and D3 (c)

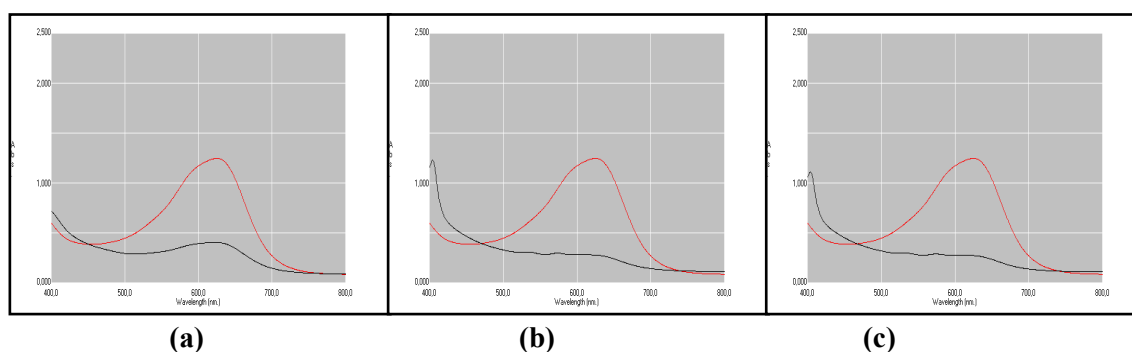


Figure 3: Spectrophotometric scanning images of the decolorization of RB 171 under static condition by A1 (a) and under agitated condition by D1 (b) and D3 (c)

The color removal percentages based on the time dependent absorbance measurements could be seen at Table 1. The highest color removal was generally obtained at 72th hour for all of the bacterial strains tested in both static and agitated conditions (Table 1). The most effective color removals were detected as 84, 83, and 75% in static conditions while the maximum decolorization rates were 30, 88, and 89% in agitated conditions for the strains of A1, D1, and D3, respectively.

Bacterial decolorization efficiency may vary depending on the type of bacteria tested, the dye, and whether the incubation conditions are static or agitated. For this purpose, different bacterial species and incubation conditions have been tested in decolorization of various dyes in many studies. Accordingly; in some of the studies in the literature, it is reported that higher decolorization rates were obtained in static conditions [20-22] and in others under shaking conditions [23-25]. This situation can be explained by the metabolic diversity of the tested bacterial species.

Six different isolates (DEC1-DEC6) of *Aeromonas hydrophila* were tested for decolorization of RB 171 dye at 100 mg/mL concentration in a work by Chen et al. [26]. It was stated that the highest decolorization activity was 80% for DEC 1, and DEC 6 isolates at the 7th day of incubation. The decolorization of azo dyes such as Reactive Violet 13 and Reactive Blue

171 at 50 and 100 ppm by *Pseudomonas stutzeri* in different media was studied by Gangavarapu and Ravuri [27]. According to the results, the maximum removal rates of Reactive Violet 13 and Reactive Blue 171 dyes (50 ppm) in Mineral Salt Media (MSM) were detected as 64% and 66%, whereas, 82% and 88% were measured as the highest decolorization values for the same dyes in Luria Broth (LB) media within 24 h, respectively. Moreover, the usability of *Bacillus cereus* RJVL 2514 for the decolorization of Reactive Violet 13 and Reactive Blue 171 dyes were investigated by the same researchers in another work. It was reported that the highest dye removal percents were detected for Reactive Violet 13 and Reactive Blue 171 at 50 ppm concentrations in MSM (92%, 91%) and also LB media (98%, 96%) within 24h, respectively [28].

Table 1: Decolorization rates of RB 171 by A1, D1, and D3 after different incubation periods

Time (h)	Dye Decolorization (%)					
	A1		D1		D3	
	Static	Agitated	Static	Agitated	Static	Agitated
24	68	14	23	77	22	78
48	81	26	49	88	41	88
72	84	30	83	88	75	89

3.2. Fungal dye decolorization

P. ostreatus effectively decolorized the color of RB 171 in both static and agitated conditions at all incubation times tested (Fig. 4).

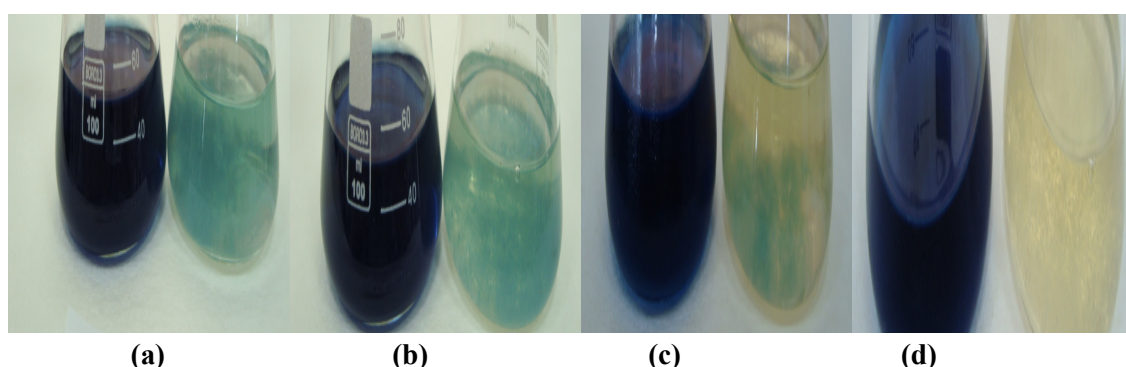


Figure 4: Macroscopic images of the decolorization of RB 171 by *P. ostreatus* under static conditions (a and c) and agitated conditions (b and d) after 3 h (a and b) and 24 h (c and d) incubation periods

Similarly, these color removals were also supported by spectrophotometric scanning (Fig. 5), and the percentages of the decolorization based on the time dependent absorbance measurements were showed at Table 2. The maximal color removal data were measured at 24th

hour as 93% for static and also agitated conditions. Vantamuri and Kaliwal tested the white rot fungus *Marasmius* sp. BBKAV79 for decolorization and degradation of Navy Blue HER (Reactive Blue 171) dye at 50 mg/L within 24 h under agitated condition. It was reported that the dye decolorization rate was 91.25% at 50 mg/L, but the decolorization value decreased to 37.50% at 150 mg/L [29]. In another work, Reactive Blue 171 dye removal rates were only 40% and 20% at 150 and 200 mg/L after the treatment with *Trichosporon beigelii* NCIM-3326 for 48 h under the static condition [30].

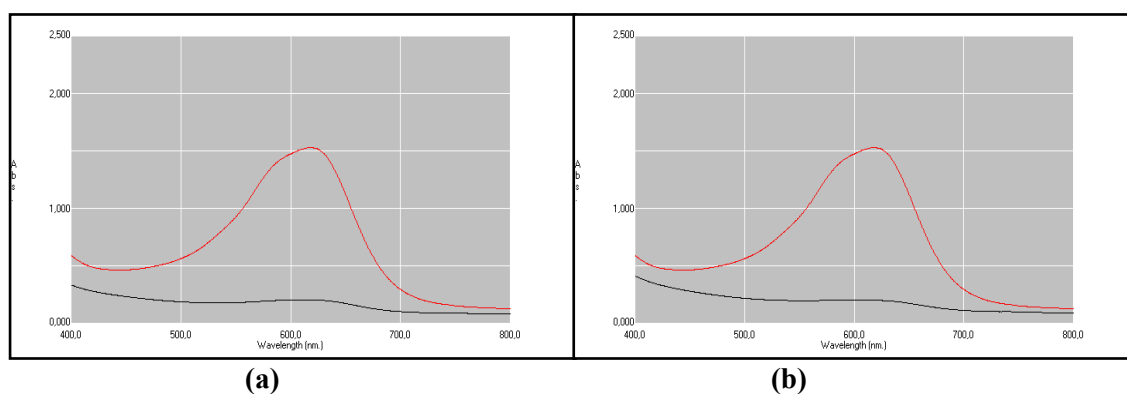


Figure 5: Spectrophotometric scanning images of the decolorization of RB 171 by *P. ostreatus* under static (a) and agitated conditions (b) after 3 h incubation period

Table 2: Decolorization rates of RB 171 by *P. ostreatus* after different incubation periods

Time (h)	Dye Decolorization (%)	
	Static	Agitated
3	87	87
6	89	90
9	91	91
12	91	92
24	93	93

3.3. Determination of *Pleurotus ostreatus* laccases by spectrophotometric analyzes

Some researchers were reported that laccase production was induced when fungi incubated in the media containing dye [31-33]. Similarly, in this study, the results of spectrophotometric measurements showed that the laccase activity in SBM+150 mg/L RB 171 was higher (ten-fold) than the enzyme activity in SBM. The increase of laccase activity in dye-containing SBM compared to SBM can be seen macroscopically in Fig. 6.

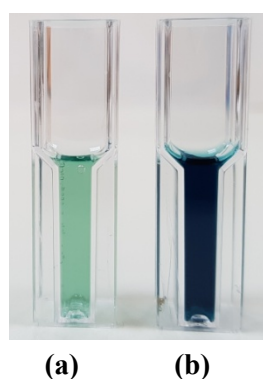


Figure 6: Spectrophotometric images resulting from the reaction of *Pleurotus ostreatus* culture fluids obtained from SBM (a) and SBM+150 mg/L RB 171 (b) with ABTS for 1 minute

3.4. Determination of *Pleurotus ostreatus* laccases by zymogram analysis

According to the zymogram analysis, one laccase band was observed in *Pleurotus ostreatus* culture liquid obtained from SBM while two laccase bands were observed in the fungal culture fluid from SBM containing 150 mg/L RB171 (Fig. 7). As mentioned above, higher laccase production in the dye-containing medium is also seen in the zymogram analysis. Accordingly, more intense laccase bands were detected in the culture fluid obtained from the dye-containing SBM compared to the culture fluid from SBM. This suggests that the dye tested (RB 171) induces the laccase production of *Pleurotus ostreatus*.

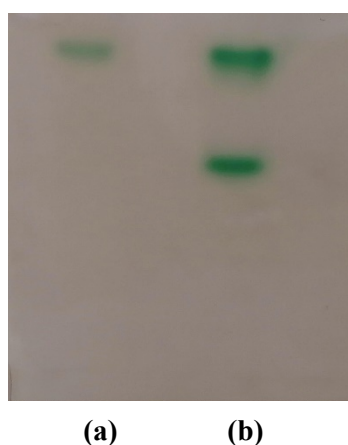


Figure 7: Zymogram of the crude laccases produced by incubation of *Pleurotus ostreatus* in SBM (a) and SBM+150 mg/L RB 171 (b)

4. Conclusion

This study performed by using three bacteria and a white rot fungus had shown that the color of RB 171 textile dye could be effectively removed in a short time and economically under static and agitated conditions. Furthermore, the spectrophotometric and zymogram analyzes showed that the laccase enzyme produced by *Pleurotus ostreatus* could play a role in

the decolorization process, and also the dye tested could also induce the laccase production. According to the data obtained from this study, an economical and rapid solution to environmental pollution can be found by removing the color of many textile dyes completely or at a high rate with various microorganisms. In addition, the production of an industrial enzyme, laccase can be increased during the decolorization process.

Acknowledgement

We would like to thank the Inonu University Research Fund Unit for their financial support (2013/66) throughout this study.

References

- [1] Karim, M.E., Dhar, K., Hossain, M.T., *Decolorization of textile reactive dyes by bacterial monoculture and consortium screened from textile dyeing effluent*, Journal of Genetic Engineering and Biotechnology, 16 (2), 375-380, 2018.
- [2] Robinson, T., Chandran, B., Nigam, P., *Removal of dyes from an artificial textile dye effluent by two agricultural waste residues, corncob and barley husk*, Environment International, 28 (1-2), 29-33, 2002.
- [3] Elshafei, A.M., Elsayed, M.A., Hassan, M.M., Haroun, B.M., Othman, A.M., Farrag, A.A., *Biodecolorization of six synthetic dyes by Pleurotus ostreatus ARC280 laccase in presence and absence of hydroxybenzotriazole (HBT)*, Annual Research and Review in Biology, 15, 1-16, 2017.
- [4] Hu, T.L., Wu, S.C., *Assessment of the effect of azo dye RP₂B on the growth of a nitrogen fixing cyanobacterium-Anabaena sp.*, Bioresource Technology, 77 (1), 93-95, 2001.
- [5] Dogan, E.E., Yesilada, E., Ozata, L., Yologlu, S., *Genotoxicity testing of four textile dyes in two crosses of Drosophila using wing somatic mutation and recombination test*, Drug and Chemical Toxicology, 28 (3), 289-301, 2005.
- [6] Saratale, R.G., Saratale, G.D., Chang, J.S., Govindwar, S.P., *Bacterial decolorization and degradation of azo dyes: A review*, Journal of the Taiwan Institute of Chemical Engineers, 42 (1), 138-157, 2011.
- [7] Banat, I.M., Nigam, P., Singh, D., Marchant, R., *Microbial decolorization of textile-dye-containing effluents: A review*, Bioresource Technology, 58 (3), 217-227, 1996.
- [8] Yesilada, O., Birhanli, E., Ozmen, N., Ercan, S., *Highly stable laccase from repeated-batch culture of Funalia trogii ATCC 200800*, Applied Biochemistry and Microbiology, 50 (1), 55-61, 2014.
- [9] Meerbergen, K., Willems, K.A., Dewil, R., Impe, J.V., Appels, L., Lievens, B., *Isolation and screening of bacterial isolates from wastewater treatment plants to decolorize azo dyes*, Journal of Bioscience and Bioengineering, 125 (4), 448-456, 2018.

[10] Singh, R.P., Singh, P.K., Singh, R.L., *Bacterial decolorization of textile azo dye acid orange by Staphylococcus hominis RMLRT03*, Toxicology International, 21 (2), 160-166, 2014.

[11] Parmar, N., Shukla, S.R., Microbial decolorization of reactive dye solutions, *Clean-Soil Air Water*, 43 (10), 1426-1432, 2015.

[12] Yeşilada, Ö., Birhanlı, E., Ercan, S., Özmen, N., Reactive dye decolorization activity of crude laccase enzyme from repeated-batch culture of *Funalia trogii*. *Turkish Journal of Biology*, 38 (1), 103-110, 2014.

[13] Boran, F., Birhanlı, E., Yeşilada, Ö., Özbey, E., Comparison of indigo carmine decolorization by *Pseudomonas aeruginosa* and crude laccase enzyme from *Funalia trogii*, *Turkish Journal of Biology*, 43 (1), 37-46, 2019.

[14] Akpınar, M., Urek, R.O., Induction of fungal laccase production under solid state bioprocessing of new agroindustrial waste and its application on dye decolorization, *3 Biotech*, 7 (2), 1-10, 2017.

[15] Jasińska, A., Góralczyk-Bińkowska, A., Soboń, A., Długoński J., Lignocellulose resources for the *Myrothecium roridum* laccase production and their integrated application for dyes removal. *International Journal of Environmental Science and Technology*, 16, 4811-4822, 2019.

[16] Liu, J., Sun, S., Han, Y., Meng, J., Chen, Y., Yu, H., Zhang, X., Ma, F., Lignin waste as co-substrate on decolorization of azo dyes by *Ganoderma lucidum*, *Journal of the Taiwan Institute of Chemical Engineers*, 122, 85-92, 2021.

[17] Kahraman, S., Kuru, F., Dogan, D., Yesilada, O., Removal of indigo carmine from an aqueous solution by fungus *Pleurotus ostreatus*, *Archives of Environmental Protection*, 38 (3), 51-57, 2012.

[18] Yesilada, O., Yildirim, S.C., Birhanli, E., Apohan. E., Asma, D., Kuru, F., The evaluation of pre-grown mycelial pellets in decolorization of textile dyes during repeated batch process, *World Journal of Microbiology and Biotechnology*, 26 (1), 33-39, 2010.

[19] Ulu, A., Birhanli, E., Boran, F., Köytepe, S., Yesilada, O., Ateş, B., Laccase-conjugated thiolated chitosan-Fe₃O₄ hybrid composite for biocatalytic degradation of organic dyes, *International Journal of Biological Macromolecules*, 150, 871-884, 2020.

[20] Tripathi, A., Srivastava, S.K., Ecofriendly treatment of azo dyes: biodecolorization using bacterial strains, *International Journal of Bioscience, Biochemistry and Bioinformatics*, 1(1), 37-40, 2011.

[21] Sneha, U., Poornima, R., Sridhar, S., Decolorization of synthetic textile dyes using *Pseudomonas putida*, *Journal of Chemical and Pharmaceutical Research*, 5(5), 219-225, 2013.

[22] Singh, R.P., Singh, P.K., Singh, R.L., Bacterial decolorization of textile azo dye acid orange by *Staphylococcus hominis* RMLRT03, *Toxicology International*, 21(2), 160-166, 2014.

[23] Li, G., Peng, L., Ding, Z., Liu, Y., Gu, Z., Zhang, L., Shi, G., Decolorization and biodegradation of triphenylmethane dyes by a novel *Rhodococcus qingshengii* JB301 isolated from sawdust, *Annals of Microbiology*, 64, 1575-1586, 2014.

[24] Ganapathy, B., Chandaran, I., Ponnaiah, P., Decolorizing palm oil mill effluent (pome) using plant polysaccharide degrading microorganisms isolated from soil, *Polish Journal of Environmental Studies*, 27 (2), 629-636, 2018.

[25] Barathi, S., Aruljothi, K.N., Karthik, C., Padikasan, I.A., Optimization for enhanced ecofriendly decolorization and detoxification of reactive blue 160 textile dye by *Bacillus subtilis*, *Biotechnology Reports* 28, e00522, 1-7, 2020.

[26] Chen, K.C., Wu, J.Y., Liou, D.J., Hwang, S.C.J., Decolorization of the textile dyes by newly isolated bacterial strains, *Journal of Biotechnology*, 101 (1), 57-68, 2003.

[27] Gangavarapu, V.L., Ravuri, J.M., Decolourization of reactive violet 13 and reactive blue 171 by *Pseudomonas stutzeri* RJVL 1514 isolated from dye contaminated soil, *International Journal of Microbiology Research*, 8 (5), 754-758, 2016.

[28] Gangavarapu, V.L., Ravuri, J.M., Decolourisation of reactive violet 13 and reactive blue 171 by *Bacillus cereus* RJVL 2514 isolated from dye contaminated soils, *International Journal of Environmental Sciences*, 7 (1), 30-39, 2016.

[29] Vantamuri, A.B., Kaliwal, B.B., Decolourization and biodegradation of navy blue her (reactive blue 171) dye from *Marasmius* sp. BBKAV79, *3 Biotech*, 7 (1), 1-7, 2017.

[30] Saratale, R.G., Saratale, G.D., Chang, J.S., Govindwar, S.P., Decolorization and biodegradation of textile dye navy blue her by *Trichosporon beigelii* NCIM-3326, *Journal of Hazardous Materials*, 166 (2-3), 1421-1428, 2009.

[31] D'Souza, D.T., Tiwari, R., Sah., A.K, Raghukumar, C., Enhanced production of laccase by a marine fungus during treatment of colored effluents and synthetic dyes, *Enzyme and Microbial Technology*, 38, 504-511, 2006.

[32] Sánchez-López, M.I., Vanhulle, S.F., Mertens, V., Guerra, G., Figueroa, S.H., Decock, C., Corbisier, A.M., Penninckx, M.J., Autochthonous white rot fungi from the tropical forest: potential of Cuban strains for dyes and textile industrial effluents decolourisation. *African Journal of Biotechnology*, 7 (12), 1983-1990, 2008.

[33] Sing, N.N., Husaini, A., Zulkharnain, A., Roslan, H.A., Decolourisation capabilities of ligninolytic enzymes produced by *Marasmius cladophyllus* UMAS MS8 on remazol brilliant blue R and other azo dyes, *Biomed Research International*, 2017, 1-8, 2017.



Estimation of Risk Measures for Transmuted Weibull Distribution

Caner TANIŞ^{1,*}

¹*Çankırı Karatekin University, Faculty of Science, Department of Statistics, 18100, Çankırı, Türkiye*
caner.tanis@gmail.com, ORCID: 0000-0003-0090-1661

Received: 23.04.2021

Accepted: 19.10.2021

Published: 31.12.2021

Abstract

In this paper, we tackle a problem of the estimation of some risk measures for transmuted Weibull distribution. In this regard, the maximum likelihood method is used to estimate the risk measures. We also obtain asymptotic confidence intervals based on the asymptotic distributions of maximum likelihood estimators of risk measures. Then, we consider a comprehensive Monte Carlo simulation study to assess the performances of these estimators at different sample sizes and parameter settings.

Keywords: Risk measures; Transmuted Weibull distribution; Point estimation; Interval estimation.

Dönüştürülmüş Weibull Dağılımı için Risk Ölçülerinin Tahmini

Öz

Bu çalışmada dönüştürülmüş Weibull dağılımı için bazı risk ölçülerinin tahmini problemini ele aldık. Bu bağlamda risk ölçülerini tahmin edebilmek için en çok olabilirlik yöntemi kullanıldı. Ayrıca risk ölçülerinin en çok olabilirlik tahmin edicilerinin asimptotik dağılımlarına dayalı yaklaşık güven aralıkları elde ettik. Sonrasında, bu tahmin edicilerin farklı örnek hacimleri ve parametre değerlerinde performanslarını değerlendirmek için geniş bir Monte Carlo benzetim çalışması tasarladık.



Anahtar Kelimeler: Risk ölçüleri; Dönüştürülmüş Weibull dağılımı; Nokta tahmini; Aralık tahmini.

1. Introduction

Transmuted Weibull distribution is suggested by [1] via quadratic transmutation map (QRTM). The QRTM is proposed by [2], and it is summarized by

$$F(x) = G(x)[1 + \lambda(1 - G(x))], \quad (1)$$

where $\lambda \in [-1,1]$, $G(x)$ refers the cumulative distribution function (CDF) of baseline distribution, and $F(x)$ denotes the CDF referring transmuted distribution which newly generated by the QTRM. Consider the baseline distribution Weibull distribution with CDF $G(x; \alpha, \beta) = 1 - \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}$ and the probability density function (PDF) $g(x; \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}$ then, the PDF and CDF of transmuted Weibull distribution are

$$F(x; \alpha, \beta, \lambda) = \left[1 - \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}\right] \left[1 + \lambda \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}\right], \quad (2)$$

and

$$f(x; \alpha, \beta, \lambda) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\} \left[1 - \lambda + 2\lambda \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}\right], \quad (3)$$

respectively, where $\beta > 0$ is a scale parameter, $\alpha > 0$ shape parameter and $\lambda \in [-1,1]$ [1]. In this study, the transmuted Weibull distribution is briefly denoted by $TW(\alpha, \beta, \lambda)$. The $TW(\alpha, \beta, \lambda)$ distribution has a potential to model the data sets in many fields such as, agriculture, biology, economics, actuarial sciences. Aryal and Tsokos [1] described some characteristic properties such as moments, variance, quantile function, reliability function, hazard function, order statistics of $TW(\alpha, \beta, \lambda)$ distribution. They emphasized that the hazard function can be increasing, decreasing or constant for $TW(\alpha, \beta, \lambda)$ distribution in [1]. In this case, it can be said that due to the flexible of the hazard function, the $TW(\alpha, \beta, \lambda)$ distribution has the potential to model many datasets having different hazard functions. Khan et al. [3] examined some statistical properties such as geometric mean, harmonic mean, entropies, mean deviation, L-moments of $TW(\alpha, \beta, \lambda)$ distribution. They also provided the log-transmuted Weibull regression model and its applications in [3]. For more details about $TW(\alpha, \beta, \lambda)$ distribution please see [1, 3].

Recently, many actuaries and insurance practitioners have focused on the measurement of financial risk. The risk measures manifest themselves in many different types of insurance problems including the determination of capital, and the estimation of possible maximum losses [4]. Therefore, we focus on the estimation of risk measures for $TW(\alpha, \beta, \lambda)$ distribution.

The main purpose of this paper to tackle the problem of point and interval estimation of risk measures for the $TW(\alpha, \beta, \lambda)$ distribution. We estimate the risk measures such as value at risk (VaR), tail value at risk (TVaR), tail variance (TV), and tail variance Premium (TVP) for the $TW(\alpha, \beta, \lambda)$ distribution. The rest of this study is organized as follows: In Section 2, we describe the risk measures for the $TW(\alpha, \beta, \lambda)$ distribution. Then, the maximum likelihood estimators (MLEs) of these risk measures and asymptotic confidence intervals based on MLEs are derived in Section 3. In Section 4, an extensive Monte Carlo simulation study designed to evaluate the performances of these estimators according to mean squares errors (MSEs) and bias.

2. Risk Measures

2.1. VaR measure

The VaR is one of the popular risk measures, and it quantifies maximum loss for investments. It is also known quantile risk measure. The VaR is generally used by firms and regulators in the financial sector in order to determine the amount of assests required to cover potential losses. The VaR of a random variable X is the q th quantile of its cdf, denoted by VaR_q , and it is defined by $VaR_q = Q(q)$ [5-7].

Let X be a random variable from $TW(\alpha, \beta, \lambda)$ distribution. The VaR is defined as follows:

$$VaR_q = \beta \left[-\log \left(1 - \frac{\lambda + 1 - \sqrt{(1 + \lambda)^2 - 4\lambda q}}{2\lambda} \right) \right]^{\frac{1}{\alpha}}, \tag{4}$$

where $q \in (0,1)$.

2.2. TVaR measure

TVaR, also known as tail conditional expectation is important risk measure. It measures the expected value of the loss given that an event outside a given probability level has occurred [6, 8, 9]. The TVaR of $TW(\alpha, \beta, \lambda)$ distribution is

$$\begin{aligned} TVaR_q &= \frac{1}{1-q} \int_{VaR_q}^1 x f(x) dx \\ &= \frac{1}{1-q} \Gamma \left(1 + \frac{1}{\alpha}, \left(\frac{VaR_q}{\beta} \right)^\alpha \right), \end{aligned} \tag{5}$$

where $\Gamma(., x)$ is incomplete gamma function, and VaR_q is given in Eqn. (4).

2.3. TV measure

The TV is important risk measure suggested by [10]. The TV of $TW(\alpha, \beta, \lambda)$ distribution is given by

$$\begin{aligned} TV_q X &= E(X^2 | X > x_q) - \{TVaR_q\}^2 \\ &= \frac{1}{1-q} \int_{VaR_q}^1 x^2 f(x) dx - \{TVaR_q\}^2 \\ &= \frac{\beta^2}{1-q} \Gamma\left(1 + \frac{2}{\alpha}, \left(\frac{VaR_q}{\beta}\right)^\alpha\right) - \{TVaR_q\}^2, \end{aligned} \tag{6}$$

where $TVaR_q$ is given in Eqn. (5).

2.4. TVP measure

The TVP is one of the significant measures of risk which play a crucial role in insurance sciences [9]. The TVP of $TW(\alpha, \beta, \lambda)$ distribution is

$$TVP_q = TVaR_q + \theta TV_q, \tag{7}$$

where $0 < \theta < 1$, $TVaR_q$ and TV_q are defined in Eqn. (5) and Eqn. (6), respectively.

3. Estimation of Risk Measures

3.1. Maximum Likelihood Estimation of risk measures

In order to obtain the MLEs of examined risk measures, we first derive MLEs of α, β and λ .

Let X_1, X_2, \dots, X_n be a random sample from $TW(\alpha, \beta, \lambda)$ distribution. Then log-likelihood function is given by [1, 2]

$$\ell(\Psi) = n \log\left(\frac{\alpha}{\beta}\right) - \sum_{i=1}^n \left(\frac{x_i}{\beta}\right)^\alpha + \sum_{i=1}^n \log\left(\frac{x_i}{\beta}\right)^{\alpha-1} + \sum_{i=1}^n \left[1 - \lambda + 2\lambda \exp\left(-\frac{x_i}{\beta}\right)^\alpha\right] \log \tag{8}$$

where $\Psi = (\alpha, \beta, \lambda)$. The MLE of Ψ is given by

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmax}}\{\ell(\Psi)\} \tag{9}$$

By using Eqns. (4)-(7) and invariant property of MLE, we can compute the MLEs of mentioned risk measures of VaR, TVaR, TV, and TVP by

$$\hat{V}aR_q = \hat{\beta} \left[-\log \left(1 - \frac{\hat{\lambda} + 1 - \sqrt{(1 + \hat{\lambda})^2 - 4\hat{\lambda}q}}{2\hat{\lambda}} \right) \right]^{\frac{1}{\hat{\alpha}}}, \tag{10}$$

$$\hat{T}VaR_q = \frac{1}{1-q} \Gamma \left(1 + \frac{1}{\hat{\alpha}}, \left(\frac{\hat{V}aR_q}{\hat{\beta}} \right)^{\hat{\alpha}} \right), \tag{11}$$

$$\hat{T}V_q = \frac{\hat{\beta}^2}{1-q} \Gamma \left(1 + \frac{2}{\hat{\alpha}}, \left(\frac{\hat{V}aR_q}{\hat{\beta}} \right)^{\hat{\alpha}} \right) - \{ \hat{T}VaR_q \}^2, \tag{12}$$

and

$$\hat{T}VP_q = \hat{T}VaR_q + \theta \hat{T}V_q, \tag{13}$$

respectively.

3.2. Asymptotic confidence interval

In this subsection, we provide the asymptotic variances and covariances of the MLEs $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\lambda}$ by entries of the inverse of the observed Fisher information matrix is given by

$$I^{-1}(\hat{\Psi}) = \begin{pmatrix} -\frac{\partial^2 \ell(\Psi)}{\partial \alpha^2} & -\frac{\partial^2 \ell(\Psi)}{\partial \alpha \partial \beta} & -\frac{\partial^2 \ell(\Psi)}{\partial \alpha \partial \lambda} \\ -\frac{\partial^2 \ell(\Psi)}{\partial \beta \partial \alpha} & -\frac{\partial^2 \ell(\Psi)}{\partial \beta^2} & -\frac{\partial^2 \ell(\Psi)}{\partial \beta \partial \lambda} \\ -\frac{\partial^2 \ell(\Psi)}{\partial \lambda \partial \alpha} & -\frac{\partial^2 \ell(\Psi)}{\partial \lambda \partial \beta} & -\frac{\partial^2 \ell(\Psi)}{\partial \lambda^2} \end{pmatrix}.$$

Now, we can obtain the variance of $Var(\hat{R})$ using delta method as $Var(\hat{R}) = B' I^{-1}(\hat{\Psi}) B$ where R denotes one of the risk measures (VaR, TVaR, TV, TVP), \hat{R} is the MLE of R and $B' = \left(\frac{\partial R}{\partial \alpha}, \frac{\partial R}{\partial \beta}, \frac{\partial R}{\partial \lambda} \right)$. By using the MLEs of α, β and λ $Var(\hat{R})$ can be estimated. The asymptotic $100(1 - \eta)\%$ confidence interval of risk measures by

$$\left(\hat{R} - z_{1-\frac{\eta}{2}} \sqrt{Var(\hat{R})}, \hat{R} + z_{1-\frac{\eta}{2}} \sqrt{Var(\hat{R})} \right)$$

where z_{η} 100 η^{th} percentile of $N(0,1)$.

4. Simulation Study

In this section, we design a comprehensive Monte Carlo simulation study to assess the performances of MLEs, of risk measures according to biases and MSEs. The simulation study is performed based on 5000 repetitions. We consider the sample size 25, 50, 100, 200, 500 and two parameter settings as follows: $(\alpha = 0.5, \beta = 1.5, \lambda = 0.2)$, $(\alpha = 1, \beta = 2, \lambda = 0.5)$. The results of simulation study are presented in Tables 1-2. Table 1 provides average of biases and MSEs of

risk measures such as VaR, TVaR, TV, and TVP. Also, Table 2 presents the average of lengths and coverage probabilities (CPs) of these risk measures.

Table 1: Average biases and MSEs of risk measures

n	α	β	λ	Sig.level	θ	bias				MSE			
						VaR	TVaR	TV	TVP	VaR	TVaR	TV	TVP
25						-0.0629	-0.0096	-0.006	-0.012	0.0883	0.0018	0.0004	0.0026
50						-0.0406	-0.0043	-0.0035	-0.0061	0.0433	0.0008	0.0001	0.0011
100	0.5	1.5	0.2	0.5	0.5	-0.0195	-0.0006	-0.0021	-0.0017	0.0188	0.0003	0.00008	0.0005
200						-0.0116	0.0005	-0.0019	-0.0004	0.0093	0.0002	0.00007	0.0003
500						-0.006	0.001	-0.0013	0.0003	0.0037	0.0001	0.00004	0.0001
25						-0.0463	-0.0405	0.0485	-0.0162	1.0149	0.0841	0.0505	0.0407
50						0.0116	0.002	0.0004	0.003	0.4698	0.0368	0.0117	0.0234
100	0.5	1.5	0.2	0.75	0.5	0.0089	0.0208	-0.0152	0.0132	0.231	0.0199	0.0047	0.014
200						0.0218	0.028	-0.0201	0.018	0.119	0.0127	0.0029	0.0093
500						0.0172	0.0249	-0.0201	0.014	0.0485	0.0069	0.0018	0.0054
25						-0.03	-0.0151	-0.0013	-0.016	0.0574	0.0019	0.001	0.0034
50						-0.0151	-0.0096	0.0022	-0.0082	0.0253	0.0008	0.0004	0.0014
100	1	2	0.5	0.4	0.6	-0.006	-0.006	0.0029	-0.0042	0.0149	0.0004	0.0002	0.0007
200						-0.0056	-0.0047	0.0027	-0.003	0.0075	0.0002	0.0001	0.0003
500						-0.0049	-0.0029	0.0014	-0.002	0.0024	0.0001	0.0001	0.0001
25						0.0448	-0.1257	0.7922	0.3495	0.7267	0.4067	3.8147	0.4188
50						0.0327	-0.0684	0.501	0.2321	0.2071	0.2191	1.8846	0.2019
100	1	2	0.5	0.8	0.6	0.0095	-0.0226	0.2892	0.1508	0.6506	0.1288	1.0662	0.1183
200						0.01287	0.0029	0.1564	0.0967	0.0663	0.0841	0.7361	0.083
500						-0.0022	0.0064	0.0941	0.0628	0.0196	0.0487	0.5343	0.0601

Table 2: Average lengths and CPs of risk measures

n	α	β	λ	Sig.level	θ	length				CP			
						VaR	TVaR	TV	TVP	VaR	TVaR	TV	TVP
25						1.0773	0.1596	0.075	0.19	0.9138	0.923	0.84	0.9208
50						0.7571	0.1129	0.0506	0.134	0.9316	0.9364	0.848	0.939
100	0.5	1.5	0.2	0.5	0.5	0.5284	0.0804	0.036	0.0958	0.943	0.9428	0.8404	0.9424
200						0.374	0.0596	0.0278	0.0716	0.9494	0.945	0.8436	0.944
500						0.2387	0.0405	0.0204	0.0494	0.949	0.9304	0.8322	0.9264
25						3.741	1.1613	0.8475	0.8683	0.8836	0.9104	0.8476	0.9192
50						2.6325	0.8108	0.4831	0.6535	0.9074	0.926	0.865	0.9344
100	0.5	1.5	0.2	0.75	0.5	1.8941	0.595	0.3286	0.4978	0.9292	0.9316	0.8786	0.9404
200						1.3459	0.4449	0.242	0.3815	0.9356	0.9262	0.8838	0.933
500						0.8693	0.3003	0.1645	0.2651	0.9452	0.9342	0.8932	0.9316
25						0.7477	0.1629	0.1262	0.2025	0.9172	0.9494	0.8898	0.9352
50	1	2	0.5	0.4	0.6	0.535	0.1113	0.0915	0.1362	0.9358	0.9582	0.9038	0.9454
100						0.3797	0.0779	0.0687	0.0934	0.9394	0.9498	0.9094	0.9472

200						0.2704	0.057	0.053	0.0658	0.9438	0.9458	0.9142	0.953
500						0.1716	0.0379	0.0374	0.0412	0.9496	0.9172	0.9126	0.9554
25						1.928	2.7358	8.9104	2.8183	0.898	0.9482	0.9524	0.9646
50						1.4025	2.0489	6.4002	1.9501	0.9094	0.9544	0.9558	0.9606
100	1	2	0.5	0.8	0.6	1.0116	1.5629	4.8082	1.4423	0.9282	0.9498	0.9314	0.8688
200						0.7256	1.2158	3.756	1.1273	0.9384	0.9286	0.881	0.7732
500						0.4699	0.8916	2.8454	0.8698	0.9442	0.8804	0.8248	0.7528

From Tables 1-2, It is seen that as the sample size increases, the MSEs and biases of risk measures decrease and approach zero. Also, we observed that the lengths decrease and CPs approach 0.95 as expected. In the case of high significance level (it is defined in Eqn. (4) as q), the MSEs and biases of the risk measures are larger than in other cases.

5. Conclusion

In this study, we provide some risk measures such as VaR, TVaR, TV, and TVP for $TW(\alpha, \beta, \lambda)$ distribution. We use the maximum likelihood method to estimate these risk measures. Then, we obtain MLEs of examined risk measures using the invariant property of MLE. Not only point estimates of risk measures but also interval estimates are discussed. Approximate confidence intervals based on the asymptotic distribution of MLE were obtained. Monte Carlo simulations are performed to observe the performance of the estimators according to MSE and bias. From the results of the simulation study, it is observed that the MLEs of risk measures provided the estimation procedures.

References

- [1] Aryal, G.R., Tsokos, C.P., *Transmuted Weibull distribution: a generalization of the Weibull probability distribution*, European Journal of Pure and Applied Mathematics, 4(2), 89-102, 2011.
- [2] Shaw, W.T., Buckley, I.R., *The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map*, arXiv:0901.0434, 2009.
- [3] Khan, M.S., King, R., Hudson, I.L., *Transmuted Weibull distribution: properties and estimation*, Communications in Statistics-Theory and Methods, 46(11), 5394-5418, 2017.
- [4] Dowd, K., Blake, D., *After VaR: the theory, estimation, and insurance applications of quantile-based risk measures*. Journal of Risk and Insurance, 73(2), 193-229, 2006.
- [5] Artzner, P., *Application of coherent risk measures to capital requirements in insurance*, North American Actuarial Journal 2(2), 11-25, 1999.
- [6] Afify, A.Z., Gemeay, A.M., Ibrahim, N.A. *The heavy-tailed exponential distribution: Risk measures, estimation, and application to actuarial data*, Mathematics, 8(8) 1276, doi:10.3390/math8081276, 2020.

[7] Taniş, C., *On transmuted power function distribution: characterization, risk measures, and estimation*, *Journal of New Theory*, (34), 72-81, 2021.

[8] Klugman, S.A., Panjer, H.H., Willmot, G.E., *Loss models: from data to decisions*, vol. 715, John Wiley & Sons, 2012.

[9] Al-Babtain, A.A., Gemeay, A.M., Afify, A.Z., *Estimation methods for the discrete poisson-lindley and discrete lindley distributions with actuarial measures and applications in medicine*, *Journal of King Saud University*, 33(2), 1-11, 2020. <https://doi.org/10.1016/j.jksus.2020.10.021>.

[10] Landsman, Z., *On the tail mean–variance optimal portfolio selection*, *Insurance: Mathematics and Economics*, 46, 547–553, 2010.



The Effective and Eco-friendly Tea Fungus for the Biosorption of Dye Pollutant from Aqueous Solutions

Tuğba ALP ARICI^{1,*}

¹*Kütahya Dumlupınar University, Emet Vocational School, Department of Chemical Technology, 43700, Kütahya, Türkiye*
tugba.alp@dpu.edu.tr, ORCID: 0000-0003-3927-9849

Received: 01.07.2021

Accepted: 25.10.2021

Published: 31.12.2021

Abstract

The present study has reported the production of the environmentally friendly Kombucha tea fungus (TF), as well as its investigation as a potential biosorbent to remove cationic methyl violet (MV) dye from the aqueous solutions. TF was characterized by using FT-IR, thermal analysis and zeta potential measurements. The biosorption of MV was significantly dependent on the pH of the solutions and the highest removal for MV was observed at pH=9.0 which was also proved with zeta potential measurement. Biosorption equilibrium was established in 10 min and a definition of the overall rate-controlling step in the biosorption of MV onto TF was carried out. The biosorption data were in good agreement with the pseudo-second-order kinetic and Langmuir isotherm models. The maximum biosorption capacity of TF for MV was determined as 1180.09 mg g⁻¹. In addition, an effective biosorption performance was observed even in the presence of high foreign ion concentrations. The results indicated that TF could be utilized as a highly efficient biosorbent for MV biosorption from aqueous solutions.

Keywords: Tea fungus; Biosorption; Methyl violet; Isotherms.



Sulu Çözeltilerdeki Boyarmadde Kirliliğinin Biyosorpsiyonu için Etkili ve Çevre Dostu Çay Mantarı

Öz

Bu çalışma, çevre dostu Kombucha çay mantarının (TF) üretimini ve sulu çözeltilerden katyonik metil viyole (MV) boyarmaddesinin uzaklaştırılmasında biyosorban olarak potansiyelinin araştırılmasını sunmaktadır. TF, FT-IR, termal analiz ve zeta potansiyel ölçümleri kullanılarak karakterize edilmiştir. MV'nin biyosorpsiyonu, çözeltilerin pH'sına önemli ölçüde bağlıdır ve en yüksek MV giderimi, zeta potansiyeli ölçümü ile de kanıtlanan pH=9.0'da gözlenmiştir. Biyosorpsiyon 10 dakika içerisinde dengeye ulaşmış ve MV'nin TF üzerine biyosorpsiyonunda hız sınırlayıcı adım belirlenmiştir. Biyosorpsiyon verileri, yalancı ikinci dereceden kinetik ve Langmuir izoterm modelleri ile iyi bir uyum göstermiştir. TF'nin MV için maksimum tek tabakalı biyosorpsiyon kapasitesi 1180.09 mg g⁻¹ olarak belirlenmiştir. Ayrıca, yüksek yabancı iyon derişimlerinin varlığında dahi etkili bir biyosorpsiyon performansı gözlenmiştir. Sonuçlar, TF'nin sulu çözeltilerden MV biyosorpsiyonu için yüksek verimli bir biyosorban olarak kullanılabileceğini göstermiştir.

Anahtar Kelimeler: Çay mantarı; Biyosorpsiyon; Metil viyole; İzotermler.

1. Introduction

Kombucha is a traditional probiotic drink that is prepared from the fermentation of sugar added black or green tea with lactic acid bacteria, acetic bacteria, and yeasts [1-3]. After fermentation, a symbiotic culture namely, *Medusomyces gisevii* (tea fungus) layer forms that floats on the tea in 7-15 days [4]. Under the culture, there is Kombucha tea with a pH in the range of 2.5-3.0. Kombucha tea is originated from China (Manchuria) and after gained rapid popularity and spread all over the world, mainly to Russia and Germany [5]. Kombucha tea contains several components with known benefits for human health such as water-soluble vitamins, minerals, amines, purines, flavonoids, catechins, tannins, essential oils and hydrolytic enzymes [6-8]. In particular, it has anti-diabetic, anti-carcinogenic and anti-inflammatory effects. Moreover, it plays a healing role in the treatment of gastric ulcers, lowering high cholesterol, improvement of the liver and the immune system [9]. As the demand for Kombucha increases, to gain more benefits from obtained culture, the application potential in various industrial areas such as textiles and cosmetics has been investigated [10-12]. In addition, the use of this fungus to eliminate the environmental damage of pollutants in industrial wastewater has attracted attention as a popular

biotechnological research subject. Among the pollutants, dyes are substances that have a high potential for use in most various industrial areas such as food, leather, textile, paper and plastics [13,14]. Methyl violet (MV) which is a cationic dye with high color intensity is extensively used as a purple dye for textile products such as silk, wool, nylon, cotton and in printing ink. In addition, MV has been widely utilized as a pH indicator [15]. It is employed as an ingredient in Gram's stain for classifying bacteria and also a bacteriostatic agent in the medical field. MV has a toxic organic structure which causes diverse health problems for humans such as cancer, skin irritation, respiratory and kidney failure and permanent blindness [16, 17]. Because of all its hazard mentioned above, MV needs to be effectively removed from the aqueous solutions before being discharged into the environment. There are diverse methods such as biological, physical and chemical that are utilized to eliminate pollutants from the aqueous solutions [18, 19]. However, their uses have many restraints such as application difficulty, process cost, low yield and occurrence of huge amount of sludge. Biosorption is a popular pollution treatment method that is ecofriendly, economic, easy to application, non-toxic and high efficiency [20, 21]. Biological origin materials such as bacteria [22], fungus [23], alg [24] and agricultural wastes [25] are used as biosorbents in the biosorption.

In this context, tea fungus (TF) was used as a biosorbent for the biosorption of methyl violet (MV) from the aqueous solutions. The characterization of the TF was carried out with Infrared Spectroscopy (IR), Thermogravimetric (TG) analysis and zeta potential measurements. The effects of pH, TF amount, temperature, contact time, initial MV concentration and ionic strength on the biosorption were investigated to determine the best biosorption conditions. The biosorption kinetics and equilibrium isotherms data were modeled and the mechanism of the biosorption process was evaluated. The maximum biosorption capacity of TF towards MV was compared with other studies in the literature.

2. Materials and Methods

2.1. Chemicals and preparation of TF

MV was used as a dye pollutant in the biosorption process. Its maximum absorption wavelength (λ_{\max}) and molecular weight were 584 nm and 393.9 g mol⁻¹, respectively. HCl and NaOH solutions in various concentrations were preferred to adjust the pH values. Deionized water was utilized for dissolving dye and preparing the target concentration of solutions. All the reagents provided for the biosorption were analytical grade and purchased from Merck. The Kombucha tea with fungus, sugar and black tea leaves were provided from the local market in Turkey. The production process of the fungus in biosorption included the following steps.; First, 3 L of water

was put in a glass bottle and boiled, then 45 g L^{-1} sugar was added into this water. After the sugar was dissolved, 10 g of black tea leaves were added into the solution and brewed for 15 min. Subsequently, the solid-liquid phase was separated by filtration and the solution was kept in a glass bottle. Kombucha tea and fungus were added to the previous solution at room temperature and then left to the incubation process for about 15 days. At the end of this process, a disc-shaped, smooth new fungus (TF) was obtained on the Kombucha tea. Before being used in the biosorption process, this fungus was washed with plenty of deionized water to eliminate any possible impurities. It was put in an oven at 65°C for drying. Then, it was ground and sieved through a $200 \mu\text{m}$ sieve to obtain a homogeneous grain-sized biosorbent (Fig. 1). Finally, it was stored in a capped bottle for utilization in all experiments.

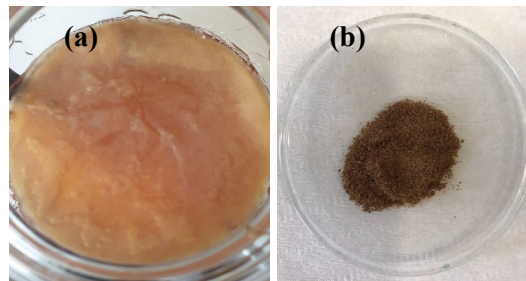


Figure 1: Kombucha tea fungus before (a) and after (b) drying, grinding and sieving

2.2. Biosorption studies

The various parameters (pH, contact time, biosorbent amount, temperature, initial pollutant concentration and ionic strength) were investigated for the biosorption of MV. The pH experiments were carried out by 20 mL of a 100 mg L^{-1} MV solution with 0.625 g L^{-1} of TF and the pH value was attentively adjusted between 2.0 and 10.0 by adding a small volume of NaOH or HCl solutions using a Hanna pH meter. MV solutions were stirred in a 25 mL glass bottle for 60 min. To determine the optimum biosorbent amount, TF was weighed from 0.250 to 1.25 g L^{-1} and added into 100 mg L^{-1} MV at pH 9.0. The temperature was changed between 15 and 40°C at an initial MV concentration of 100 mg L^{-1} . In order to execute kinetic analysis, contact time for MV biosorption onto TF was applied in the range of 1–90 min. The effect of initial MV concentration on the biosorption capacity of TF was examined by changing the concentrations from 50 to 800 mg L^{-1} at the confirmed conditions. To investigate the ionic strength effect on the biosorption of MV, the biosorption experiments were occurred with 100 mg L^{-1} of MV solutions containing NaCl in the concentration range of $0.02\text{--}0.1 \text{ mol L}^{-1}$ under the optimum biosorption conditions. After each application of biosorption experiment, the TF was separated from the medium via filtration. All experiments were repeated in triplicate. The obtained samples were measured by using a UV-vis spectroscopy (Shimadzu UV-2600 spectrophotometer). The obtained

data were evaluated by biosorption amount and biosorption (%) of MV using the equations which were given as below

$$q_e = \frac{V \times (C_0 - C_e)}{m} \quad (1)$$

$$\text{Biosorption (\%)} = \frac{C_0 - C_e}{C_0} \quad (2)$$

In this equation, q_e : The amount of MV biosorbed on TF at equilibrium (mg g^{-1}), m : The amount of TF (g), C_0 : The initial concentration (mg L^{-1}) of MV, C_e : equilibrium concentration (mg L^{-1}) of MV, V : The volume of MV solution (L).

2.3. Characterizations

FT-IR analyses were (Bruker Tensor 27 FT-IR spectrometer) carried out in the 400–4000 cm^{-1} range to identify the surface functional groups of TF before and after MV biosorption. Thermogravimetric analysis (TGA) of TF in the temperature range of 30–700°C was done by using Perkin Elmer Diamond TG/DTA instrument with a rate of 10°C min^{-1} under the air atmosphere. The pH-dependent surface charges of TF were specified using a zeta potential measurement (ZEN 3600 Model Zetasizer Nano-ZS). For this measurement, 0.50 g L^{-1} of TF was added into 20 mL of deionized water and the pH was changed in the range of 2-10.

3. Results and Discussion

3.1. Characterization of FT

TF was characterized by FT-IR spectroscopy and thermal analysis. As mentioned in the introduction, TF has several components in its ingredient. In the FT-IR spectrum of TF, the broad band appearing at 3381 cm^{-1} can be owing to $\nu(\text{O-H})$ stretching vibrations of phenolic alcohols. Aliphatic $\nu(\text{C-H})$ stretching vibrations appear at 2937 cm^{-1} and 2898 cm^{-1} . The picks observed at 1727 cm^{-1} and 1638 cm^{-1} are assigned to the carbonyl $\nu(\text{C=O})$ and imine $\nu(\text{C=N})$ stretching vibrations, respectively. Aromatic $\nu(\text{C=C})$ stretching vibration is observed at 1424 cm^{-1} . After MV biosorption, the pick observed at 3381 cm^{-1} shifts to high wavenumber and also the intensity of aliphatic $\nu(\text{C-H})$ stretching vibration decreases. They can be due to the fact that phenolic $-\text{OH}$, imine and aliphatic $-\text{CH}-$ groups are covered by dye molecules. The carbonyl pick at 1727 cm^{-1} disappears and some peaks shift (Fig. 2). They show the biosorption of MV onto TF. TG/DTG curves are acquired to determine the decomposition stage of TF. As illustrated in Fig. 3, the first stage in the temperature range 30-100°C is the elimination of water molecules. After this step, the

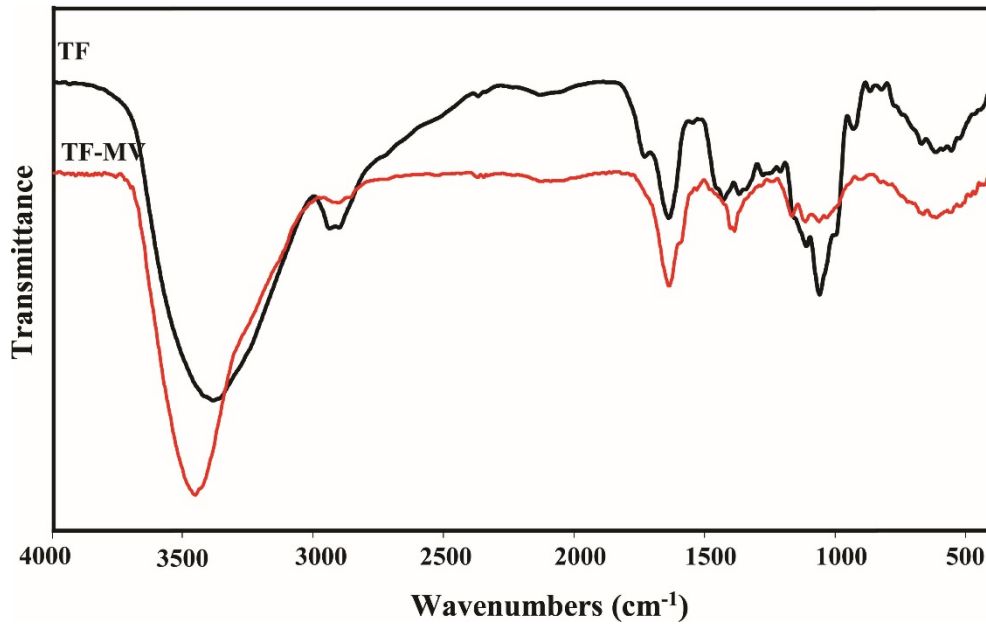


Figure 2: FT-IR spectra of TF and TF-MV

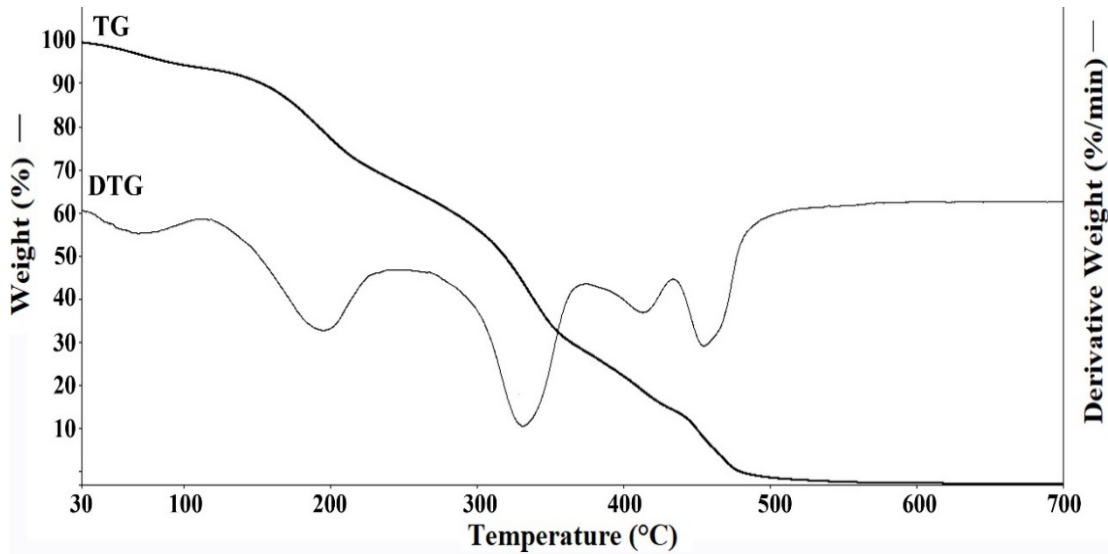


Figure 3: TG/DTG curves of TF

3.2. Biosorption parameters

3.2.1. Effect of pH

Biosorption amount is generally affected by the pH of the dye solution because of affecting the functional groups on the biosorbent and the ionization degree of the dye molecules. Thus, it is a major experimental parameter that should be considered. Fig. 4 displays the results for the

biosorption of MV onto TF when the pH is changed at a range of 2.0–10.0 of the MV solutions at room temperature. The biosorption capacity of TF for MV removal increases with the increasing pH of the biosorption medium and reaches an equilibrium value of 154.46 mg g^{-1} at pH 9.0 (Fig. 4). After this pH value, no significant change appears in the MV biosorption ability of TF. Therefore, pH 9.0 is selected as the pH value of the biosorption media for further experiments. As the acidity of the biosorption medium increases, the positive charge density increases due to the protonation of the functional groups on the biosorbent surface. Thus, the interaction of cationic dye molecules with the biosorbent surface is restricted, resulting in lower biosorption. The negative charge density related to the concentration of OH^- ions rises on the biosorbent surface when the pH increases. Thus, the electrostatic interactions between cationic dye MV and TF are enhanced. The surface charge density values of TF in deionized water at various pHs which support the obtained data from the pH parameter are presented in Fig. 4. It is showed that the zeta potential of TF decreases considerably from -1.60 mV at pH 2.0 to -30.40 mV at pH 10.0. TF has no points of zero charge and the biosorbent surface has the highest negative charge density (-35.40 mV) at pH=9.0 where the maximum amount of the biosorption is achieved.

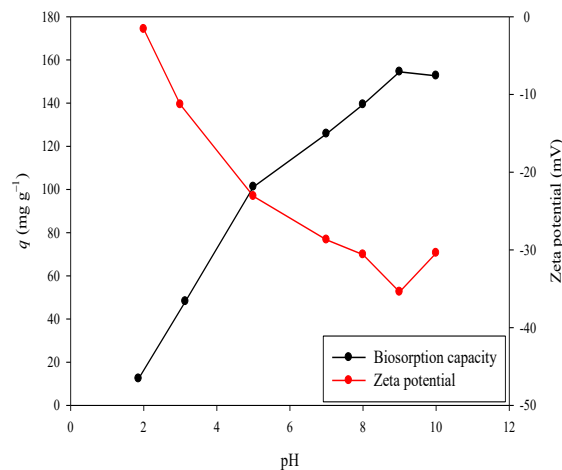


Figure 4: Zeta potential of TF in deionized water and pH effect on the MV biosorption

3.2.2. Effect of TF amount and temperature

The biosorbent amount is another important parameter that highly affects the biosorption process yield and cost. It has been indicated that the pollutants are removed at high efficiency with the increase in the biosorbent amount which gives rise to the increase of active sites on the biosorbent surface [26]. Hence, the effect of TF amount on the MV biosorption was examined in the range of $0.25\text{--}1.00 \text{ g L}^{-1}$ with pH of 9.0 (Fig 5a). As seen in Fig 5a, the biosorption yield

increases from 84.75% to 96.83% as the amount of TF increases from 0.250 to 0.500 g L⁻¹. It is noted that after 0.750 g L⁻¹ amount of TF, the biosorption yield decreases to 86.66%. This decrease can be explained with the fact/inference that the biosorbent can aggregate at higher amounts and its active binding sites can be lower for interaction with dye molecules. With a very economic amount of 0.500 g L⁻¹, the biosorption reaches its highest efficiency. For this reason, the optimum amount of TF to be used in further biosorption experiments was determined as 0.500 g L⁻¹.

To investigate whether the biosorption was temperature-dependent, the biosorption experiments were realized at the temperature range of 15–40°C at the initial MV concentration of 100 mg L⁻¹, TF amount of 0.500 g L⁻¹ and the contact time of 60 min. As depicted in Fig. 5b, the biosorption amounts of MV change from 190.20 to 191.10 mg g⁻¹ with the increase of the temperature from 15 to 40°C. The results show an insignificant effect of temperature in the biosorption ability of TF for MV removal. Similar results for the effect of temperature are observed in the literature [27-28]. As a result, this system is very advantageous as it has the potential for the elimination of MV from the solution with very high efficiency at any temperature.

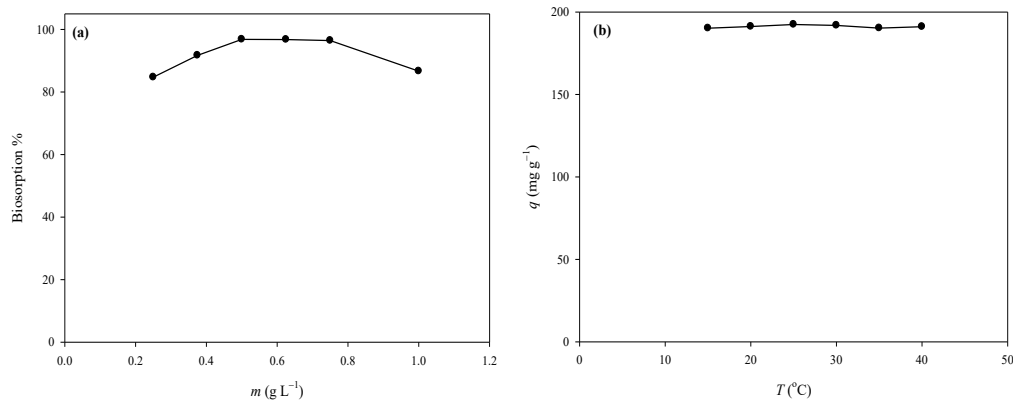


Figure 5: The effect of TF amount (a) and temperature (b) on the MV biosorption

3.2.3. Effect of contact time and biosorption kinetics

To provide information about the biosorption behavior of MV onto TF as a function of contact time, the biosorption experiments were performed in the contact time range of 1–90 min at 20°C. The effect of contact time on biosorption is demonstrated in Fig 6a. MV is rapidly removed from the aqueous solution in 10 min and after this point, the biosorption equilibrium is established. This rapid biosorption is due to the binding of the MV molecules to a great number of vacant active biosorption sites on TF. The biosorption amount of MV rises from 180.52 to 188.86 mg g⁻¹ in the time range of 1-10 min. The biosorption reaching equilibrium quite fast is

an important advantage that supports the practicality of the method. A similar study was also reported for biosorption equilibrium time as 10 min [29]. The biosorption kinetic applications are commonly utilized to reveal the control mechanism of the biosorption process.

In this study, to explore the MV biosorption nature, the obtained experimental data from contact time were applied to Lagergren-first-order [30], pseudo-second-order [31] and intra-particle diffusion kinetic models [32]. The kinetic parameters obtained from fitting the experimental data to these three kinetic model equations are presented in Table 1. When the highest R^2 values are compared, it is easy to see that the R^2 of the pseudo-second-order model for TF ($R^2=0.999$) is significantly higher than Lagergren-first-order ($R^2=0.597$) and intra-particle diffusion kinetic models ($R^2=0.787$). The plot of the pseudo-second-order kinetic model for the biosorption of MV onto TF is exhibited in Fig. 6b. The obtained data for the biosorption of MV onto TF are better fitted to the pseudo-second-order kinetic model. This finding indicates that the rate-limiting step in the biosorption process can be chemical biosorption which includes electron exchange or electron sharing between MV and TF.

Table 1: Kinetic parameters for the biosorption of MV

Model	Parameters	Values
Lagergren-first-order model	k_1	1.97×10^{-2}
$\log(q_e - q_t) = \log(q_e) - \frac{k_1}{2.303} t$	q_e	7.96
	R^2	0.597
Pseudo-second-order model	k_2	1.31×10^{-2}
$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \left(\frac{1}{q_e}\right) t$	q_e	194.18
	R^2	0.999
Intraparticle diffusion model	k_p	3.11
$q_t = K_p t^{1/2} + C$	C	178.94
	R^2	0.787

(where, q_t (mg g^{-1}): The biosorption capacity of TF at t time, k_1 (min^{-1}), k_2 ($\text{g mg}^{-1} \text{min}^{-1}$), k_p ($\text{mg g}^{-1} \text{min}^{-1/2}$): The rate constants, C (mg g^{-1}): The intercept.)

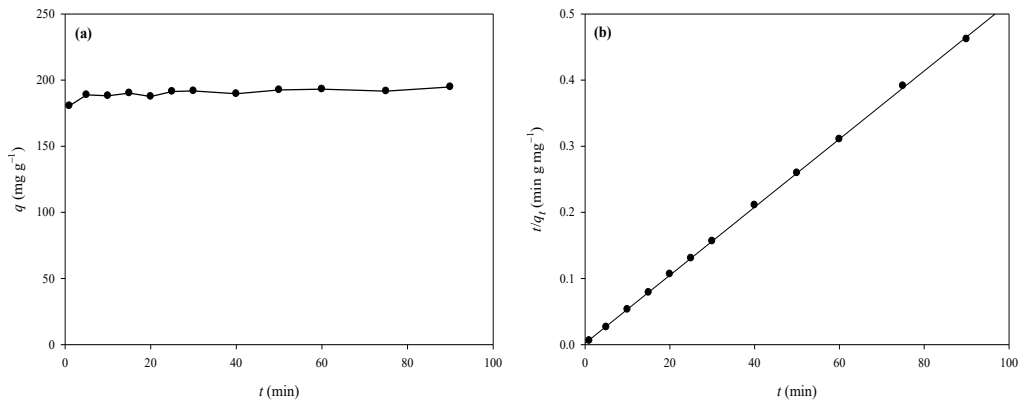


Figure 6: The effect of contact time (a) and the plot of the pseudo-second-order kinetic model for the MV biosorption (b)

3.2.4. Effect of initial MV concentration and biosorption isotherms

Various isotherm models used to interpret experimental biosorption data at equilibrium give significant information about the biosorption mechanism and maximum biosorption capacity of the biosorbent. In this study, the obtained data in equilibrium were applied to Langmuir [33] and Freundlich [34] isotherm models. According to the general isotherm plot (Fig. 7a), the obtained data prove that the biosorption capacity of TF increases from 87.20 to 1149.0 mg g^{-1} with increasing the initial concentration of MV from 50 to 800 mg L^{-1} in the solutions. This can be due to the occupation of the biosorption sites on TF as the concentration of the MV increases. The biosorption process reaches equilibrium when the MV concentration is 700 mg L^{-1} and after this concentration value, the biosorption amount of MV is constant.

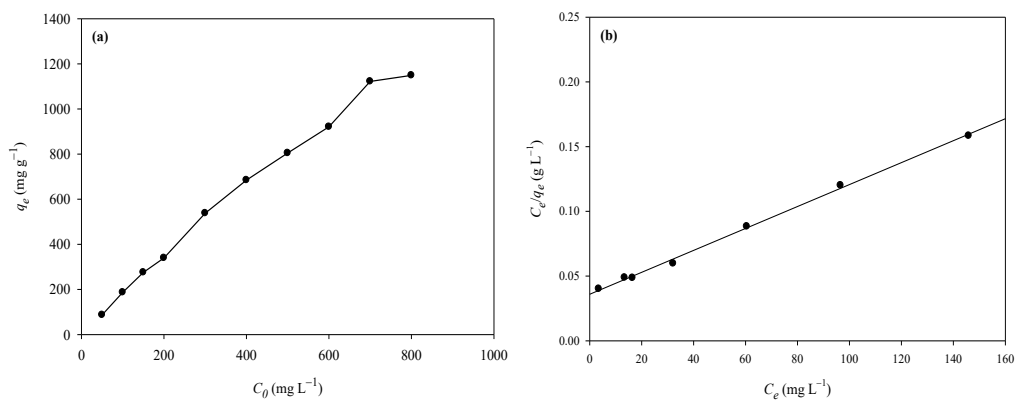


Figure 7: General isotherm plot (a) and Langmuir isotherm model for the MV biosorption (b)

Table 2 shows the Langmuir and Freundlich isotherm parameter values for MV biosorption onto TF. From the results, the data are well fitted to the linearized Langmuir isotherm model than the Freundlich isotherm model because of the highest R^2 value. The values of q_{max} and K_L are identified from the slope and intercept of the linear Langmuir plot of C_e/q_e versus C_0 (Fig 7b). The maximum biosorption capacity of TF is determined as 1180.09 mg g⁻¹ for MV cationic dye from this model. The R_L value of 0.175 demonstrates the favorable interaction between the TF surface and MV molecules. From the results, it can be assumed that the maximum biosorption amount corresponds to saturated monolayered MV molecules on the TF surface with constant energy. The maximum biosorption capacity of TF for MV is quite high or comparable when compared to several biosorbents in the literature (Table 3). TF is considered environmentally friendly, as it not only has an excellent biosorption capacity but also is not treated with any chemicals.

Table 2: Isotherm parameters for the biosorption of MV

Isotherm model	Equations	Parameters	R^2
Langmuir	$\frac{C_e}{q_e} = \frac{1}{q_{max}K_L} + \frac{C_e}{q_{max}}$ $R_L = \frac{1}{1 + K_L C_0}$	$q_{max}=1180.09, K_L=2.36 \times 10^{-2}$ $R_L=0.175$	0.998
Freundlich	$\ln q_e = \ln K_F + \frac{1}{n} \ln C_e$	$n=1.608, K_F=50.60$	0.956

(where, q_{max} : The maximum MV biosorption capacity of TF (mg g⁻¹), K_L and K_F : The Langmuir (L mg⁻¹) and Freundlich (L g⁻¹) constants, respectively, R_L : The separation factor constant and n : Heterogeneity factor.)

Table 3: Comparison of biosorption capacity of TF for MV with other sorbents reported previously

Biosorbent	q_m (mg g ⁻¹)	Reference
Calcined lotus leaf	26.315	[35]
PFPCS	3756.33	[36]
Dithiocarbamate-grafted star-like polymer	1239	[37]
Graphene oxide hydrogel composite	1052.63	[38]
<i>Carya illinoensis</i> powder	642.0	[39]
<i>Anethum graveolens</i>	833.33	[40]
Pu-erh tea powder	277.78	[41]
Breadfruit core	307	[42]
<i>Phragmites australis</i> activated carbon	351.33	[43]
TF	1180.09	Present study

3.2.5. Ionic strength effect

The ionic strength effect on the MV biosorption onto TF was examined under optimum conditions (pH: 9, TF amount: 0.500 g L^{-1} , t: 10 min, initial MV concentration: 100 mg L^{-1} , T: 20°C). For this purpose, NaCl solutions were added to MV solutions in such a way that the final salt concentrations in the solutions were between 0.02 mol L^{-1} and 0.1 mol L^{-1} . As seen in Figure 8, the biosorption capacity of TF results in a reduction of approximately 12% when the salt concentration is 0.02 mol L^{-1} and the biosorption efficiency of MV gradually decreases with increasing NaCl concentration. The biosorption yield of MV is evaluated as 72.20% when the concentration of NaCl reaches 0.08 mol L^{-1} and after this point, the increase in the salt concentration does not influence the biosorption. This can be expressed by a competition between cationic MV molecules and Na^+ ions, which decreases the electrostatic charge on the surface or raises ionic strength [42-43].

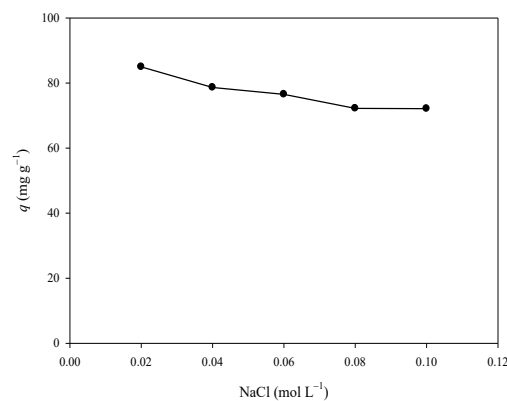


Figure 8: The ionic strength effect on the MV biosorption

4. Conclusion

Kombucha tea fungus (TF) was simply prepared with sugar, black tea leaves and tea containing Kombucha culture. It was used as a biosorbent for methyl violet (MV) biosorption from aqueous solutions. TF was identified with various characterization techniques such as FT-IR, TG analysis and zeta potential measurements. The results demonstrated that pH was an important parameter on the biosorption of MV. The maximum biosorption capacity for TF for the biosorption of MV was achieved at pH=9.0 and the biosorption amount of MV was evaluated as 188.86 mg g^{-1} in a short 10 min contact time with 0.500 g L^{-1} of TF. The observed extremely short equilibrium time is a significant advantage in saving time in the biosorption process. The changes in temperature from 15 to 40°C did not lead to any changes in the MV biosorption. The biosorption process was well defined with the Langmuir isotherm model and the calculated

maximum biosorption capacity value was 1180.09 mg g⁻¹. As a result, TF was superior to many biosorbents because of its high biosorption capacity for MV and also its being a green biosorbent that did not undergo any chemical treatment. All these findings have proven that TF was very effective in dye pollutant removal as well as in other areas of use in industry.

References

- [1] Marsh, A.J., O'Sullivan, O., Hill, C., Ross, R.P., Cotter, P.D., *Sequence-based analysis of the bacterial and fungal compositions of multiple kombucha (tea fungus) samples*, *Food Microbiology*, 38, 171-178, 2014.
- [2] Dutta, H., Paul, S.K., *Kombucha drink: production, quality, and safety aspects*, *Production and Management of Beverages*, 1, 259-288, 2019.
- [3] Taheur, F.B., Mansour, C., Jeddou, K.B., Machreki, Y., Kouidhi, B., Abdulkhakim, J.A., Chaieb, K., *Aflatoxin B1 degradation by microorganisms isolated from Kombucha culture*, *Toxicon*, 179, 76-83, 2020.
- [4] Hesseltine, C.W., *A millennium of fungi, food, and fermentation*, *Mycologia*, 57, 149-197, 1965.
- [5] Blanc, P.J., *Characterization of the tea fungus metabolites*, *Biotechnology Letters*, 18, 139-142, 1996.
- [6] Pasha, C., Reddy, G., *Nutritional and medicinal improvement of black tea by yeast fermentation*, *Food Chemistry*, 89, 449-453, 2005.
- [7] Malbaša, R.V., Lončar, E.S., Vitas, J.S., Čanadanović-Brunet, J.M., *Influence of starter cultures on the antioxidant activity of kombucha beverage*, *Food Chemistry*, 127, 1727-1731, 2011.
- [8] Velićanski, A., Cvetković, D., Markov, S., *Characteristics of Kombucha fermentation on medicinal herbs from Lamiaceae family*, *Romanian Biotechnological Letters*, 18, 8034-8042, 2013.
- [9] Villarreal-Soto, S.A., Beaufort, S., Bouajila, J., Souchard, J.P., Taillandier, P., *Understanding kombucha tea fermentation: a review*, *Journal of Food Science*, 83, 580-588, 2018.
- [10] Amarasekara, A.S., Wang, D., Grady, T.L., *A comparison of kombucha SCOBY bacterial cellulose purification methods*, *SN Applied Sciences*, 2, 1-7, 2020.
- [11] Dima, S.-O., Panaitescu, D.-M., Orban, C., Ghiurea, M., Doncea, S.-M., Fierascu, R.C., Nistor, C.L., Alexandrescu, E., Nicolae, C.-A., Trică, B., *Bacterial nanocellulose from side-streams of kombucha beverages production: Preparation and physical-chemical properties*, *Polymers*, 9, 374, 2017.
- [12] Soares, M.G., de Lima, M., Schmidt, V.C.R., *Technological aspects of kombucha, its applications and the symbiotic culture (SCOBY), and extraction of compounds of interest: A literature review*, *Trends in Food Science & Technology*, 110, 539-550, 2021.
- [13] Srivastava, S., Sinha, R., Roy, D., *Toxicological effects of malachite green*, *Aquatic Toxicology*, 66, 319-329, 2004.
- [14] Al-Fawwaz, A.T., Abdullah, M., *Decolorization of methylene blue and malachite green by immobilized *Desmodesmus* sp. isolated from North Jordan*, *International Journal of Environmental Science and Development*, 7, 95, 2016.

- [15] Bale, M., *Management of the umbilicus with crystal violet solution*, Canadian Medical Association Journal, 124, 372, 1981.
- [16] Mittal, A., Gajbe, V., Mittal, J., *Removal and recovery of hazardous triphenylmethane dye, Methyl Violet through adsorption over granulated waste materials*, Journal of Hazardous Materials, 150, 364-375, 2008.
- [17] Parshetti, G., Saratale, G., Telke, A., Govindwar, S., *Biodegradation of hazardous triphenylmethane dye methyl violet by Rhizobium radiobacter (MTCC 8161)*, Journal of Basic Microbiology, 49, S36-S42, 2009.
- [18] Rott, U., Minke, R., *Overview of wastewater treatment and recycling in the textile processing industry*, Water Science and Technology, 40, 137-144, 1999.
- [19] El Haddad, M., Slimani, R., Mamouni, R., Laamari, M.R., Rafqah, S., Lazar, S., *Evaluation of potential capability of calcined bones on the biosorption removal efficiency of safranin as cationic dye from aqueous solutions*, Journal of the Taiwan Institute of Chemical Engineers, 44, 13-18, 2013.
- [20] Wang, J., Chen, C., *Biosorbents for heavy metals removal and their future*, Biotechnology Advances, 27, 195-226, 2009.
- [21] Mosier, A.P., Behnke, J., Jin, E.T., Cady, N.C., *Microbial biofilms for the removal of Cu²⁺ from CMP wastewater*, Journal of Environmental Management, 160, 67-72, 2015.
- [22] Nguyen, T.A., Fu, C.-C., Juang, R.-S., *Biosorption and biodegradation of a sulfur dye in high-strength dyeing wastewater by Acidithiobacillus thiooxidans*, Journal of Environmental Management, 182, 265-271, 2016.
- [23] Fan, H., Yang, J., Gao, T., Yuan, H., *Removal of a low-molecular basic dye (Azure Blue) from aqueous solutions by a native biomass of a newly isolated Cladosporium sp.: kinetics, equilibrium and biosorption simulation*, Journal of the Taiwan Institute of Chemical Engineers, 43, 386-392, 2012.
- [24] Moghazy, R.M., Labena, A., Husien, S., *Eco-friendly complementary biosorption process of methylene blue using micro-sized dried biosorbents of two macro-algal species (Ulva fasciata and Sargassum dentifolium): Full factorial design, equilibrium, and kinetic studies*, International Journal of Biological Macromolecules, 134, 330-343, 2019.
- [25] Ezeonuegbu, B.A., Machido, D.A., Whong, C.M., Japhet, W.S., Alexiou, A., Elazab, S.T., Qusty, N., Yaro, C.A., Batiha, G.E.-S., *Agricultural waste of sugarcane bagasse as efficient adsorbent for lead and nickel removal from untreated wastewater: Biosorption, equilibrium isotherms, kinetics and desorption studies*, Biotechnology Reports, 30, e00614, 2021.
- [26] Sintakindi, A., Ankamwar, B., *Fungal biosorption as an alternative for the treatment of dyes in waste waters: a review*, Environmental Technology Reviews, 10, 26-43, 2021
- [27] Alp Arici, T., *Highly reusable plant-based biosorbent for the selective methylene blue biosorption from dye mixture in aqueous media*, International Environmental Science and Technology, 1-12, 2021. <https://doi.org/10.1007/s13762-021-03238-w>
- [28] Yang, Y., Wei, X., Sun, P., Wan, J., *Preparation, characterization and adsorption performance of a novel anionic starch microsphere*, Molecules, 15, 2872-2885, 2010.
- [29] Deniz, F., Kepekci, R.A., *Dye biosorption onto pistachio by-product: A green environmental engineering approach*, Journal of Molecular Liquids, 219, 194-200, 2016.
- [30] Lagergren, S., *Zur theorie der sogenannten adsorption gelöster stoffe*, Kungliga Svenska Vetenskapsakademiens. Handlingar, 24, 1-39 1898.

- [31] Ho, Y.-S., McKay, G., *Kinetic models for the sorption of dye from aqueous solution by wood*, Process Safety and Environmental Protection, 76, 183-191, 1998.
- [32] Weber Jr, W.J., Morris, J.C., *Kinetics of adsorption on carbon from solution*, Journal of the Sanitary Engineering Division, 89, 31-59, 1963.
- [33] Langmuir, I., *The adsorption of gases on plane surfaces of glass, mica and platinum*, Journal of the American Chemical society, 40, 1361-1403, 1918.
- [34] Freundlich, H., *Über die adsorption in lösungen*, Zeitschrift für physikalische Chemie, 57, 385-470, 1907.
- [35] Sharafzad, A., Tamjidi, S., Esmaeili, H., *Calcined lotus leaf as a low-cost and highly efficient biosorbent for removal of methyl violet dye from aqueous media*, International Journal of Environmental Analytical Chemistry, 1-24, 2020.
- [36] Chen, K., Du, L., Gao, P., Zheng, J., Liu, Y., Lin, H., *Super and selective adsorption of cationic dyes onto carboxylate-modified passion fruit peel biosorbent*, Frontiers in Chemistry, 9, 376, 2021.
- [37] Liu, Y., Zhao, Y., Cheng, W., Zhang, T., *Targeted reclaiming cationic dyes from dyeing wastewater with a dithiocarbamate-functionalized material through selective adsorption and efficient desorption*, Journal of Colloid and Interface Science, 579, 766-777, 2020.
- [38] Makhado, E., Pandey, S., Ramontja, J., *Microwave assisted synthesis of xanthan gum-cl-poly (acrylic acid) based-reduced graphene oxide hydrogel composite for adsorption of methylene blue and methyl violet from aqueous solution*, International Journal of Biological Macromolecules, 119, 255-269, 2018.
- [39] Yamil, L.d.O., Georgin, J., Franco, D.S., Netto, M.S., Grassi, P., Picilli, D.G., Oliveira, M.L., Dotto, G.L., *Powdered biosorbent from pecan pericarp (Carya illinoensis) as an efficient material to uptake methyl violet 2B from effluents in batch and column operations*, Advanced Powder Technology, 31, 2843-2852, 2020.
- [40] Hamitouche, A., Haffas, M., Boudjemaa, A., Benammar, S., Sehailia, M., Bachari, K., *Efficient biosorption of methylene blue, malachite green and methyl violet organic pollutants on biomass derived from Anethum graveolens: An eco-benign approach for wastewater treatment*, Desalination Water Treatment, 5, 225-236, 2017.
- [41] Li, P., Su, Y.-J., Wang, Y., Liu, B., Sun, L.-M., *Bioadsorption of methyl violet from aqueous solution onto Pu-erh tea powder*, Journal of Hazardous Materials, 179, 43-48, 2010.
- [42] Priyantha, N., Lim, L.B., Tennakoon, D., Liaw, E.T., Ing, C.H., Liyandeniya, A.B., *Biosorption of cationic dyes on breadfruit (Artocarpus altilis) peel and core*, Applied Water Science, 8, 1-11, 2018.
- [43] Chen, S., Zhang, J., Zhang, C., Yue, Q., Li, Y., Li, C., *Equilibrium and kinetic studies of methyl orange and methyl violet adsorption on activated carbon derived from Phragmites australis*, Desalination, 252, 149-156, 2010.



Optimization Based Undersampling for Imbalanced Classes

Fatih SAĞLAM^{1,*}, Mervenur SÖZEN², Mehmet Ali CENGİZ³

¹*Ondokuz Mayıs University, Faculty of Science and Literature, Department of Statistics, 55100, Samsun, Türkiye*

fatih.saglam@omu.edu.tr, ORCID: 0000-0002-2084-2008

²*Ondokuz Mayıs University, Faculty of Science and Literature, Department of Statistics, 55100, Samsun, Türkiye*

mervenur.pala1@gmail.com, ORCID: 0000-0001-5603-5382

³*Ondokuz Mayıs University, Faculty of Science and Literature, Department of Statistics, 55100, Samsun, Türkiye*

macengiz@omu.edu.tr, ORCID: 0000-0002-1271-2588

Received: 21.02.2021

Accepted: 25.10.2021

Published: 31.12.2021

Abstract

The classification methods consider the probability of predicting the majority class to be high when the number of class observations is different. To address this problem, there are some methods such as resampling methods in the literature. Undersampling, one of the resampling methods, creates balance by removing data from the majority class. This study aims to compare different optimization methods to determine the most suitable observations to be taken from the majority class while undersampling. Firstly, a simple simulation study was conducted and graphs were used to analyze the discrepancy between the resampled datasets. Then, different classifier models were constructed for different imbalanced data sets. In these models, random undersampling, undersampling with genetic algorithm, undersampling with differential evolution algorithm, undersampling with an artificial bee colony, and under-sampling with particle herd optimization were compared. The results were given rank numbers differing depending on the



classifiers and data sets and a general mean rank was obtained. As a result, when undersampling, artificial bee colony was seen to perform better than other methods of optimization.

Keywords: Imbalanced classes; Classification; Undersampling; Optimization.

Dengesiz Sınıflamada Optimizasyona Dayalı Azörnekleme

Öz

Sınıflama yöntemleri, sınıf gözlemlerinin sayısı farklı olduğunda çoğunluk sınıfını tahmin etme olasılığının yüksek olduğunu düşünür. Bu sorunu gidermek için literatürde yeniden örnekleme yöntemleri gibi bazı yöntemler bulunmaktadır. Yeniden örnekleme yöntemlerinden biri olan azörnekleme, çoğunluk sınıftan verileri silerek denge oluşturur. Bu çalışma, az örnekleme yapılırken çoğunluk sınıftan alınacak en uygun gözlemleri belirlemek için farklı optimizasyon yöntemlerini karşılaştırmayı amaçlamaktadır. İlk olarak, basit bir simülasyon çalışması yapılmış ve yeniden örneklenen veri setleri arasındaki farklılığı analiz etmek için grafikler kullanılmıştır. Daha sonra, farklı dengesiz veri setleri için farklı sınıflayıcı modelleri oluşturulmuştur. Bu modellerde rastgele azörnekleme, genetik algoritma ile azörnekleme, diferansiyel evrim algoritması ile azörnekleme, yapay arı kolonisi ile azörnekleme ve parçacık sürüsü optimizasyonu ile azörnekleme karşılaştırılmıştır. Sonuçlara sınıflandırıcılara ve veri setlerine göre değişen sıra numaraları verilmiş ve genel bir ortalama sıra elde edilmiştir. Sonuç olarak, yetersiz örnekleme yapıldığında, yapay arı kolonisinin diğer optimizasyon yöntemlerinden daha iyi performans gösterdiği görülmüştür.

Anahtar Kelimeler: Dengesiz sınıflar; Sınıflama; Azörnekleme; Optimizasyon.

1. Introduction

Class imbalance is one of the major problems in machine learning. Class imbalance leads to bias in the learning process from the data set. This bias causes incorrect predictions and makes it difficult to evaluate the model. Resampling is one of the most frequently used methods for dealing with this problem. In this method, balance is achieved by increasing the number of data of minority class observations and/or by reducing the number of data of majority class observations.

Undersampling methods are the methods that provide balance by reducing the majority class observations. Oversampling methods are the methods that provide balance by increasing the minority class observations. The most basic one of the undersampling methods is random undersampling (RUS) method. It is often used in literature to handle class imbalance problem.

For example, Chen et al. [1] reduced the aliasing artifacts and improved image quality by using a hybrid scheme in the form of RUS-based singular value decomposition and compressed sensing. Liu and Tsoumakas [2] used the RUS to improve the learning method of the Ensemble of Classifier Chains against class imbalance. Noise detection is a popular approach to select samples to be removed. Tomek Link [3] is one of these noise detection methods frequently used in the literature to select samples to be discarded in undersampling applications [4-6]. Edited Nearest Neighbor [7] is another noise detection method that can be used for undersampling. It uses three-nearest neighbor and single-nearest neighbor sequentially to reduce the number of samples. Laurikkala [8] proposed a neighborhood cleaning rule to undersample the majority class. The method uses Edited Nearest Neighbor to remove noisy samples not only in the majority class but also in the minority class.

Removal of observations to achieve class balance causes undesirable loss of information available. To minimize the loss of information, it would be more appropriate to select the observations by non-random methods. Many methods have been suggested in the literature for this purpose [9-13]. The purpose of this study is to use optimization to select observations that are to be removed in undersampling. Optimization is the process of achieving the most appropriate solution for a specific purpose under certain constraints. In other words, optimization is the process of obtaining the best result under given conditions. Different optimization techniques have been developed to solve the problems encountered. There are two types of algorithms for solving optimization problems. One of these is the classical algorithms that can obtain the optimal solution by scanning the entire solution space. The other is the heuristic algorithms which intuitively reach the solution in a short time without scanning the solution's entire space. Genetic Algorithm (GA) [14], Particle Swarm Optimization (PSO) [15], Differential Evolution Algorithm (DE) [16], and Artificial Bee Colony (ABC) [17] are some examples of heuristic algorithms.

Methods of optimization were used in various ways when applying methods of resampling. García and Herrera [18] have proposed eight different methods for selecting the observations to be taken in undersampling with evolutionary algorithms. Inspired by this method, Roshan and Asadi [19] increased the performance of bagging by detecting the best performing bonds with the multipurpose evolutionary algorithm. Yu et al. [20] improved the performance of the Support Vector Machines classifier in imbalanced class data sets by setting an optimized decision limit.

In the introductory section of this study, class imbalance problem and optimization methods are mentioned and related studies in the literature are given. In the second section, the problem of class imbalance is explained. Then, random sampling and sampling methods with optimization are introduced. In the third section, the performance criteria used in the model evaluation while

there is a class imbalance are explained. In the fourth section, models with different data sets and different classifiers are formed. In these models, performance results are obtained with and without resampling. In the fifth section, the averages of all results are given and the significance of the difference between these results is tested. Finally, the results are compared and suggestions are made in the sixth section.

2. Materials and Methods

2.1. Class imbalance problem

Algorithms, where the dependent variable is categorical and where these categories are estimated using certain independent variables are referred to as classification. In classification, the categories included in the dependent variable take the class name. Classification methods were developed to determine the order in the observation data and categorize the sample based on this order [21]. Most classification methods establish a model assuming that the number of observations of these classes are equal. In this case, when the number of observations in the classes varies, the models show bias in favor of the class with many observations. This is called class imbalance problem. In almost all actual data, the number of observations of classes is not equal, i.e., there is a class imbalance. However, the term class imbalance is used for situations where a significant imbalance rate exists. This is called between-classes imbalance and is a condition often encountered [22]. When there is an inter-classes imbalance, the class with fewer observations is called the positive class, and the class with more observations is called the negative class. There are some methods used to solve the problem of class imbalance. These methods include the following:

- In resampling methods, data derivation or data reduction may be applied until the data is balanced [23].
- Different weights can be given to the observations.
- Different weights can be given to the classes.
- By using ensemble algorithms, positive classes can be predicted more easily [24].

Resampling methods are methods in which the data set is modified to make the observation classes in the data set more balanced [25].

The resampling method can be grouped under three titles. These are oversampling, undersampling, and hybrid methods. In oversampling methods, the aim is to increase the balance of the positive class data in a certain way. In undersampling methods, some observations from

the majority class observations are supposed to be balanced either randomly or in a controlled way. In hybrid methods, undersampling and oversampling methods are used in conjunction.

2.2. Random undersampling

The most basic of undersampling methods is the random undersampling method. In this method, a certain amount of data is randomly selected from the negative class and the remainder is removed from the data set. Thus, balance can be ensured. RUS deletes the observation from the original data set. The procedure for RUS is given in Algorithm 1. Random observations from T_{neg} of the entered T training set are selected and removed. However, discarding existing information in this method causes loss of information. This is also undesirable. Consequently, RUS is a simple method of adjusting the balance of the T data set [25].

Algorithm 1: RUS (Exact Balance)

Input	:	Training dataset, T
Function	:	
		1. T_{neg} and $T_{pos} \leftarrow$ datasets belonging to classes
		2. N_{neg} and $N_{pos} \leftarrow$ number of observations in T_{neg} and T_{pos} respectively
		3. $L = N_{neg} - N_{pos}$ (Number of samples required for exact balance)
		4. $U \leftarrow L$ -length series randomly selected within $1, 2, \dots, N_{neg}$ series
		5. $T_{neg}^{un} \leftarrow T_{neg}$ after discarding samples selected in U
Output	:	$T_{un} \leftarrow T_{neg}^{un}$ and T_{pos} datasets together

2.3. Undersampling with optimization

The purpose of this method is to determine the most useful observations to be taken from samples of negative class. For this, it is necessary to determine the data set, which provides the best separation in the T training data set. Algorithm 2 specifies the degree to which the separation is accomplished. This algorithm is the objective function of the optimization method used in Algorithm 3. Vector Z with length N_{neg} to be tested in optimization are the parameters to be optimized. These parameters are set as $0 \leq Z_i \leq 1, i = 1, 2, \dots, N_{neg}$ in optimization method. Z parameters are rounded in the objective function. Thus, all parameters take values of either 0 or 1. This is because the method of optimization to be used may not optimize discrete parameters. Data sets of negative and positive classes are determined and two data sets are notated as T_{neg} and T_{pos} are obtained. Then, negative observations in T_{neg} , which have a value of 1 in the Z vector, are removed and T_{neg}^{un} is obtained. Then T_{neg}^{un} and T_{pos} are combined and the new undersampled data set, T^{un} is obtained. A decision tree ($cp = 0.01$), K , is trained in T_{neg}^{un} . The

performance of K in dataset T^{un} is used to determine the performance of the resampling method. For this, independent variables of T^{un} are used to obtain the predictions of K and area under ROC curve (AUC) is used to measure the performance. The subset T^{un} of T , maximizing the AUC level obtained from the estimation gives us the optimal subset. The performance of T^{un} obtained with this method depends on multiple factors. The optimization method to be used is, obviously, one of the factors. Different optimization methods may find different subsets. Iteration numbers in optimization methods may change the results. The type of the classifier (CART decision tree in this case) and the evaluation criteria (AUC in this case) are also important factors.

Undersampling with the optimization method optimizes too many parameters. Therefore, it takes too much time compared to resampling methods without optimization. Too many parameters need too many combinations and more iterations are required to optimize them. Each iteration also takes longer because the parameter number is high.

Algorithm 2: Objective Function of undersampling with optimization

Input	:	Train data set, T ; Z vector of length N_{neg} with values in range $[0,1]$
Function	:	<ol style="list-style-type: none"> 1. $Z = [Z + 0.5]$ (rounding) 2. T_{neg} and $T_{pos} \leftarrow$ datasets of classes 3. $T_{neg}^{un} \leftarrow T_{neg}$ observations corresponding to value of 1 in the vector Z 4. $T_{un} \leftarrow T_{neg}^{un}$ and T_{pos} datasets together 5. $K \leftarrow$ decision tree model trained in T^{un} 6. $Pr(Y = \text{positive class} X = x) \leftarrow$ positive class probabilities of K
Output	:	$Perf \leftarrow$ AUC calculated using $Pr(Y = \text{positive class} X = x)$

Algorithm 3: Undersampling with optimization

Input	:	Train data set, T
Function	:	<ol style="list-style-type: none"> 1. T_{neg} and $T_{pos} \leftarrow$ datasets of classes 2. N_{neg} and $N_{pos} \leftarrow$ number of observations in T_{neg} and T_{pos} respectively 3. Optimize vector $0 \leq Z_i \leq 1, i = 1, 2, \dots, N_{neg}$ using the objective function in Algorithm 2 3. $Z \leftarrow$ optimized parameters 4. $Z_{opt} = [Z + 0.5]$ (rounding) 5. $T_{neg}^{un} \leftarrow T_{neg}$ observations corresponding to the value of 1 in the vector Z_{opt}

Output : $T_{un} \leftarrow T_{neg}^{un}$ and T_{pos} datasets together

2.4 Performance criteria

In order to determine the classification models developed for a data set with class imbalance problem, the most suitable performance criterion should be selected. Although there are many performance criteria in the literature, F_1 , G-mean and AUC are frequently used in imbalanced data sets instead of misleading criteria such as accuracy or error rate [26, 27, 28, 29, 30]. F_1 and G-mean are calculated using confusion matrix. The confusion matrix is given in Table 1. Some of the performance criteria that were calculated through confusion matrix are presented in Table 2. Since accuracy is biased towards the negative class, imbalanced data sets produce biased results. While sensitivity and precision consider positive class performance to be very high, specificity gives the negative class more importance. G-mean is the geometric mean of sensitivity and specificity. F_1 score is the harmonic mean of sensitivity and precision. Since G-mean and F_1 consider both classes, they give a more realistic and general performance in imbalanced datasets.

Table 1: Confusion matrix

		Prediction	
		Positive	Negative
Truth	Positive	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)

Table 2: Some performance criteria used to evaluate classification models

Criterion	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
True positive rate (Sensitivity)	$\frac{TP}{TP + FN}$
True negative rate (Specificity)	$\frac{TN}{TN + FP}$
Positive prediction value (Precision)	$\frac{TP}{TP + FP}$
G-mean	$\sqrt{\frac{\text{Sensitivity} \times \text{Specificity}}{2}}$
F_1	$\frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{Sensitivity}}}$

AUC is the most commonly used criterion for graphical performance. Calculation of AUC can be given as follows:

$$\overline{AUC} = \int_0^1 \overline{ROC}(t) dt \quad (1)$$

3. Simulation

Figure 1 shows the flow chart of forming the model with or without resampling. The study was conducted on simulation data and then 11 different data sets which are available online [31]. Table 3 shows the structural characteristics of the data sets. Undersampling methods were applied to achieve balance in the train dataset by removing observations from the negative class. Methods of undersampling used are; RUS, undersampling with genetic algorithm (US with GA), undersampling with differential evolution algorithm (US with DE), undersampling with an artificial bee colony (US with ABC), undersampling with particle flock optimization (US with PSO).

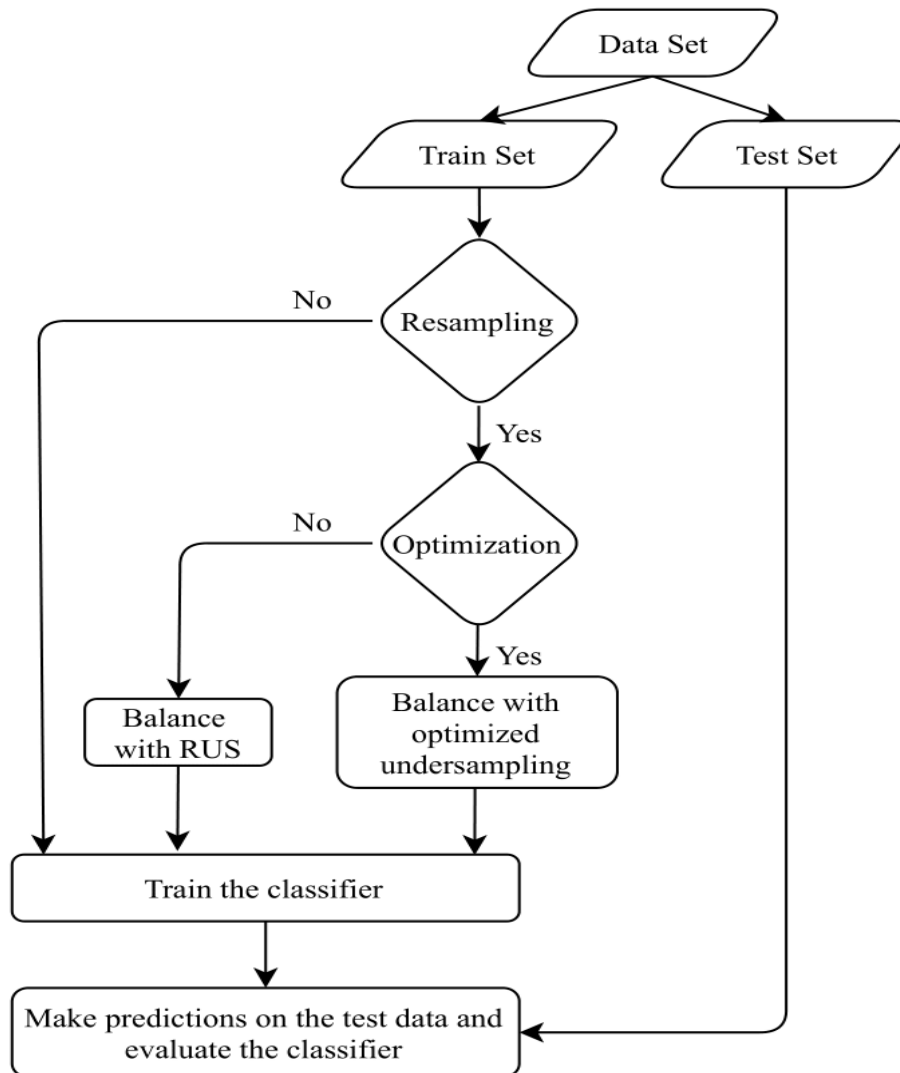


Figure 1: Flow chart of modelling with and without

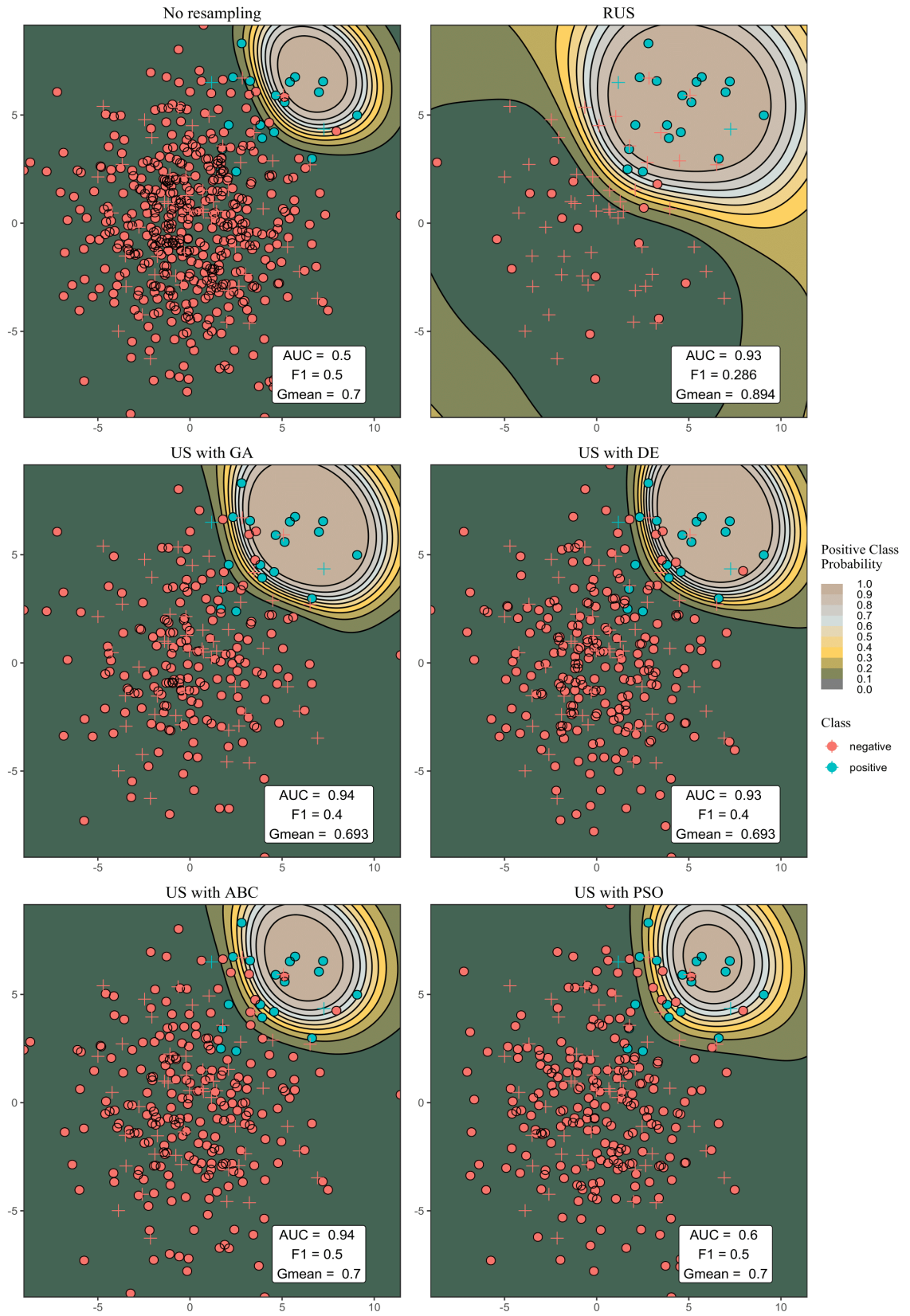


Figure 2: Decision boundaries of radial SVM in the simulation data set (“+” signs indicate test set, circles indicate training set)

The maximum number of iterations was set to 100 in all optimization methods. The class observations were made in the simulation study as two dimensional $X_{neg} \sim N(0,3)$ with length 500 and $X_{pos} \sim N(5,2)$ with length 20. The training and test data set were stratified as 90% and 10% respectively. The radial SVM model was formed on the imbalanced training set. Secondly, the balance was achieved with RUS, and the model was formed. Subsequently, four different optimization methods were used and undersamplings were applied. Figure 2 shows the decision boundaries and performances of the models. Although according to the original data set, there was no increase in F_1 score, US with GA and US with ABC reached 0.94 AUC. Compared to other methods, US with PSO was more inefficient in the simulation study. In the datasets undersampled by optimization methods, it is seen that less data is discarded from the majority class compared to the RUS method. If we consider that the US with GA, US with DE, and US with ABC methods perform close to or better than RUS, we see that we can achieve the desired results with less information loss. But the same is difficult to say for US with PSO.

Table 3: Structural properties of data sets

Dataset	Independent Variable Number	N	N_{pos}	N_{neg}	Imbalance Ratio (N_{neg}/N_{pos})
breast	9	277	81	196	2.420
bupa	5	345	145	200	1.379
cleveland	5	296	34	262	7.706
ecoli	5	335	34	301	8.853
haberman	3	305	81	224	2.765
Sonar	60	120	23	97	4.217
glass0	9	214	70	144	2.057
kyphosis	3	81	17	64	3.765
newthyroid1	5	215	35	180	5.143
Seeds	7	199	66	133	2.015
Vertebral	6	310	100	210	2.100

Eleven different data sets and 5 different classifiers were used to analyze the performances of these methods. AUC, F_1 and G-mean criteria were used to evaluate the performance of the models. The 10-fold cross-validity was repeated 10 times and the performance criteria were averaged. In addition, the performance of the methods was graded from 1 to 6 with $10 \times 10 = 100$ models set up for each case. The sequence numbers were properly synchronized when the same performances were achieved in this ranking. To give an example, all $(1 + 2 + 3) / 3 = 2$ were assigned rank numbers when the three methods were in the first place, and the subsequent methods were assigned 4-5-6 rank numbers. The obtained rank numbers were averaged for 100 models and added to the results.

In the study, J48 decision tree ($C = 0.25$, $M = 2$), K nearest neighborhood (KNN, $K = 5$), radial support vector machines (SVM, $\sigma = 1$, $C = 0.25$), naive bayes and logistic regression classifiers were used.

The AUC measurements of the J48 classifier are given in Table 4 and the average rank in Table 5. In the 11 data sets, the best results were obtained once without resampling, four times in RUS, three times in US with GA, one time in US with DE, and two times in US with ABC. The best ranking was achieved with US with GA, and the worst with US with PSO when the rankings were analyzed. While GA was the optimization method that gave the best result in AUC rankings, the method that gave the worst result was PSO.

Table 4: AUC values for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.6168±0.11	0.6069±0.1225	0.6047±0.1292	0.629±0.1082	0.6192±0.1145	0.6262±0.1214
bupa	0.6168±0.09	0.6285±0.1023	0.6126±0.0972	0.614±0.0981	0.6123±0.0962	0.617±0.0932
cleveland	0.5209±0.0575	0.5878±0.1374	0.5288±0.1267	0.5455±0.1177	0.5379±0.1145	0.525±0.1075
ecoli	0.7921±0.1858	0.7988±0.1271	0.7949±0.1511	0.7888±0.1471	0.803±0.1659	0.7587±0.1748
haberman	0.5727±0.0953	0.6269±0.0979	0.6137±0.1087	0.6147±0.1144	0.6193±0.1008	0.58±0.0995
Sonar	0.6982±0.1976	0.7125±0.1931	0.7627±0.1898	0.7598±0.1786	0.7164±0.2081	0.7492±0.1917
glass0	0.7939±0.0975	0.7874±0.0912	0.8049±0.1035	0.7973±0.089	0.7868±0.1028	0.7999±0.1031
kyphosis	0.5905±0.1948	0.7181±0.1765	0.6371±0.2016	0.6015±0.1819	0.6461±0.193	0.5886±0.1787
newthyroid1	0.9501±0.0899	0.937±0.0734	0.9569±0.0784	0.95±0.0808	0.94±0.0912	0.9547±0.0735
Seeds	0.918±0.0833	0.8965±0.0821	0.9024±0.079	0.9085±0.0709	0.9125±0.0811	0.9165±0.0762
Vertebral	0.8384±0.0862	0.8385±0.0755	0.8399±0.0814	0.8389±0.0792	0.8474±0.0863	0.8261±0.0924

Table 5: AUC mean ranks for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	3.68	3.565	3.75	3.2	3.445	3.36
bupa	3.56	3.24	3.615	3.525	3.545	3.515
cleveland	3.715	2.785	3.65	3.39	3.685	3.775
ecoli	3.245	3.58	3.51	3.66	3.35	3.655
haberman	3.855	2.98	3.215	3.42	3.31	4.22
Sonar	3.655	3.97	3.3	3.295	3.535	3.245
glass0	3.505	3.685	3.275	3.57	3.715	3.25
kyphosis	3.965	2.67	3.425	3.72	3.305	3.915
newthyroid1	3.02	4.205	3.29	3.535	3.535	3.415
Seeds	3.18	3.935	3.57	3.645	3.415	3.255
Vertebral	3.385	3.68	3.47	3.505	3.195	3.765
Mean	3.524	3.481	3.461	3.497	3.458	3.579

The F_1 measurements of the J48 classifier are given in Table 6 and the average rank in Table 7. In the 11 datasets, the best results were obtained ten times without resampling and one time in US with GA. In terms of rank averages, the best ranking was reached when no resampling was performed, while the worst ranking was achieved when RUS was applied. While the optimization method that gave the best result in F_1 rankings was ABC, the method that gave the worst result was GA.

Table 6: F_1 values for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.8206±0.046	0.7038±0.0929	0.7608±0.0756	0.7708±0.0749	0.755±0.084	0.7637±0.0753
bupa	0.7063±0.0706	0.6644±0.107	0.5852±0.1437	0.5763±0.1495	0.6267±0.1624	0.6363±0.1305
cleveland	0.9337±0.0162	0.7415±0.1415	0.9102±0.0424	0.9255±0.0258	0.9137±0.037	0.9173±0.0362
ecoli	0.9614±0.0196	0.8683±0.0629	0.954±0.0271	0.9549±0.0251	0.9533±0.027	0.9563±0.028
haberman	0.8084±0.0537	0.7604±0.075	0.7859±0.0684	0.7791±0.0895	0.7821±0.0734	0.7803±0.0716
Sonar	0.5043±0.2811	0.4947±0.2217	0.5827±0.2566	0.5709±0.2669	0.5506±0.2752	0.5758±0.2785
glass0	0.848±0.0655	0.7997±0.0696	0.8285±0.0821	0.8136±0.0798	0.8278±0.0729	0.8218±0.0814
kyphosis	0.8298±0.0831	0.7222±0.1746	0.7931±0.1185	0.7841±0.1371	0.7891±0.1377	0.7907±0.1235
newthyroid1	0.9837±0.0219	0.9601±0.0354	0.9787±0.0249	0.976±0.0244	0.9753±0.0276	0.9755±0.026
Seeds	0.89±0.0949	0.8404±0.1068	0.8523±0.099	0.8563±0.091	0.8784±0.0931	0.8662±0.0997
Vertebral	0.8668±0.0547	0.8377±0.0659	0.8496±0.0561	0.8429±0.0588	0.8637±0.0551	0.8416±0.0605

Table 7: F_1 mean ranks for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	1.945	4.82	3.66	3.325	3.615	3.635
bupa	2.48	3.07	4.21	4.325	3.38	3.535
cleveland	2.415	5.74	3.44	2.915	3.315	3.175
ecoli	2.625	5.835	3.16	3.215	3.245	2.92
haberman	2.555	4.22	3.45	3.57	3.525	3.68
Sonar	3.765	4.145	3.305	3.24	3.435	3.11
glass0	2.735	4.2	3.39	3.9	3.34	3.435
kyphosis	2.805	4.28	3.615	3.585	3.35	3.365
newthyroid1	3.025	4.345	3.305	3.495	3.385	3.445
Seeds	2.93	3.965	3.79	3.775	3.175	3.365
Vertebral	2.885	3.82	3.56	3.885	2.92	3.93
Mean	2.742	4.404	3.535	3.566	3.335	3.418

The G-mean measurements of the J48 classifier are given in Table 8 and the average rank in Table 9. In the 11 datasets, the best results were obtained two times without resampling, six times in RUS, two times in US with GA, and one time in US with DE in 11 data sets. When examining the averages of the rankings, it was seen that the best ranking was achieved when applying RUS, and the worst ranking was achieved when no resampling was applied. While GA was the optimization method that gave the best result in G-mean rankings, the method that gave the worst result was PSO.

Table 8: G-mean values for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.51±0.1689	0.5718±0.1185	0.5602±0.1376	0.5841±0.1292	0.5822±0.1175	0.5698±0.1514
bupa	0.5841±0.0857	0.5897±0.1153	0.5581±0.119	0.555±0.1207	0.5619±0.1108	0.569±0.1021
cleveland	0.0203±0.1001	0.5257±0.2173	0.1597±0.2633	0.1614±0.2603	0.1577±0.2675	0.1378±0.2534
ecoli	0.704±0.233	0.7902±0.1074	0.7423±0.1893	0.738±0.2026	0.7228±0.2308	0.7113±0.2235
haberman	0.4054±0.2724	0.6052±0.1263	0.5796±0.1355	0.5679±0.153	0.5795±0.1466	0.5355±0.1505
Sonar	0.6133±0.3036	0.6766±0.236	0.7191±0.2513	0.6998±0.2701	0.6691±0.2825	0.6947±0.2794
glass0	0.7812±0.0946	0.7706±0.0801	0.7916±0.1016	0.7765±0.1209	0.7741±0.0982	0.774±0.0999
kyphosis	0.3025±0.3502	0.6312±0.2596	0.4504±0.3615	0.3735±0.3632	0.4316±0.3576	0.3402±0.362
newthyroid1	0.9434±0.1019	0.9314±0.0818	0.9419±0.0994	0.9314±0.1059	0.9239±0.113	0.9338±0.0981
Seeds	0.9114±0.0753	0.8821±0.0835	0.8863±0.0783	0.8912±0.0709	0.9046±0.0735	0.8974±0.0789
Vertebral	0.7471±0.0987	0.8054±0.076	0.7801±0.0818	0.7854±0.0837	0.7559±0.0982	0.7657±0.0903

Table 9: G-mean rank means for J48 classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	4.22	3.47	3.67	3.12	3.275	3.245
bupa	3.345	3.125	3.625	3.83	3.53	3.545
cleveland	4.355	1.745	3.71	3.645	3.715	3.83
ecoli	3.485	3.4	3.395	3.53	3.55	3.64
haberman	4.04	2.84	3.295	3.545	3.245	4.035
Sonar	3.83	4.01	3.295	3.23	3.49	3.145
glass0	3.43	3.84	3.245	3.5	3.43	3.555
kyphosis	4.095	2.655	3.275	3.635	3.455	3.885
newthyroid1	2.965	4.035	3.295	3.575	3.605	3.525
Seeds	2.95	3.97	3.795	3.77	3.17	3.345
Vertebral	4.21	2.755	3.235	3.255	3.915	3.63
Mean	3.72	3.259	3.44	3.512	3.489	3.58

The AUC measurements of the KNN classifier are given in Table 10 and the average ranks are given in Table 11. In the 11 datasets, the best results were obtained five times without resampling, four times with RUS, and two times in US with GA. When examining the averages of the rankings, it was seen that the best ranking was achieved without resampling and the worst ranking was achieved when RUS was applied. While the optimization method which gave the best result in AUC rankings was ABC, the method that gave the worst result was PSO.

Table 10: AUC values for KNN classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.6473±0.1052	0.6617±0.1104	0.6553±0.101	0.6584±0.0966	0.6573±0.0943	0.6611±0.1091
bupa	0.6614±0.0842	0.6648±0.0862	0.6518±0.0853	0.6613±0.085	0.6611±0.0874	0.6573±0.0785
cleveland	0.5378±0.1342	0.5581±0.1596	0.5418±0.1415	0.551±0.1387	0.5492±0.1483	0.5322±0.1363
ecoli	0.8913±0.0988	0.9134±0.0595	0.9092±0.0768	0.9156±0.0737	0.91±0.0804	0.9099±0.0764
haberman	0.6531±0.0961	0.6529±0.1051	0.6443±0.1013	0.651±0.1044	0.648±0.1051	0.6456±0.105
Sonar	0.917±0.1157	0.8745±0.1331	0.9036±0.127	0.8969±0.1315	0.9046±0.1217	0.9023±0.1257
glass0	0.8749±0.0823	0.8528±0.0779	0.8541±0.0834	0.8526±0.0819	0.864±0.0857	0.8454±0.0867
kyphosis	0.6472±0.2394	0.6596±0.2136	0.6611±0.237	0.647±0.2407	0.6448±0.2327	0.6402±0.2424
newthyroid1	0.9919±0.0409	0.9793±0.0459	0.9878±0.0405	0.9878±0.0288	0.9869±0.0414	0.987±0.0412
Seeds	0.9693±0.0391	0.9725±0.0336	0.9734±0.0337	0.9731±0.0319	0.969±0.0395	0.97±0.0354
Vertebral	0.9092±0.0525	0.9036±0.0477	0.9044±0.0548	0.9074±0.0495	0.9086±0.0541	0.907±0.0505

Table 11: AUC rank means for KNN classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	3.88	3.29	3.505	3.435	3.61	3.28
bupa	3.36	3.4	3.775	3.415	3.535	3.515
cleveland	3.52	3.325	3.53	3.47	3.335	3.82
ecoli	3.555	3.705	3.435	3.23	3.44	3.635
haberman	3.31	3.42	3.785	3.42	3.49	3.575
Sonar	2.745	4.41	3.385	3.635	3.31	3.515
glass0	2.455	3.81	3.75	3.895	3.03	4.06
kyphosis	3.49	3.495	3.325	3.495	3.575	3.62
newthyroid1	2.69	4.28	3.48	3.61	3.49	3.45
Seeds	3.53	3.545	3.435	3.29	3.465	3.735
Vertebral	3.155	3.855	3.62	3.63	3.21	3.53
Mean	3.245	3.685	3.548	3.502	3.408	3.612

The F_1 measurements of the KNN classifier are given in Table 12, and the average ranks are given in Table 13. In the 11 datasets, the best results were obtained ten times without resampling and one time in RUS. When examining the average rankings, it was seen that the best ranking was achieved without resampling and the worst ranking was achieved when RUS was applied. While the optimization method which gave the best result in F_1 rankings was ABC, the method that gave the worst result was DE.

Table 12: F_1 values for KNN classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.8164±0.0446	0.7128±0.0781	0.7607±0.0682	0.7591±0.072	0.7488±0.0757	0.7597±0.0773
bupa	0.7151±0.064	0.6658±0.0834	0.6084±0.1053	0.5954±0.1084	0.6768±0.1044	0.6336±0.1002
cleveland	0.9341±0.0134	0.7179±0.0885	0.9036±0.0336	0.9041±0.0334	0.9033±0.0306	0.9092±0.0286
ecoli	0.9561±0.0236	0.8896±0.0506	0.9525±0.0244	0.9527±0.0277	0.9528±0.0261	0.953±0.027
haberman	0.8228±0.0467	0.7198±0.0718	0.7768±0.0571	0.7694±0.0669	0.7742±0.0657	0.771±0.0621
Sonar	0.7201±0.2115	0.5336±0.1536	0.6404±0.2064	0.6122±0.1911	0.6593±0.2196	0.62±0.1903
glass0	0.8368±0.0693	0.771±0.0898	0.7869±0.1011	0.7832±0.0854	0.8124±0.0847	0.7744±0.0918
kyphosis	0.8549±0.0627	0.6627±0.1778	0.8271±0.0997	0.8094±0.1187	0.8239±0.0966	0.806±0.1107
newthyroid1	0.9776±0.0213	0.961±0.0343	0.9757±0.0246	0.9758±0.0261	0.9752±0.026	0.9743±0.0263
Seeds	0.8397±0.1078	0.8703±0.0954	0.8586±0.0993	0.8619±0.1029	0.8522±0.0981	0.8566±0.0992
Vertebral	0.8795±0.0507	0.8464±0.056	0.8551±0.0599	0.8589±0.0573	0.8741±0.0543	0.8657±0.0562

Table 13: F_1 rank means for KNN classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	1.815	4.795	3.565	3.545	3.79	3.49
bupa	1.91	3.205	4.425	4.78	2.73	3.95
cleveland	1.6	6	3.465	3.425	3.405	3.105
ecoli	2.625	5.86	3.275	3.06	3.105	3.075
haberman	1.625	5.4	3.345	3.62	3.445	3.565
Sonar	2.14	5.055	3.27	3.775	3.09	3.67
glass0	2.13	4.355	3.68	3.855	2.895	4.085
kyphosis	2.47	5.285	3.14	3.315	3.23	3.56
newthyroid1	3.155	4.355	3.36	3.315	3.4	3.415
Seeds	3.855	3.18	3.47	3.42	3.55	3.525
Vertebral	2.665	4.405	3.925	3.625	2.995	3.385
Mean	2.363	4.718	3.538	3.612	3.24	3.53

The G-mean measurements of the KNN classifier are given in Table 14, and the average ranks are given in Table 15. In the 11 data sets, the best results were achieved once without resampling, seven times in RUS, and two times in US with DE. When the rank averages are examined, it was seen that the best ranking is reached when RUS is applied and the worst ranking is reached when no resampling is applied. While the optimization method which gave the best result in G-mean rankings was DE, the method that gave the worst result was PSO.

Table 14: G-mean values for KNN classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.491±0.1433	0.6063±0.1029	0.5965±0.121	0.6024±0.1235	0.5889±0.1112	0.5852±0.1324
bupa	0.6066±0.0887	0.62±0.0858	0.6017±0.0882	0.5993±0.0876	0.6065±0.0915	0.6116±0.0845
cleveland	0±0	0.4971±0.2218	0.0627±0.172	0.0399±0.1362	0.0227±0.115	0.036±0.1321

ecoli	0.6932±0.2592	0.8305±0.1195	0.7691±0.2262	0.7959±0.1877	0.7573±0.2195	0.7694±0.2189
haberman	0.4515±0.1882	0.5998±0.1053	0.5593±0.1238	0.5744±0.1242	0.5837±0.123	0.5486±0.1303
Sonar	0.8406±0.1701	0.7461±0.1368	0.8062±0.1682	0.7983±0.1487	0.8127±0.1711	0.8013±0.165
glass0	0.7489±0.1184	0.7562±0.0934	0.7609±0.1004	0.7613±0.0898	0.7484±0.1086	0.7436±0.0973
kyphosis	0.0932±0.2581	0.538±0.3016	0.2994±0.3863	0.3243±0.3748	0.2402±0.3525	0.2942±0.357
newthyroid1	0.8628±0.1572	0.8996±0.1166	0.8929±0.1299	0.8905±0.1346	0.8896±0.1277	0.8888±0.1323
Seeds	0.8704±0.0871	0.9037±0.0733	0.8915±0.0774	0.8963±0.0809	0.884±0.0806	0.8901±0.0772
Vertebral	0.8097±0.082	0.8301±0.0622	0.8246±0.0711	0.8324±0.0637	0.814±0.0776	0.8286±0.0722

Table 15: G-mean rank means for KNN classifier

Dataset	No resampling	RUS	GA	DE	ABC	PSO
breast	4.765	3.25	3.135	3.15	3.31	3.39
bupa	3.5	3.115	3.605	3.745	3.55	3.485
cleveland	4.095	1.38	3.78	3.875	3.975	3.895
ecoli	3.905	3.205	3.49	3.275	3.65	3.475
haberman	4.91	2.84	3.435	3.185	2.955	3.675
Sonar	2.17	4.995	3.28	3.775	3.105	3.675
glass0	3.34	3.565	3.325	3.365	3.54	3.865
kyphosis	4.315	2.585	3.395	3.35	3.785	3.57
newthyroid1	3.425	4.015	3.38	3.345	3.4	3.435
Seeds	3.885	3.17	3.46	3.42	3.55	3.515
Vertebral	3.76	3.38	3.59	3.275	3.665	3.33
Mean	3.825	3.227	3.443	3.433	3.499	3.574

The AUC measurements of the radial SVM classifier are given in Table 16, and the average ranks are given in Table 17. In the 11 datasets, the best results were obtained five times without resampling, four times in RUS, and two times in US with GA. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was achieved when RUS was applied. While the optimization method which gave the best result in AUC rankings was ABC, the method that gave the worst result was PSO.

Table 16: AUC values for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.6454±0.1174	0.5424±0.1658	0.6282±0.1137	0.6305±0.1238	0.6092±0.1343	0.6226±0.1318
bupa	0.7124±0.0737	0.6899±0.0824	0.6839±0.081	0.6822±0.0833	0.7022±0.0734	0.6884±0.0779
cleveland	0.6946±0.1735	0.5804±0.2161	0.6666±0.1737	0.6711±0.168	0.6599±0.1717	0.6379±0.1879
ecoli	0.8878±0.1397	0.9284±0.0562	0.9062±0.11	0.9088±0.1041	0.8936±0.1273	0.903±0.1093
haberman	0.6543±0.1122	0.6688±0.1087	0.6884±0.1157	0.6928±0.1084	0.6965±0.1052	0.6904±0.1055
Sonar	0.6496±0.3671	0.4933±0.2564	0.3483±0.3553	0.4027±0.3668	0.481±0.3894	0.3379±0.34
glass0	0.8801±0.0767	0.8676±0.0796	0.8641±0.0795	0.8674±0.0809	0.8743±0.0791	0.8716±0.0772
kyphosis	0.8211±0.1889	0.7748±0.2553	0.8464±0.1699	0.8277±0.1926	0.8426±0.1741	0.7962±0.2171
newthyroid1	0.9945±0.0127	0.9955±0.0141	0.9956±0.0156	0.9966±0.0114	0.9971±0.012	0.996±0.0123
Seeds	0.9852±0.0206	0.9828±0.0231	0.9844±0.0214	0.9813±0.0279	0.9838±0.0218	0.9842±0.0227
Vertebral	0.8843±0.0664	0.8831±0.0585	0.8841±0.0645	0.8894±0.0602	0.8827±0.0646	0.8813±0.063

Table 17: AUC rank means for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	2.99	4.085	3.57	3.315	3.67	3.37
bupa	2.495	3.755	3.93	3.9	3.02	3.9
cleveland	2.685	3.855	3.605	3.485	3.655	3.715
ecoli	3.155	3.99	3.4	3.345	3.565	3.545

haberman	4.295	4.105	3.175	3.215	2.92	3.29
Sonar	2.645	3.11	3.99	3.85	3.535	3.87
glass0	2.92	3.64	4.01	3.625	3.165	3.64
kyphosis	3.41	3.91	3.125	3.505	3.275	3.775
newthyroid1	3.755	3.51	3.445	3.44	3.39	3.46
Seeds	3.285	3.675	3.435	3.73	3.505	3.37
Vertebral	3.085	4.015	3.515	3.135	3.385	3.865
Mean	3.156	3.786	3.564	3.504	3.371	3.618

The F_1 measurements of the radial SVM classifier are given in Table 18 and the average ranks are given in Table 19. In the 11 datasets, the best results were obtained nine times without resampling and two times in RUS. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was achieved when RUS was applied. While the optimization method which gave the best result in F_1 rankings was ABC, the method that gave the worst result was DE.

Table 18: F_1 values for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.8194±0.0214	0.5094±0.1433	0.7538±0.127	0.7804±0.0703	0.7335±0.1058	0.7759±0.0791
bupa	0.753±0.0551	0.7251±0.06	0.5706±0.156	0.5056±0.1405	0.6801±0.1355	0.6579±0.1484
cleveland	0.9352±0.0138	0.597±0.1404	0.9313±0.0197	0.9314±0.0166	0.9321±0.0157	0.9299±0.0193
ecoli	0.9552±0.0211	0.9344±0.0313	0.9548±0.0242	0.9529±0.0251	0.9536±0.0228	0.9548±0.0234
haberman	0.84±0.0313	0.7209±0.0815	0.7989±0.0522	0.7946±0.06	0.7956±0.062	0.7996±0.0599
Sonar	0±0	0.2256±0.1296	0±0	0±0	0±0	0±0
glass0	0.865±0.0446	0.8364±0.0753	0.8303±0.0676	0.8327±0.066	0.8588±0.0546	0.835±0.07
kyphosis	0.8749±0.0802	0.7525±0.1731	0.8602±0.1082	0.858±0.1002	0.8582±0.1009	0.8478±0.0995
newthyroid1	0.9779±0.0259	0.9202±0.0464	0.9369±0.0461	0.9425±0.041	0.9445±0.0451	0.9426±0.0442
Seeds	0.8808±0.0942	0.8906±0.0961	0.882±0.091	0.8792±0.0918	0.877±0.0921	0.8854±0.0866
Vertebral	0.8796±0.0401	0.8493±0.0577	0.8592±0.0498	0.8614±0.0488	0.8748±0.0466	0.8643±0.0479

Table 19: F_1 rank means for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	1.97	5.79	3.38	3.03	3.715	3.115
bupa	1.95	2.86	4.49	5.225	2.98	3.495
cleveland	2.68	6	3.025	3.145	3.04	3.11
ecoli	2.915	5.03	3.155	3.455	3.325	3.12
haberman	1.86	5.44	3.43	3.48	3.46	3.33
Sonar	3.88	1.6	3.88	3.88	3.88	3.88
glass0	2.59	3.755	4.14	3.855	2.735	3.925
kyphosis	2.925	4.71	3.13	3.285	3.305	3.645
newthyroid1	1.87	4.645	3.775	3.605	3.485	3.62
Seeds	3.55	3.29	3.49	3.6	3.65	3.42
Vertebral	2.56	4.425	3.925	3.72	2.815	3.555
Mean	2.614	4.322	3.62	3.662	3.308	3.474

The G-mean measurements of the radial SVM classifier are given in Table 20, and the average ranks are given in Table 21. The best results were obtained two times without resampling and nine times in the RUS. When the rank averages were examined, it was seen that the best ranking was achieved when RUS was applied and the worst ranking was achieved when no

resampling was applied. While the optimization method which gave the best result in G-mean rankings was ABC, the method that gave the worst result was PSO.

Table 20: G-mean values for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.1087±0.1739	0.4599±0.1615	0.3307±0.2134	0.3666±0.2347	0.3846±0.2455	0.3167±0.2392
bupa	0.6257±0.0824	0.6185±0.0792	0.5701±0.1112	0.5489±0.1041	0.6068±0.0874	0.5926±0.1067
cleveland	0.0049±0.049	0.473±0.2252	0.0627±0.1712	0.0531±0.1613	0.0444±0.1523	0.0317±0.1265
ecoli	0.549±0.2918	0.7246±0.2439	0.6533±0.2643	0.6579±0.2568	0.6091±0.2776	0.6315±0.2772
haberman	0.2752±0.2021	0.6192±0.1089	0.5344±0.1509	0.5702±0.1236	0.5697±0.1311	0.5288±0.1673
Sonar	0±0	0.0294±0.1017	0±0	0±0	0±0	0±0
glass0	0.7377±0.1137	0.7699±0.1084	0.7463±0.1042	0.7568±0.0974	0.7544±0.111	0.7468±0.1112
kyphosis	0.3236±0.3847	0.6165±0.3033	0.4577±0.4007	0.4429±0.3898	0.3801±0.3967	0.3884±0.3827
newthyroid1	0.9718±0.0435	0.924±0.0429	0.9368±0.051	0.9418±0.0456	0.9413±0.0527	0.9424±0.0453
Seeds	0.9027±0.0779	0.9163±0.0772	0.9066±0.0742	0.9025±0.0757	0.9013±0.0759	0.9072±0.0726
Vertebral	0.7749±0.0852	0.7962±0.0781	0.7874±0.0743	0.7941±0.0781	0.7844±0.0803	0.7881±0.075

Table 21: G-mean rank means for radial SVM classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	5.085	2.545	3.555	3.165	3.06	3.59
bupa	2.82	3.19	3.865	4.345	3.26	3.52
cleveland	4.095	1.505	3.785	3.8	3.865	3.95
ecoli	3.85	3.385	3.34	3.46	3.62	3.345
haberman	5.6	2.36	3.455	3.03	3	3.555
Sonar	3.54	3.3	3.54	3.54	3.54	3.54
glass0	3.67	3.125	3.78	3.395	3.36	3.67
kyphosis	3.86	2.865	3.37	3.51	3.645	3.75
newthyroid1	2	4.555	3.745	3.585	3.515	3.6
Seeds	3.585	3.17	3.5	3.62	3.675	3.45
Vertebral	3.755	3.185	3.57	3.275	3.555	3.66
Mean	3.805	3.017	3.591	3.52	3.463	3.603

The AUC measurements of the Naive Bayes classifier are given in Table 22, and the average rank in Table 23. In the 11 datasets, the best results were obtained seven times without resampling, two times with US with GA, and two times with US with PSO. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was reached when RUS was applied. While the optimization method which gave the best results in AUC rankings was ABC, the method that gave the worst results was DE.

Table 22: AUC values for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.72±0.0998	0.7156±0.1009	0.7253±0.1017	0.7215±0.098	0.7235±0.0992	0.7239±0.0962
bupa	0.6041±0.0917	0.5989±0.0938	0.6019±0.0929	0.5898±0.089	0.6032±0.0897	0.6025±0.0889
cleveland	0.6386±0.1591	0.6314±0.1702	0.631±0.1591	0.6321±0.1621	0.633±0.1553	0.6206±0.1584
ecoli	0.92±0.0702	0.9104±0.0808	0.9184±0.0719	0.9199±0.0709	0.9188±0.0708	0.9168±0.0714
haberman	0.6865±0.1091	0.6782±0.1193	0.6816±0.11	0.6796±0.1132	0.6832±0.109	0.6856±0.112
Sonar	0.8886±0.1289	0.8916±0.112	0.8919±0.1248	0.8891±0.1347	0.8896±0.1241	0.8895±0.1173
glass0	0.8926±0.0681	0.8664±0.072	0.8698±0.0827	0.8722±0.0656	0.8844±0.0739	0.8635±0.081
kyphosis	0.7975±0.1888	0.7869±0.1861	0.8001±0.1917	0.777±0.1946	0.8011±0.1902	0.8086±0.1768
newthyroid1	0.9674±0.0399	0.9483±0.0531	0.9519±0.0476	0.9503±0.0514	0.9561±0.043	0.9575±0.043
Seeds	0.9646±0.0385	0.9615±0.04	0.9631±0.0396	0.9628±0.0391	0.9637±0.0393	0.9629±0.0385
Vertebral	0.8397±0.0748	0.8383±0.0772	0.8374±0.0758	0.8406±0.0756	0.8388±0.074	0.8439±0.0725

Table 23: AUC rank means for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	3.77	4.035	3.05	3.515	3.355	3.275
bupa	3.295	3.635	3.43	3.72	3.46	3.46
cleveland	3.03	3.455	3.615	3.385	3.645	3.87
ecoli	3.285	3.715	3.51	3.3	3.39	3.8
haberman	3.225	3.62	3.54	3.705	3.525	3.385
Sonar	3.515	3.635	3.42	3.4	3.52	3.51
glass0	2.855	3.755	3.675	3.685	3.04	3.99
kyphosis	3.4	3.68	3.455	3.72	3.395	3.35
newthyroid1	2.935	3.78	3.685	3.735	3.455	3.41
Seeds	3.255	3.765	3.505	3.51	3.4	3.565
Vertebral	3.53	3.695	3.86	3.19	3.675	3.05
Mean	3.281	3.706	3.522	3.533	3.442	3.515

The F_1 measurements of the Naive Bayes classifier are given in Table 24 and the average rank in Table 25. In the 11 datasets, the best results were obtained ten times without resampling and one time with US with PSO. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was reached when RUS was applied. While the optimization method which gave the best result in F_1 rankings gave ABC, the method that gave the worst result was DE.

Table 24: F_1 values for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.8257±0.0372	0.7982±0.0712	0.8211±0.0564	0.8185±0.059	0.8212±0.0545	0.821±0.0512
bupa	0.6746±0.0753	0.604±0.1004	0.526±0.1055	0.5153±0.0964	0.6229±0.1117	0.5537±0.1141
cleveland	0.9266±0.0302	0.7282±0.0763	0.8877±0.0415	0.8864±0.0406	0.8858±0.0422	0.891±0.0399
ecoli	0.937±0.0355	0.9009±0.0467	0.9284±0.0374	0.924±0.0405	0.9299±0.0382	0.9291±0.0386
haberman	0.8311±0.0406	0.771±0.0647	0.7981±0.0563	0.802±0.0546	0.7977±0.0563	0.7991±0.0495
Sonar	0.6375±0.2684	0.5887±0.197	0.6117±0.2594	0.6477±0.2398	0.6293±0.2589	0.6513±0.2555
glass0	0.8051±0.0834	0.7306±0.1047	0.7286±0.1121	0.7378±0.0969	0.7728±0.1121	0.73±0.101
kyphosis	0.8899±0.0626	0.8093±0.1145	0.872±0.0905	0.8549±0.0906	0.8725±0.0967	0.8618±0.09
newthyroid1	0.9695±0.0321	0.9493±0.0402	0.9573±0.0337	0.9555±0.0339	0.9602±0.0309	0.9594±0.029
Seeds	0.8549±0.0914	0.8424±0.0987	0.8502±0.0896	0.8469±0.0895	0.8482±0.0892	0.8506±0.0925
Vertebral	0.8143±0.0649	0.7939±0.0698	0.7983±0.0664	0.7998±0.0697	0.8079±0.0679	0.8043±0.0648

Table 25: F_1 rank means for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	3.295	4.235	3.33	3.375	3.385	3.38
bupa	1.645	3.185	4.575	4.67	2.775	4.15
cleveland	1.375	5.95	3.4	3.525	3.52	3.23
ecoli	2.255	5.41	3.25	3.715	3.165	3.205
haberman	1.9	4.765	3.685	3.495	3.65	3.505
Sonar	3.2	4.265	3.65	3.19	3.47	3.225
glass0	2.225	3.945	4.095	3.89	2.87	3.975
kyphosis	2.73	4.715	3.255	3.625	3.19	3.485
newthyroid1	2.72	4.045	3.575	3.76	3.455	3.445
Seeds	3.255	3.77	3.455	3.57	3.565	3.385
Vertebral	2.79	3.965	3.84	3.755	3.215	3.435
Mean	2.49	4.386	3.646	3.688	3.296	3.493

G-mean measurements of the Naive Bayes classifier are given in Table 26 and average ranks are given in Table 27. In the 11 datasets, the best results were obtained four times without resampling and seven times in the RUS. In terms of rank averages, the best ranking was achieved by RUS, and the worst by US with GA. While the optimization method which gave the best result in G-mean rankings was ABC, the method that gave the worst result was the GA.

Table 26: G-mean values for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.3682±0.2003	0.6469±0.1213	0.5802±0.1327	0.6015±0.1294	0.5918±0.1413	0.591±0.1232
bupa	0.5828±0.0885	0.5844±0.0925	0.5455±0.0842	0.5413±0.0799	0.5708±0.0908	0.5572±0.0867
cleveland	0.2689±0.294	0.5469±0.207	0.3662±0.2895	0.3865±0.2796	0.3542±0.2772	0.3022±0.295
ecoli	0.7919±0.1345	0.8374±0.1012	0.8026±0.1306	0.8015±0.137	0.799±0.1292	0.8018±0.1356
haberman	0.467±0.1733	0.5992±0.1322	0.5502±0.14	0.5735±0.1484	0.5827±0.1287	0.569±0.1374
Sonar	0.7262±0.26	0.7672±0.1661	0.7163±0.266	0.7496±0.2347	0.7246±0.2469	0.7499±0.2397
glass0	0.7849±0.0842	0.7363±0.0886	0.733±0.0974	0.7414±0.0842	0.7652±0.1006	0.7314±0.0952
kyphosis	0.3715±0.3912	0.6034±0.3323	0.4756±0.3999	0.4635±0.3887	0.4601±0.4029	0.44±0.4073
newthyroid1	0.9632±0.0462	0.9512±0.0378	0.9586±0.0319	0.9569±0.0319	0.9587±0.0341	0.959±0.0312
Seeds	0.8903±0.0719	0.8838±0.0771	0.8895±0.0691	0.8871±0.0707	0.8866±0.0691	0.8877±0.0726
Vertebral	0.7768±0.0757	0.77±0.0711	0.77±0.0733	0.7728±0.0747	0.7738±0.0762	0.7733±0.0722

Table 27: G-mean rank means for Naive Bayes classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	5.435	2.48	3.41	3.205	3.245	3.225
bupa	2.865	2.955	4.025	4.03	3.485	3.64
cleveland	3.615	2.29	3.69	3.565	3.835	4.005
ecoli	2.915	3.97	3.49	3.745	3.375	3.505
haberman	4.94	2.8	3.625	3.155	3.155	3.325
Sonar	3.245	4.065	3.66	3.235	3.53	3.265
glass0	2.535	3.74	3.99	3.78	2.97	3.985
kyphosis	3.69	3.195	3.43	3.55	3.51	3.625
newthyroid1	2.81	3.995	3.525	3.74	3.485	3.445
Seeds	3.255	3.77	3.455	3.57	3.565	3.385
Vertebral	3.14	3.795	3.69	3.645	3.325	3.405
Mean	3.495	3.369	3.635	3.565	3.407	3.528

The AUC measurements of the Logistic Regression classifier are given in Table 28, and the average rank in Table 29. In the 11 datasets, the best results were obtained five times without resampling, one time in US with ABC and five times in US with PSO. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was achieved when RUS was applied. While the optimization method which was the best results in AUC rankings was ABC, the method that gave the worst results was DE.

Table 28: AUC values for logistic regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.7215±0.1073	0.7128±0.105	0.7224±0.1056	0.7252±0.1096	0.7202±0.1072	0.7225±0.1064
bupa	0.7107±0.0884	0.7092±0.0903	0.7042±0.0886	0.7036±0.0896	0.7077±0.0887	0.7047±0.0884
cleveland	0.7403±0.1578	0.7225±0.1555	0.7328±0.1556	0.7363±0.155	0.7377±0.1597	0.7359±0.1527
ecoli	0.9067±0.0708	0.8986±0.0737	0.9048±0.0732	0.9038±0.0723	0.9056±0.0719	0.9036±0.0718
haberman	0.6807±0.1273	0.6784±0.1221	0.6807±0.1348	0.6806±0.1225	0.6792±0.1232	0.6814±0.1272
Sonar	0.7314±0.1561	0.5458±0.212	0.6897±0.1929	0.5949±0.1948	0.6628±0.1994	0.6175±0.1964

glass0	0.8216±0.0858	0.8114±0.086	0.8134±0.0861	0.8106±0.0866	0.8225±0.083	0.8156±0.0865
kyphosis	0.8357±0.1641	0.8168±0.1701	0.8361±0.1628	0.8296±0.1722	0.8357±0.1632	0.8395±0.1625
newthyroid1	0.9778±0.054	0.9878±0.0442	0.9864±0.0513	0.985±0.0422	0.9861±0.0365	0.9932±0.0316
Seeds	0.9888±0.0336	0.9719±0.0486	0.9698±0.0526	0.9752±0.0503	0.9826±0.0379	0.9833±0.0363
Vertebral	0.9312±0.0425	0.9305±0.0402	0.9302±0.0442	0.932±0.0433	0.9299±0.0445	0.9319±0.0421

Table 29: AUC rank means for logistic regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	3.485	4	3.405	3.295	3.45	3.365
bupa	3.14	3.38	3.62	3.795	3.5	3.565
cleveland	3.115	3.895	3.8	3.4	3.31	3.48
ecoli	3.195	3.785	3.485	3.575	3.42	3.54
haberman	3.455	3.685	3.285	3.475	3.555	3.545
Sonar	2.655	4.27	3.13	3.985	3.31	3.65
glass0	3.195	3.55	3.695	3.77	3.13	3.66
kyphosis	3.31	3.85	3.455	3.58	3.52	3.285
newthyroid1	3.68	3.51	3.355	3.61	3.57	3.275
Seeds	2.995	3.875	3.79	3.57	3.29	3.48
Vertebral	3.47	3.595	3.48	3.335	3.71	3.41
Mean	3.245	3.763	3.5	3.581	3.433	3.478

The F_1 measures of the Logistic Regression classifier are given in Table 30, and the average rank in Table 31. In the 11 datasets, the best results were obtained ten times without resampling and one time with US with PSO. When the averages of the rankings were examined, it was seen that the best ranking was achieved without resampling and the worst ranking was reached when RUS was applied. While the optimization method which gave the best result in F_1 rankings was ABC, the method that gave the worst result was DE.

Table 30: F_1 values for Logistic Regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.8355±0.0474	0.7303±0.0773	0.785±0.0733	0.786±0.0807	0.7734±0.0833	0.781±0.0824
bupa	0.7337±0.0657	0.6787±0.08	0.584±0.0908	0.5675±0.1012	0.6892±0.1035	0.6255±0.0915
cleveland	0.9412±0.0142	0.7924±0.0638	0.9264±0.0317	0.9274±0.03	0.9252±0.0302	0.929±0.0262
ecoli	0.9573±0.0191	0.8928±0.0526	0.9534±0.0216	0.953±0.0223	0.9569±0.0202	0.9532±0.0213
haberman	0.8449±0.0253	0.7911±0.073	0.8417±0.0433	0.8364±0.055	0.8376±0.0458	0.842±0.0457
Sonar	0.4384±0.243	0.326±0.184	0.411±0.2199	0.3397±0.1981	0.4052±0.2274	0.3542±0.2112
glass0	0.8265±0.0646	0.7616±0.076	0.7734±0.0761	0.7651±0.0804	0.807±0.064	0.7778±0.0803
kyphosis	0.8743±0.0901	0.7844±0.1455	0.8561±0.1082	0.8579±0.1093	0.8626±0.1112	0.8551±0.1101
newthyroid1	0.9856±0.0221	0.9812±0.0283	0.986±0.023	0.9859±0.0227	0.9856±0.0211	0.9884±0.0198
Seeds	0.9341±0.0779	0.9274±0.0798	0.9253±0.0742	0.9236±0.0779	0.9263±0.0778	0.9199±0.0792
Vertebral	0.888±0.0464	0.8697±0.0531	0.872±0.0589	0.8704±0.0548	0.8834±0.0467	0.8773±0.0545

Table 31: F_1 rank means for Logistic Regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	1.735	5.15	3.375	3.325	3.79	3.625
bupa	1.75	2.755	4.92	5.09	2.62	3.865
cleveland	2.07	6	3.235	3.155	3.35	3.19
ecoli	2.64	5.705	3.235	3.28	2.765	3.375
haberman	3.09	5.07	2.99	3.34	3.455	3.055
Sonar	2.82	3.99	3.375	3.875	3.22	3.72
glass0	2.165	4.24	3.81	4.105	2.785	3.895
kyphosis	2.995	4.695	3.365	3.335	3.24	3.37
newthyroid1	3.51	3.745	3.475	3.445	3.545	3.28
Seeds	3.245	3.46	3.49	3.565	3.46	3.78
Vertebral	2.65	3.965	3.815	4.04	3.085	3.445
Mean	2.606	4.434	3.553	3.687	3.21	3.509

The G-mean measurements of the Logistic Regression classifier are given in Table 32, and the average of the rankings are given in Table 33. In the 11 datasets, the best results were obtained two times without resampling, seven times in RUS, one time in US with DE, and one time in US with PSO. When the rank averages are examined, it was seen that the best ranking was achieved when RUS was applied and the worst ranking was reached when no resampling was applied. While the optimization method which gave the best result in G-mean rankings was ABC, the method that gave the worst result was PSO.

Table 32: G-mean values for Logistic Regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	0.5676±0.1264	0.6443±0.0986	0.6534±0.1153	0.6659±0.1156	0.6579±0.1107	0.6533±0.1217
bupa	0.6277±0.0944	0.6512±0.0817	0.6028±0.0728	0.5917±0.0789	0.6219±0.0877	0.6188±0.0776
cleveland	0.1495±0.2492	0.6424±0.1875	0.3951±0.3086	0.3728±0.3211	0.3499±0.313	0.3277±0.3099
ecoli	0.6304±0.2396	0.8087±0.1089	0.6991±0.2258	0.7334±0.1961	0.6913±0.2328	0.7038±0.2171
haberman	0.3366±0.1824	0.6211±0.1288	0.5238±0.1594	0.5384±0.1656	0.528±0.1766	0.5237±0.1607
Sonar	0.5984±0.2861	0.5068±0.242	0.5932±0.2585	0.5143±0.2604	0.5719±0.2703	0.5313±0.2745
glass0	0.6978±0.1216	0.7263±0.086	0.7201±0.1008	0.7238±0.0948	0.7111±0.0915	0.7196±0.1032
kyphosis	0.4359±0.3947	0.6308±0.3119	0.4907±0.3897	0.5615±0.3737	0.5065±0.3945	0.5442±0.3718
newthyroid1	0.9531±0.0893	0.9611±0.0886	0.9575±0.0909	0.9633±0.0836	0.9661±0.0616	0.9714±0.0575
Seeds	0.9503±0.0619	0.9477±0.061	0.9452±0.059	0.9436±0.0611	0.9461±0.0601	0.9413±0.0612
Vertebral	0.8248±0.0714	0.8393±0.0655	0.8318±0.0734	0.8336±0.0697	0.8287±0.0682	0.8317±0.0748

Table 33: G-mean rank means for Logistic Regression classifier

Dataset	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
breast	4.925	3.59	3.14	2.89	3.215	3.24
bupa	3.325	2.42	4.045	4.315	3.51	3.385
cleveland	4.6	1.94	3.39	3.42	3.805	3.845
ecoli	3.855	2.935	3.69	3.4	3.43	3.69
haberman	5.275	1.94	3.47	3.33	3.395	3.59
Sonar	2.895	3.995	3.315	3.835	3.29	3.67
glass0	3.8	3.255	3.4	3.435	3.615	3.495
kyphosis	3.595	3.445	3.705	3.375	3.45	3.43
newthyroid1	3.52	3.765	3.465	3.435	3.535	3.28
Seeds	3.245	3.46	3.49	3.565	3.46	3.78
Vertebral	3.73	3.36	3.525	3.435	3.535	3.415
Mean	3.888	3.1	3.512	3.494	3.476	3.529

4. Results and Discussion

The rank means for all cases is given in Table 34. Kruskal Wallis H test was used to test whether the difference between the rank means found for all cases was statistically significant. Since there was a significant difference between all cases, Dunn test was performed for pairwise comparisons. Significance levels of all these statistical tests are given in Table 35. The following inferences were made on the basis of these results:

- For AUC, F_1 and G-mean, there is a significant difference between the rank averages according to resampling status.
- For AUC, F_1 and G-mean, there is no significant difference between the no resampling and RUS rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the no resampling and GA rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the RUS and GA rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the no resampling and DE rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the RUS and DE rank averages.
- For AUC, there is no significant difference between the GA and DE rank averages. For F_1 , there is a significant difference between the GA and DE rank averages. For G-mean, there is no significant difference between the GA and DE rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the no resampling and ABC rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the RUS and ABC rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the GA and ABC rank averages.
- For AUC and F_1 there is a significant difference between the DE and ABC rank averages. There is no significant difference between the DE and ABC rank averages for G-mean.
- For AUC, F_1 and G-mean, there is a significant difference between the no resampling and PSO rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between RUS and PSO rank averages.
- For AUC and G-mean, there is a significant difference between GA and PSO rank averages. For F_1 , there is no significant difference between GA and PSO rank averages.
- For AUC, there is no significant difference between the DE and PSO rank averages. For F_1 and G-mean, there is a significant difference between the DE and PSO rank averages.
- For AUC, F_1 and G-mean, there is a significant difference between the ABC and PSO rank averages.

Table 34: Rank means for all cases

Criterion	No resampling	RUS	US with GA	US with DE	US with ABC	US with PSO
AUC	3.290	3.684	3.519	3.523	3.422	3.560
F_1	2.563	4.453	3.579	3.643	3.278	3.485
G-mean	3.747	3.194	3.524	3.505	3.467	3.563

Table 35: Kruskal-Wallis H test and pairwise comparisons for the mean of all cases

Test		p_{AUC}	p_{F_1}	p_{Gmean}
Kruskall-Wallis H Test		<0.001*	<0.001*	<0.001*
	1 - 2	<0.001*	<0.001*	<0.001*
	1 - 3	<0.001*	<0.001*	<0.001*
	2 - 3	<0.001*	<0.001*	<0.001*
	1 - 4	<0.001*	<0.001*	<0.001*
	2 - 4	<0.001*	<0.001*	<0.001*
	3 - 4	0.439	0.012	0.252
Dunn Test	1 - 5	<0.001*	<0.001*	<0.001*
	2 - 5	<0.001*	<0.001*	<0.001*
	3 - 5	<0.001*	<0.001*	0.019
	4 - 5	<0.001*	<0.001*	0.079
	1 - 6	<0.001*	<0.001*	<0.001*
	2 - 6	<0.001*	<0.001*	<0.001*
	3 - 6	0.070	0.001	0.100
	4 - 6	0.092	<0.001*	0.026
	5 - 6	<0.001*	<0.001*	<0.001*

5. Conclusion

The problem with class imbalance is that the dependent variable contains different class observation numbers. This situation affects most classification methods. Resampling methods are the most common methods for solving the class imbalance problem. One of the resampling methods is undersampling method. The aim of the undersampling method is to achieve balance by extracting certain observations within the positive class. It is very important to correctly select these observations, as discarding observations causes information loss.

In this study, optimization methods were used in the selection of observations to be extracted in undersampling. The optimization methods used are GA, DE, ABC and PSO. According to the general rank means of these methods, the best optimization method was found to be ABC. The difference between the ABC method and other methods was found to be significant as a result of the Kruskal Wallis H and Dunn tests.

Four optimization methods were used in this study. Since these optimization methods were used in the default settings, they are open to investigating how effective they are at different settings. In addition, it is possible to use different classifiers and different performance criteria in the objective function.

References

- [1] Chen, L., Bao, L., Li, J., Cai, S., Cai, C., Chen, Z., *An aliasing artifacts reducing approach with random undersampling for spatiotemporally encoded single-shot MRI*, *Journal of Magnetic Resonance*, 237, 115-124, 2013.
- [2] Liu, B., Tsoumakas, G., *Dealing with class imbalance in classifier chains via random undersampling*, *Knowledge-Based Systems*, 192:105292, 2019.
- [3] Tomek, I., *Two modifications of CNN*, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6 (11), 769-772, 1976.
- [4] Elhassan, T., Aljourf, M., Al-Mohanna, F., Shoukri, M., *Classification of imbalance data using tomek link (t-link) combined with random under-sampling (rus) as a data reduction method*, *Global Journal of Technology and Optimization*, S1, 2017.
- [5] Pereira, R.M., Costa, Y.M., Silla Jr, C.N., *MLTL: A multi-label approach for the Tomek Link undersampling algorithm*, *Neurocomputing*, 383, 95-105, 2020.
- [6] Devi, D., Purkayastha, B., *Redundancy-driven modified Tomek-link based undersampling: A solution to class imbalance*, *Pattern Recognition Letters*, 93, 3-12, 2017.
- [7] Wilson, D. L., *Asymptotic properties of nearest neighbor rules using edited data*, *IEEE Transactions on Systems, Man, and Cybernetics*, 3, 408-421, 1972. <https://doi.org/10.1109/TSMC.1972.4309137>.
- [8] Laurikkala, J., *Improving identification of difficult small classes by balancing class distribution*, In *Conference on Artificial Intelligence in Medicine in Europe* (pp. 63-66), Springer, Berlin, Heidelberg, 2001. https://doi.org/10.1007/3-540-48229-6_9.
- [9] Bach, M., Werner, A., Palt, M., *The Proposal of Undersampling Method for Learning from Imbalanced Datasets*, *Procedia Computer Science*, 159, 125-134, 2019.
- [10] Lu, W., Li, Z., Chu, J., *Adaptive ensemble undersampling-boost: a novel learning framework for imbalanced data*, *Journal of Systems and Software*, 132, 272-282, 2017.
- [11] Lin, W.C., Tsai, C.F., Hu, Y.H., Jhang, J.S., *Clustering-based undersampling in class-imbalanced data*, *Information Sciences*, 409, 17-26, 2017.
- [12] Ofek, N., Rokach, L., Stern, R., Shabtai, A., *Fast-CBUS: A fast clustering-based undersampling method for addressing the class imbalance problem*, *Neurocomputing*, 243, 88-102, 2017.
- [13] Körzdörfer, G., Pfeuffer, J., Kluge, T., Gebhardt, M., Hensel, B., Meyer, C.H., Nittka, M., *Effect of spiral undersampling patterns on FISP MRF parameter maps*, *Magnetic Resonance Imaging*, 2019.
- [14] Holland, J.H., *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, MIT press, 1992.
- [15] Eberhart, R., Kennedy, J., *Particle swarm optimization*, In *Proceedings of the IEEE International Conference on Neural Networks* 4, 1942-1948, 1995.
- [16] Storn, R., Price, K., *Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces*, *Journal of Global Optimization*, 11(4), 341-359, 1997.
- [17] Karaboga, D., *An idea based on honey bee swarm for numerical optimization*, Technical Report-tr06, Erciyes University, Engineering Faculty, Computer Engineering Department, 200, 1-10, 2005.
- [18] García, S., Herrera, F., *Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy*, *Evolutionary computation*, 17(3), 275-306, 2009.
- [19] Roshan, S.E., Asadi, S., *Improvement of Bagging performance for classification of imbalanced datasets using evolutionary multi-objective optimization*, *Engineering Applications of Artificial Intelligence*, 87, 103319, 2020.
- [20] Yu, H., Mu, C., Sun, C., Yang, W., Yang, X., Zuo, X., *Support vector machine-based optimized decision threshold adjustment strategy for classifying imbalanced data*, *Knowledge-Based Systems*, 76, 67-78, 2015.

- [21] Lavine, B.K., *Clustering and classification of analytical data*, Encyclopedia of Analytical Chemistry: Instrumentation and Applications, 2000.
- [22] He, H., Garcia, E.A., *Learning from imbalanced data*, IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263-1284, 2009.
- [23] Das, B., Krishnan, N.C., Cook, D.J., *RACOG and wRACOG: Two probabilistic oversampling techniques*, IEEE Transactions on Knowledge and Data Engineering, 27(1), 222-234, 2014.
- [24] Fernàndes, E. R., de Carvalho, A.C., *Evolutionary inversion of class distribution in overlapping areas for multi-class imbalanced learning*, Information Sciences, 494, 141-154, 2019.
- [25] Fernández, A., García, S., Galar, M., Prati, R.C., Krawczyk, B., Herrera, F., *Learning from imbalanced data sets*, Berlin: Springer, 1-377, 2018.
- [26] Chawla, N.V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W.P., *SMOTE: synthetic minority over-sampling technique*, Journal of Artificial Intelligence Research, 16, 321-357, 2002.
- [27] Kamalov, F., *Kernel density estimation based sampling for imbalanced class distribution*, Information Sciences, 2019.
- [28] Maldonado, S., Weber, R., Famili, F., *Feature selection for high-dimensional class-imbalanced data sets using Support Vector Machines*, Information Sciences, 286, 228-246, 2014.
- [29] Moayedikia, A., Ong, K.L., Boo, Y.L., Yeoh, W.G., Jensen, R., *Feature selection for high dimensional imbalanced class data using harmony search*, Engineering Applications of Artificial Intelligence, 57, 38-49, 2017.
- [30] Wong, M.L., Seng, K., Wong, P.K., *Cost-sensitive ensemble of stacked denoising autoencoders for class imbalance problems in business domain*, Expert Systems with Applications, 141, 112918, 2020.
- [31] Dua, D., Graff, C., *UCI Machine Learning Repository* [<http://archive.ics.uci.edu/ml>], Irvine, CA: University of California, School of Information and Computer Science, 2019.



Prediction of Human Development Index with Health Indicators Using Tree-Based Regression Models

Pelin AKIN¹, Tuba KOÇ^{2,*}

¹Çankırı Karatekin University, Faculty of Science, Department of Statistics, 18100, Çankırı, Türkiye
pekinakin@karatekin.edu.tr, ORCID: 0000-0003-3798-4827

²Çankırı Karatekin University, Faculty of Science, Department of Statistics, 18100, Çankırı, Türkiye
tubakoc@karatekin.edu.tr, ORCID: 0000-0001-5204-0846

Received: 11.03.2021

Accepted: 19.11.2021

Published: 31.12.2021

Abstract

Machine learning is a field of artificial intelligence that allows computers to predict and model future events by making inferences from past information with mathematical and statistical operations. In this study, we used tree-based regression models, one of the machine learning methods, to determine and predict the effect of health indicators of 191 countries on the human development index (HDI) between 2014 and 2018 years. When tree-based regression models were compared according to model performance criteria, it was found that the best model was the gradient boosting model with the highest $R^2 = 0.9962$ and the smallest RMSE = 0.0094. With the gradient boosting model, the three most important variables to HDI are; current health expenditure per capita, physicians and nurses, and midwives, respectively. By selecting the ten countries with the highest HDI values and Turkey, HDI values were estimated for 2018-2019 with a gradient boosting model. The countries for which HDI values are best predicted by the gradient boosting method are Netherlands, Sweden, Norway, Iceland, Denmark, Turkey, Ireland, Germany, Australia, and China.



Keywords: Machine learning algorithms; Tree-based regression model; Gradient boosting method; Human development index; Health indicators.

Ağaç Tabanlı Regresyon Modelleri Kullanılarak Sağlık Göstergeleri ile İnsani Gelişme Endeksinin Tahmini

Öz

Makine öğrenmesi, bilgisayar yardımıyla geçmişteki bilgileri kullanarak matematiksel ve istatistiksel işlemlerle çıkarımlar elde eden ve gelecekteki olaylar hakkında tahmin yürütülmesi modelleme yapılmasına imkân veren bir yapay zekâ alanıdır. Bu çalışmada 191 ülkenin 2014-2018 yıllarında sağlık göstergelerinin insani gelişim endeksi (İGE) üzerindeki etkisini belirlemek ve tahmin yapmak için makine öğrenmesi yöntemlerinden ağaç tabanlı regresyon modelleri kullanılmıştır. Ağaçlı tabanlı regresyon modelleri model performans kriterlerine göre karşılaştırıldığında en iyi modelin en yüksek $R^2 = 0.9962$ ve en küçük $RMSE = 0.0094$ değeri ile gradyan artırma model olduğu bulunmuştur. Gradyan artırma model ile İGE indeksine en fazla etki eden 3 değişken sırasıyla: kişi başına cari sağlık harcaması, doktorların sayısı ve hemşireler ile ebelerin sayısı olarak bulunmuştur. İGE değeri en yüksek olan 10 ülke ve Türkiye seçilerek gradyan artırma model ile 2018-2019 yılları için İGE değerleri tahmin edilmiştir. Gradyan artırma yöntemi ile İGE değeri en iyi tahmin edilen ülkeler sırasıyla Hollanda, İsveç, Norveç, İzlanda, Danimarka, Türkiye, İrlanda, Almanya, Avustralya ve Çin şeklindedir.

Anahtar Kelimeler: Makine öğrenmesi algoritmaları; Ağaç tabanlı regresyon modelleri; Gradyan artırma model; İnsani gelişim endeksi; Sağlık değişkenleri.

1. Introduction

Machine learning, a new and promising sub-branch of algorithmic data analysis, has rapidly advanced in recent years. Due to the rapid increase in available storage space, processing power, and network connectivity, there has been great progress in data collection, sharing, and processing technologies. Also, given the recent increase in the volume of data from all sources, it is possible to apply learning methods in increasingly complex data that are impossible to analyze with prior technology. Machine learning programs computers use statistical theory to create mathematical models to optimize a performance criterion using sample data or past experiences [1]. Several algorithms are used in machine learning to categorize data sets, analyze their results, and make predictions. Machine learning algorithms are widely used in many sectors such as education, economy, marketing, health, etc. In healthcare, machine learning algorithms are used in many areas such as prediction, diagnosis, disease tracking, and processing of unstructured data [2, 3].

Health services are considered one of the main determinants of economic and social development. Health data belonging to health services guide the development of countries in the field of health. The Human Development Index (HDI) is a concept introduced by the United Nations to measure countries' human development levels. HDI determines the human development levels of societies through three main areas: health, education, and income. Yakut et al. [4] using infant mortality rate, gross national product, high school enrollment rate, growth, foreign direct investment, energy consumption, energy production, inflation, exports, number of internet users, unemployment, imports, mobile phone subscribers, and health expenditures. They made a classification using ordered logistic regression analysis and artificial neural networks on the Human Development Index of 81 countries. Zhang et al. [5] applied the Gradient Boosting Decision Tree Algorithm (GBDT) to analyze health data and make predictions. They concluded that the GBDT algorithm performs better than the traditional least-squares method, ridge regression, lasso regression, ElasticNet, SVR, and KNN algorithm methods. König et al. [6] analyzed the impact of multimorbidity on health care costs in Germany on all sectors of care using an advanced tree-based graphic model. Rençber and Mete [7] classified countries according to the Human Development Index (HDI) using machine learning techniques such as Artificial Neural Network (ANN) and Adaptive Neural Fuzzy Inference System (ANFIS) and compared the results with the HDI. Yakut and Korkmaz [8] created decision trees with C5.0 and Gini algorithms using data from 79 countries from 2010-2017. They determined the HDI factors by the decision trees method and classified the countries as very high, high, medium, and low-level developed countries. They determined that the variables that affect most the HDI are education, employment, and health indicators. Dos Santos et al. [9] used the SMOReg data mining algorithm to predict Latin American countries' human development index and life expectancy. Hu et al. [10] examined the critical factors of high costs for breast cancer patients using the Quantile Regression Forests approach, a flexible tree-based machine learning technique using health data. Saboo et al. [11] compared the ANN and linear regression-based approach in estimating the HDI. Coşar [12] made a classification using Naive Bayes, ANN, and logistic regression methods to determine the effect of healthcare indicators in OECD (Organization for Economic Co-operation and Development) countries on the HDI.

In this study, we used tree-based regression models, one of the machine learning methods, to determine and predict the effect of health indicators of all countries on the human development index. The article is divided as follows: In Section 2, decision tree, random forest, extreme gradient boosting, and gradient boosting from Tree-Based Regression models are defined. In Section 3, the application of Tree-Based Regression models is explained with data for all countries. Finally, a brief discussion is given in Section 4.

2. Tree-Based Regression Models

Tree-based regression models use one or more decision trees. We considered four tree-based machine learning methods, decision tree (CART), random forest regression, extreme gradient boosting (XGBoost), and gradient boosting model.

2.1. Decision tree

CART was created by [13]. The CART programs construct classification or regression models of a very general structure using a two-step process; the resulting models may be represented as binary trees [14]. The mean squared error is used for the split data in the CART algorithm. Mean squared error (MSE) for a specific node is defined as;

$$MSE_{node} = \frac{1}{m_{node}} \sum (y_i - \bar{y}_{node})^2. \quad (1)$$

If it is assumed that there is a binary split on each node on the tree, it will be divided into left and right. For each division, the error term with

$$MSE_{left} = \frac{1}{m_{left}} \sum (y_i - \bar{y}_{left})^2, \quad (2)$$

$$MSE_{right} = \frac{1}{m_{right}} \sum (y_i - \bar{y}_{right})^2. \quad (3)$$

For each attribute j , the following formula is calculated,

$$\min(MSE_{left} + MSE_{right}). \quad (4)$$

The smallest of the values is chosen. The dataset splits recursively, which means that the subsets that meet a partition are partitioned until they reach a predetermined expiration criterion [14].

2.2. Random Forest regression

The ensemble learning algorithms produce a prediction model by combining the strong points of a group of simpler, a lot of basic models [15]. The most widely used ensemble learning algorithms are bagging and Random Forest algorithms. Breiman's Random Forest classification is an improved version of the bagging technique achieved by adding the randomness feature. The following steps are taken for the Random Forest algorithm: firstly, n bootstrap samples are taken from the original data set. Then CART is created for each bootstrap sampling. A new estimate is made by combining the estimates made by n trees separately. Estimation is made by taking the average of the results made in regression trees [16].

2.3. Extreme gradient boosting

Chen and Guestrin [17] proposed a scalable, end-to-end tree strengthening system called XGBoost in 2016. XGBoost, its algorithm is also called a regular gradient boosting technique. The XGBoost algorithm is optimized by using different arrangements of the Gradient Boosting algorithm. XGBoost is fast to interpret, prevents overlearning, and can handle large-sized datasets well.

2.4. Gradient boosting model

Friedman [18, 19] established the foundation for the next generation of boosting algorithms. Ridgeway [20] proposed a gradient boosting model using the link between upgrade and optimization. This algorithm can be preferred for regression and classification problems. It creates a model with a combination of its weak models. Increasing the gradient here aims to reach the minimum error values by updating the predictions according to the learning rate.

2.5. Evaluation metrics for regression models

The mean absolute error (MAE), the mean squared error (MSE), the root mean squared error (RMSE), and the coefficient of determination (R^2) is used for model selection. R^2 is used to measure the wellness of the fit by the trained models. MAE, MSE, RMSE are the average error measures [25]. The error measures and R^2 are defined as follows

$$\begin{aligned}
 \text{MAE} &= \frac{1}{N} (\sum |y_i - \hat{y}|), \\
 \text{MSE} &= \frac{1}{N} \sum (y_i - \hat{y}), \\
 \text{RMSE} &= \sqrt{\frac{1}{N} \sum (y_i - \hat{y})}, \\
 R^2 &= 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}.
 \end{aligned} \tag{5}$$

3. Application Part

In this study, the health indicators and HDI of 2014 and 2018 were used for 191 countries. The data obtained are available URL1-2 [21, 22]. The variables used in the study are given in Table 1.

Table 1: Description of the variables

Variable	Description
HDI	Human development index
x_1	Gross domestic product growth (annual %) (GDP)
x_2	Current health expenditure (% of GDP)
x_3	Hospital beds (per 1,000 people)
x_4	Specialist surgical workforce (per 100,000 population)
x_5	Current health expenditure per capita (current US\$)
x_6	Nurses and midwives (per 1,000 people)
x_7	Physicians (per 1,000 people)

During the preparation of the data set for analysis, the multiple assignments (MICE) method with chained equations was used for the missing value. In the MICE methods, a statistical distribution is obtained over the data set. Then thanks to this distribution, a link is used that fills in the missing value [23]. After the data set was completed, the data for the years 2014-2017 were divided as training and the data for 2018 as test data. CART, random forest regression, XGBoost, and gradient boosting model were applied. Analyzes were performed using version 3.5.2 of the R software.

Performance criteria for training and test data are given in Table 2 to compare models.

Table 2: Performance measurement for Models

MODELS	Train Data				Test Data			
	R ²	RMSE	MAE	MSE	R ²	RMSE	MAE	MSE
CART	0.8309	0.0620	0.0441	0.0038	0.8704	0.0546	0.0395	0.0029
Random forest	0.9884	0.0170	0.0116	0.0003	0.8865	0.0579	0.0416	0.0034
XGBoost	0.9884	0.0170	0.0117	0.0002	0.7181	0.0884	0.0664	0.0078
Gradient boosting model	0.9962	0.0094	0.0073	0	0.8942	0.0539	0.0394	0.0028

The model with the highest R² value and lower error rates than RMSE, MAE, and MSE shows the best performance. In Table 2, the algorithm that gives the best performance in predicting training data and test data is the gradient boosting model. When the gradient boosting model is applied to the training data, the visible results of the model are as shown in Fig. 1.

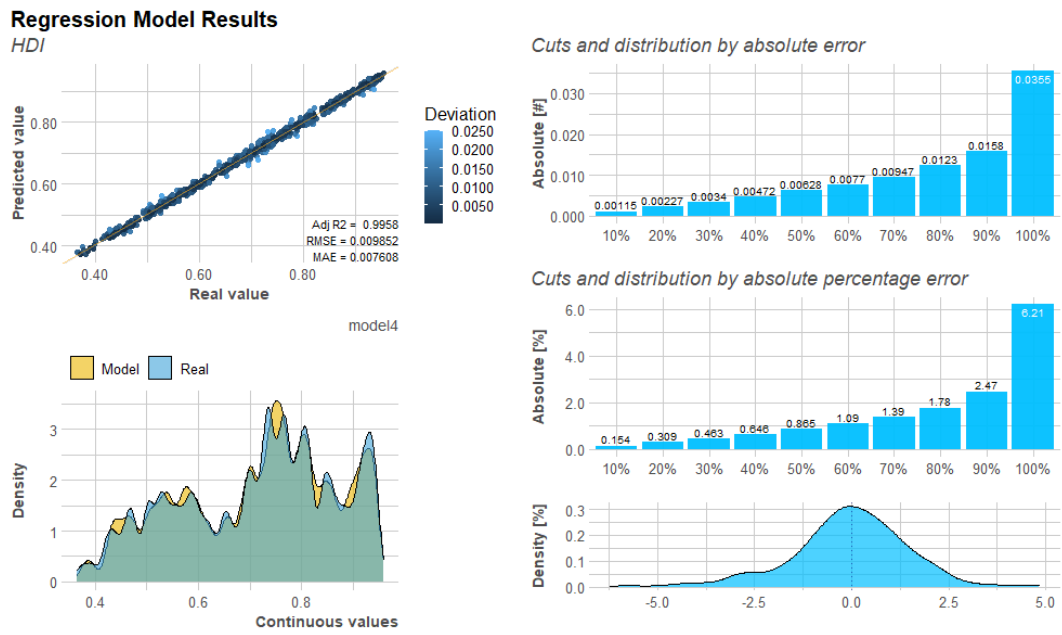


Figure 1: Gradient boosting regression model results for train data

Figure 1 shows the regression results plot, errors plot, and distribution plot. When the graphs are examined, it is seen that the gradient boosting model is suitable for the data. Feature importance graph describes which features are relevant.

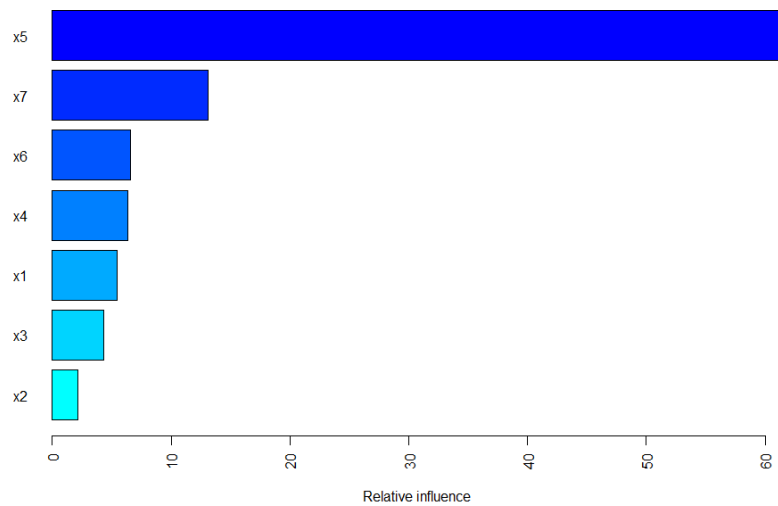


Figure 2: Gradient boosting model feature importance bar chart

In Fig. 2, it is seen that the three most important variables for the gradient boosting model used to determine the effect of health indicators on HDI are x5, x7, and x6, respectively.

When the gradient boosting model is applied to the data of the ten countries with the highest HDI index and Turkey, the predicted values for 2018 and 2019 are as follows.

Table 3: HDI index prediction with gradient boosting model

Countries	HDI ₂₀₁₈	Predicted ₂₀₁₈	Predicted ₂₀₁₉
Norway	0.947	0.956	0.95304
Switzerland	0.953	0.955	0.90801
Ireland	0.922	0.951	0.93866
Germany	0.908	0.946	0.91917
China	0.879	0.946	0.92261
Iceland	0.958	0.946	0.95179
Australia	0.893	0.943	0.92787
Sweden	0.951	0.943	0.90990
Netherlands	0.943	0.942	0.93263
Denmark	0.952	0.939	0.93321
Turkey	0.804	0.817	0.85975

When the 2018 actual values (HDI) of the countries selected from Table 3 and the predicted values for 2018 and 2019 are examined, the country with the best-predicted HDI value with the gradient boosting method is found to be the Netherlands, and the worst predicted country China. HDI forecast values change charts of selected countries for 2018 and 2019 are given in Fig. 3.

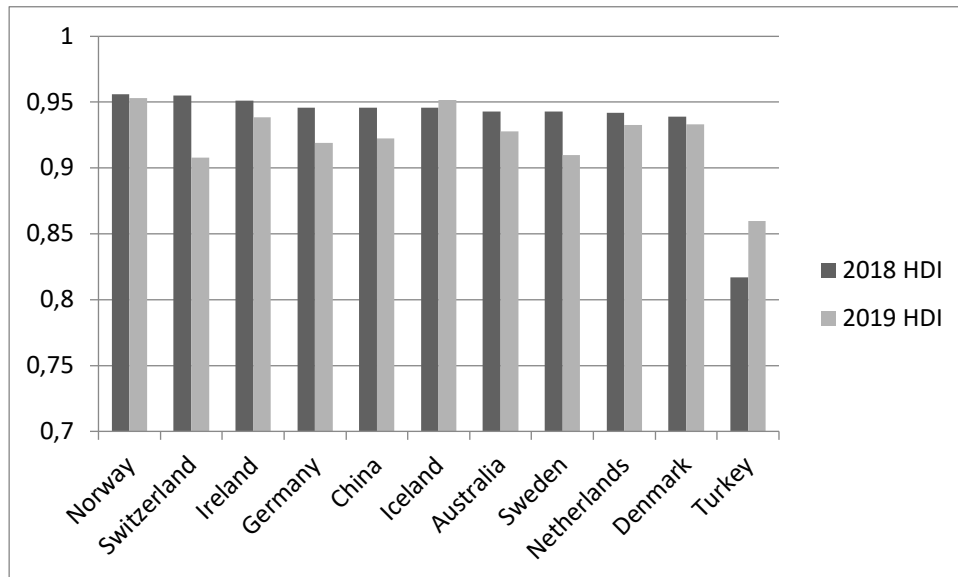


Figure 3: HDI forecast values comparison graph of countries

Figure 3 shows HDI predictions with health indicators, and it is seen that HDI values of Iceland and Turkey increase, while HDI values of Norway, Switzerland, Ireland, Germany, China, Australia, Sweden, Netherlands, and Denmark countries decrease.

4. Conclusion and Discussion

Human development, which is a concept that aims to raise the living standards of societies to the living standards of the modern world, is defined as the process that enables individuals to live their lives as they value and to exercise their basic human rights [25]. Health is one of the three most important components of the HDI.

In this study, the health indicators and HDI of 2014 and 2018 were used for 191 countries. Health indicators; GDP, current health expenditure, hospital beds specialist surgical workforce, current health expenditure per capita, nurses and midwives, and physicians have been selected from several indicators that may potentially impact the HDI. First, tree-based decision tree, random forest, extreme gradient boosting, and gradient boosting model methods, which are machine learning algorithms, were applied to the training data to determine the effect of health indicators on HDI of all countries and to make predictions. When the model performance criteria were examined, it was found that the best model was the gradient boosting model with the highest $R^2 = 0.9962$ and the smallest RMSE = 0.0094. According to the gradient boosting model results, the three variables that have the greatest effect on the HDI index are current health expenditure per capita, physicians, nurses, and midwives. Then ten countries with the highest HDI values and Turkey were chosen, and HDI values were estimated for 2018-2019 with the gradient boosting model. The countries with the best HDI value estimated by gradient boosting method are Netherlands, Sweden, Norway, Iceland, Denmark, Turkey, Ireland, Germany, Australia, and China, respectively. The gradient boosting method has estimated HDI values with the same mistake for Turkey and Denmark. However, Turkey has increased the HDI in the years, lagged behind many developed countries. Considering the analysis results, it is seen that the effect of the level of health expenditure on the HDI is quite high. Therefore, countries should allocate more resources to the field of health. Also, given the impact of numbers of physicians, nurses, and midwives on human development, it is obvious that more healthcare staff investment will increase human development at higher rates. These results are very consistent with the literature [8]. The limitation of this study is that the HDI index of 2020 and 2021 could not be predicted. HDI predictive values can be calculated when current data on health variables of countries are available.

References

- [1] Alpaydın, E., *Yapay öğrenme*, Boğaziçi Üniversitesi Yayınevi, 2011.
- [2] Alonso, D.H., Wernick, M.N., Yang, Y.Y., Germano, G., Berman, D.S., Slomka, P., *Prediction of cardiac death after adenosine myocardial perfusion SPECT based on machine learning*, *Journal of Nuclear Cardiology*, 26(5), 1746-1754, 2019.
- [3] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., and Chouvarda, I., *Machine learning and data mining methods in diabetes research*, *Computational and Structural Biotechnology Journal*, 15, 104-116, 2017.
- [4] Yakut, E., Gündüz, M., and Demirci, A., *Comparison of classification success of human development index by using ordered logistic regression analysis and artificial neural network methods*, *Journal of Applied Quantitative Methods*, 10(3), 15-34, 2015.
- [5] Zhang, B., Ren, J., Cheng, Y., Wang, B., Wei, Z., *Health data driven on continuous blood pressure prediction based on gradient boosting decision tree algorithm*, *IEEE Access*, 7, 32423-32433, 2019.
- [6] Konig, H.H., Leicht, H., Bickel, H., Fuchs, A., Gensichen, J., Maier, W., Mergenthal, K., Riedel-Heller, S., Schafer, I., Schon, G., Weyerer, S., Wiese, B., van den Bussche, H., Scherer, M., Eckardt, M., Grp, M.S., *Effects of multiple chronic conditions on health care costs: an analysis based on an advanced tree-based regression model*, *BMC Health Services Research* 13(1), 1-13, 2013.
- [7] Rençber, Ö.F., Sinan, M., *Reclassification Of Countries According To Human Development Index: An Application With Ann And Anfis Methods*, *Business & Management Studies: An International Journal*, 6(3), 228-252, 2018.
- [8] Yakut, E., Korkmaz, A., *İnsani Gelişmişlik Endeksinin Karar Ağacı Algoritmaları ile Modellenmesi: BM'de Bir Uygulama 2010-2017 Dönemi*, *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 20(2), 65-84, 2020.
- [9] dos Santos, C.B., Pilatti, L.A., Pedroso, B., Carvalho, D.R., Guimaraes, A.M., *Forecasting the human development index and life expectancy in Latin American countries using data mining techniques*, *Ciência & Saúde Coletiva*, 23(11), 3745-3757, 2018.
- [10] Hu, L., Li, L., Ji, J., Sanderson, M., *Identifying and understanding determinants of high healthcare costs for breast cancer: a quantile regression machine learning approach*, *BMC Health Services Research*, 20(1), 1066, 2020.
- [11] Saboo, A., Parakh, R., Trivedi, P., Potdar, M., *A Comparative Study of Artificial Neural Networks and Multiple Linear Regression by Predicting Human Development Index*, *International Journal of Scientific & Engineering Research*, 7(9), 424-428, 2016.
- [12] Çoşar, K., *OECD sağlık verilerinin veri madenciliği yöntemleri ile analizi*, Marmara University, İstanbul, 2020.
- [13] Breiman, L., Friedman, J., Olshen, R., Stone, C., *Classification and regression trees*, Chapman and Hall/CRC, 1998.
- [14] Therneau, T.M., Atkinson, E.J., *An introduction to recursive partitioning using the RPART routines*, Technical report Mayo Foundation, 1997.
- [15] Friedman, C., Sandow, S., *Utility-based learning from data*, Machine learning & pattern recognition series, Boca Raton: Chapman & Hall/CRC, 397 pages, 2011.
- [16] Liaw, A., Wiener, M., *Classification and regression by random Forest*, *R News*, 2(3), 18-22, 2002.

- [17] Chen, T., Guestrin, C., *Xgboost: A scalable tree boosting system*, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
- [18] Friedman, J. H., *Greedy function approximation: a gradient boosting machine*, *Annals of Statistics*, 1189-1232, 2001.
- [19] Friedman, J.H., *Stochastic gradient boosting*, *Computational Statistics & Data Analysis*, 38(4), 367-378, 2002.
- [20] Ridgeway, G., *Generalized boosted models: A guide to the gbm package*, *Update*, 1(1), 2007, 2007.
- [21] URL-1, <https://data.worldbank.org/> Erişim tarihi 25.12.2020.
- [22] URL-2, <http://hdr.undp.org/en/content/human-development-index-hdi> Erişim tarihi 25.12.2020
- [23] Şeker, Ş.E., Eşmekaya, E., *Eksik verilerin tamamlanması (imputation)*, *YBS Ansiklopedi*, 4, 10-17, 2017.
- [24] Uğuz, S., *Makine Öğrenmesi Teorik Yönleri ve Python Uygulamaları ile Bir Yapay Zeka Ekolü*, Nobel, Ankara, 2019.
- [25] Uygur, S., Yıldırım, F., *Cinsiyete bağlı insani gelişme endeks yaklaşımları: Türkiye örneği*, *Tisk Akademi*, 30-59, 2011.



Invariant and Lacunary Invariant Statistical Equivalence of Order β for Double Set Sequences

Uğur ULUSU^{1,*}, Erdinç DÜNDAR², Fatih NURAY³

¹Sivas Cumhuriyet University, 58140, Sivas, Türkiye

ugurulusu@cumhuriyet.edu.tr, ORCID: 0000-0001-7658-6114

²Afyon Kocatepe University, Department of Mathematics, 03200, Afyonkarahisar, Türkiye

edundar@aku.edu.tr, ORCID: 0000-0002-0545-7486

³Afyon Kocatepe University, Department of Mathematics, 03200, Afyonkarahisar, Türkiye

fnuray@aku.edu.tr, ORCID: 0000-0003-0160-4001

Received: 24.06.2021

Accepted: 24.11.2021

Published: 31.12.2021

Abstract

In this study, as a new approach to the concept of asymptotical equivalence in the Wijsman sense for double set sequences, the new concepts which are called asymptotical invariant statistical equivalence of order β and asymptotical lacunary invariant statistical equivalence of order β ($0 < \beta \leq 1$) in the Wijsman sense for double set sequences are introduced and explained with examples. In addition, the existence of some relations between these concepts and furthermore, the relationships between these concepts and previously studied asymptotical equivalence concepts in the Wijsman sense for double set sequences are investigated.

Keywords: Asymptotical equivalence; Convergence in the Wijsman sense; Double lacunary sequence; Invariant statistical convergence; Order β ; Sequences of sets.

Çift Küme Dizileri için β ncı Mertebeden İnvaryant ve Lacunary İnvaryant İstatistiksel Denklik



Öz

Bu çalışmada, çift küme dizileri için Wijsman anlamında asimptotik denklik kavramına yeni bir yaklaşım olarak, çift küme dizileri için Wijsman anlamında β ($0 < \beta \leq 1$) yncı mertebeden asimptotik invaryant istatistiksel denklik ve asimptotik lacunary invaryant istatistiksel denklik olarak adlandırılan yeni kavramlar tanıtıldı ve örneklerle açıklandı. Ayrıca, bu kavramlar arasında bazı ilişkilerin varlığı ve dahası bu kavramlar ve daha önceden çift küme dizileri için Wijsman anlamında çalışılmış asimptotik denklik kavramları arasındaki ilişkiler incelendi.

Anahtar Kelimeler: Asimptotik denklik; Wijsman anlamında yakınsaklık; Çift lacunary dizi; İnvaryant istatistiksel yakınsaklık; β ncı mertebe; Küme dizisi.

1. Introduction

Long after the concept of convergence for double sequences was introduced by Pringsheim [1], using the concepts of statistical convergence, double lacunary sequence and σ -convergence, this concept was extended to new convergence concepts for double sequences by some authors [2-4]. Recently, for double sequences, on two new convergence concepts called double almost statistical and double almost lacunary statistical convergence of order α were studied by Savaş [5, 6]. Moreover, for double sequences, the concept of asymptotical equivalence was introduced by Patterson [7].

Over the years, on the various convergence concepts for set sequences were studied by several authors. One of them, discussed in this study, is the concept of convergence in the Wijsman sense [8, 9]. Using the concepts of statistical convergence, double lacunary sequence and σ -convergence, this concept was extended to new convergence concepts for double set sequences by some authors [10-12]. In [11], Nuray and Uluslu studied on the concepts of invariant statistical and lacunary invariant statistical convergence in the Wijsman sense for double set sequences. Furthermore, for double set sequences, the concepts of asymptotical equivalence in the Wijsman sense were introduced by Nuray et al. [13] and then these concepts were studied by some authors.

In this paper, using order β , we studied on new asymptotical equivalence concepts in the Wijsman sense for double set sequences.

More information on the concepts of convergence or asymptotical equivalence for real or set sequences can be found in [14-24].

2. Preliminaries

First of all, let us recall the basic notions necessary for a better understanding of our study [3, 10, 12, 13, 25].

For a metric space (Y, d) , $\rho(y, C)$ denote the distance from y to C where

$$\rho(y, C) := \rho_y(C) = \inf_{c \in C} d(y, c),$$

for any $y \in Y$ and any non empty set $C \subseteq Y$.

For a non empty set Y , let a function $g: \mathbb{N} \rightarrow 2^Y$ (the power set of Y) is defined by $g(m) = C_m \in 2^Y$ for each $m \in \mathbb{N}$. Then the sequence $\{C_m\} = \{C_1, C_2, \dots\}$, which is the codomain elements of g , is called set sequences.

Throughout this study, (Y, d) will be considered as a metric space and C, C_{mn}, D_{mn} will be considered as any non empty closed subsets of Y .

A double set sequence $\{C_{mn}\}$ is called convergent to the set C in the Wijsman sense if each $y \in Y$,

$$\lim_{m, n \rightarrow \infty} \rho_y(C_{mn}) = \rho_y(C).$$

A double set sequence $\{C_{mn}\}$ is called statistical convergent to the set C in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{p, q \rightarrow \infty} \frac{1}{pq} |\{(m, n): m \leq p, n \leq q, |\rho_y(C_{mn}) - \rho_y(C)| \geq \xi\}| = 0.$$

A double sequence $\theta_2 = \{(j_s, k_t)\}$ is called a double lacunary sequence if there exist increasing sequences (j_s) and (k_t) of the integers such that

$$j_0 = 0, h_s = j_s - j_{s-1} \rightarrow \infty \text{ and } k_0 = 0, \bar{h}_t = k_t - k_{t-1} \rightarrow \infty \text{ as } s, t \rightarrow \infty.$$

In general, the following notations is used for any double lacunary sequence:

$$\begin{aligned} \ell_{st} &= j_s k_t, \quad h_{st} = h_s \bar{h}_t, \quad I_{st} = \{(m, n): j_{s-1} < m \leq j_s \text{ and } k_{t-1} < n \leq k_t\}, \\ q_s &= \frac{j_s}{j_{s-1}} \text{ and } q_t = \frac{k_t}{k_{t-1}}. \end{aligned}$$

Throughout this study, $\theta_2 = \{(j_s, k_t)\}$ will be considered as a double lacunary sequence.

A double set sequence $\{C_{mn}\}$ is called lacunary statistical convergent to the set C in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{s,t \rightarrow \infty} \frac{1}{h_{st}} \left| \left\{ (m, n) \in I_{st} : |\rho_y(C_{mn}) - \rho_y(C)| \geq \xi \right\} \right| = 0.$$

The term $\rho_y \left(\frac{C_{mn}}{D_{mn}} \right)$ is defined as follows:

$$\rho_y \left(\frac{C_{mn}}{D_{mn}} \right) = \begin{cases} \frac{\rho(y, C_{mn})}{\rho(y, D_{mn})} & , \quad y \notin C_{mn} \cup D_{mn} \\ \lambda & , \quad y \in C_{mn} \cup D_{mn}. \end{cases}$$

Double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are called asymptotically equivalent in the Wijsman sense if each $y \in Y$,

$$\lim_{m,n \rightarrow \infty} \rho_y \left(\frac{C_{mn}}{D_{mn}} \right) = 1$$

and denoted by $C_{mn} \overset{W}{\sim} D_{mn}$.

Let σ be a mapping such that $\sigma: \mathbb{N}^+ \rightarrow \mathbb{N}^+$ (the set of positive integers). A continuous linear functional ψ on ℓ_∞ is called an invariant mean (or a σ -mean) if it satisfies the following conditions:

1. $\psi(x_u) \geq 0$, when the sequence (x_u) has $x_u \geq 0$ for all u ,
2. $\psi(e) = 1$, where $e = (1,1,1, \dots)$ and
3. $\psi(x_{\sigma(u)}) = \psi(x_u)$ for all $(x_u) \in \ell_\infty$.

The mappings σ are assumed to be one to one and such that $\sigma^m(u) \neq u$ for all $m, u \in \mathbb{N}^+$, where $\sigma^m(u)$ denotes the m th iterate of the mapping σ at u . Thus ψ extends the limit functional on c , in the sense that $\psi(x_u) = \lim x_u$ for all $(x_u) \in c$.

Double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are called asymptotically invariant statistical equivalent to multiple λ in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{p,q \rightarrow \infty} \frac{1}{pq} \left| \left\{ (m, n) : m \leq p, n \leq q, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| = 0$$

uniformly in u, v .

The set of all asymptotically invariant statistical equivalent to multiple λ double set sequences in the Wijsman sense is denoted by $\{W_2^\lambda(S_\sigma)\}$.

Double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are called asymptotically lacunary invariant statistical equivalent to multiple λ in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{s,t \rightarrow \infty} \frac{1}{h_{st}} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| = 0$$

uniformly in u, v .

3. Main Results

In this section, for double set sequences, the concepts of asymptotical invariant statistical and asymptotical lacunary invariant statistical equivalence of order β ($0 < \beta \leq 1$) in the Wijsman sense were introduced. Also, the inclusion relations between them were investigated.

Definition 1. Double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are asymptotically invariant statistical equivalent to multiple λ of order β in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{p,q \rightarrow \infty} \frac{1}{(pq)^\beta} \left| \left\{ (m, n) : m \leq p, n \leq q, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| = 0$$

uniformly in u, v , where $0 < \beta \leq 1$ and we denote this in $C_{mn} \overset{W_2^\lambda(S_\sigma^\beta)}{\sim} D_{mn}$ format, and simply called asymptotically invariant statistical equivalent of order β in the Wijsman sense if $\lambda = 1$.

Example 1. Let $Y = \mathbb{R}^2$ and double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ be defined as following:

$$C_{mn} := \begin{cases} \left\{ (a, b) \in \mathbb{R}^2 : a^2 + (b - 1)^2 = \frac{1}{mn} \right\} & ; \text{ if } m \text{ and } n \text{ are square integers} \\ \{(1,0)\} & ; \text{ otherwise.} \end{cases}$$

and

$$D_{mn} := \begin{cases} \left\{ (a, b) \in \mathbb{R}^2 : a^2 + (b + 1)^2 = \frac{1}{mn} \right\} & ; \text{ if } m \text{ and } n \text{ are square integers} \\ \{(1,0)\} & ; \text{ otherwise.} \end{cases}$$

In this case, the double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are asymptotically invariant statistical equivalent of order β ($0 < \beta \leq 1$) in the Wijsman sense.

Remark 1. For $\beta = 1$, the concept of asymptotical invariant statistical equivalence of order β in the Wijsman sense coincides with the concept of asymptotical invariant statistical equivalence in the Wijsman sense for double set sequences in [25].

Definition 2. Double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are asymptotically lacunary invariant statistical equivalent to multiple λ of order β in the Wijsman sense if every $\xi > 0$ and each $y \in Y$,

$$\lim_{s,t \rightarrow \infty} \frac{1}{h_{st}^\beta} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| = 0$$

uniformly in u, v , where $0 < \beta \leq 1$ and we denote this in $C_{mn} \overset{W_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} D_{mn}$ format, and simply called asymptotical lacunary invariant statistical equivalent of order β in the Wijsman sense if $\lambda = 1$.

The set of all asymptotically lacunary invariant statistically equivalent double set sequences to multiple λ of order β in the Wijsman sense is denoted by $\{W_2^\lambda(S_{\sigma\theta}^\beta)\}$.

Example 2. Let $Y = \mathbb{R}^2$ and double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ be defined as following:

$$C_{mn} := \begin{cases} \{(a, b) \in \mathbb{R}^2 : (a - m)^2 + (b + n)^2 = 1\} & ; & \text{if } (m, n) \in I_{st}, \\ & & m \text{ and } n \text{ are square integer} \\ \{(-1, -1)\} & ; & \text{otherwise.} \end{cases}$$

and

$$D_{mn} := \begin{cases} \{(a, b) \in \mathbb{R}^2 : (a + m)^2 + (b - n)^2 = 1\} & ; & \text{if } (m, n) \in I_{st}, \\ & & m \text{ and } n \text{ are square integer} \\ \{(-1, -1)\} & ; & \text{otherwise.} \end{cases}$$

In this case, the double set sequences $\{C_{mn}\}$ and $\{D_{mn}\}$ are asymptotically lacunary invariant statistical equivalent of order β ($0 < \beta \leq 1$) in the Wijsman sense.

Remark 2. For $\beta = 1$, the concept of asymptotical lacunary invariant statistical equivalence of order β in the Wijsman sense coincide with the concept of asymptotical lacunary invariant statistical equivalence in the Wijsman sense for double set sequences in [25].

Theorem 1. If $\liminf_s q_s^\beta > 1$ and $\liminf_t q_t^\beta > 1$ where $0 < \beta \leq 1$, then

$$C_{mn} \overset{W_2^\lambda(S_{\sigma}^\beta)}{\sim} D_{mn} \Rightarrow C_{mn} \overset{W_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} D_{mn}.$$

Proof. Let $0 < \beta \leq 1$ and suppose that $\liminf_s q_s^\beta > 1$ and $\liminf_t q_t^\beta > 1$. Then, there exist $\eta, \mu > 0$ such that $q_s^\beta \geq 1 + \eta$ and $q_t^\beta \geq 1 + \mu$ for all s, t , which implies that

$$\frac{h_{st}}{\ell_{st}^\beta} \geq \frac{\eta\mu}{(1+\eta)(1+\mu)} \Rightarrow \frac{h_{st}^\beta}{\ell_{st}^\beta} \geq \frac{\eta^\beta \mu^\beta}{(1+\eta)^\beta (1+\mu)^\beta}$$

For every $\xi > 0$ and each $y \in Y$, we have

$$\begin{aligned} & \frac{1}{\ell_{st}^\beta} \left| \left\{ (m, n) : m \leq j_s, n \leq k_t, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ & \geq \frac{1}{\ell_{st}^\beta} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ & = \frac{h_{st}^\beta}{\ell_{st}^\beta h_{st}^\beta} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ & \geq \frac{\eta^\beta \mu^\beta}{(1+\eta)^\beta (1+\mu)^\beta} \frac{1}{h_{st}^\beta} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right|, \end{aligned}$$

for all u, v . If $C_{mn} \stackrel{w_2^\lambda(S_\sigma^\beta)}{\sim} D_{mn}$, then for each $y \in Y$ the term on the left side of the above inequality convergent to 0 and this implies that

$$\frac{1}{h_{st}^\beta} \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \rightarrow 0$$

uniformly in u, v . Thus, we get $C_{mn} \stackrel{w_2^\lambda(S_{\sigma\theta}^\eta)}{\sim} D_{mn}$.

Theorem 2. If $\limsup_s q_s < \infty$ and $\limsup_t q_t < \infty$, then

$$C_{mn} \stackrel{w_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} D_{mn} \Rightarrow C_{mn} \stackrel{w_2^\lambda(S_\sigma^\beta)}{\sim} D_{mn},$$

where $0 < \beta \leq 1$.

Proof. Let $\limsup_s q_s < \infty$ and $\limsup_t q_t < \infty$. Then, there exist $M, N > 0$ such that $q_s < M$ and $q_t < N$ for all s, t . Also, we suppose that $C_{mn} \stackrel{w_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} D_{mn}$ (where $0 < \beta \leq 1$) and $\xi > 0$, and let

$$\kappa_{st} := \left| \left\{ (m, n) \in I_{st} : \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right|.$$

Then, there exist $s_0, t_0 \in \mathbb{N}$ such that for every $\xi > 0$, each $y \in Y$ and all $s \geq s_0, t \geq t_0$

$$\frac{\kappa_{st}}{h_{st}^\beta} < \xi,$$

for all u, v . Now, let

$$\gamma := \max\{\kappa_{st}: 1 \leq s \leq s_0, 1 \leq t \leq t_0\},$$

and let p and q be any integers satisfying $j_{s-1} < p \leq j_s$ and $k_{t-1} < q \leq k_t$. Then, for each $y \in Y$ we have

$$\begin{aligned} & \frac{1}{(pq)^\beta} \left| \left\{ (m, n): m \leq p, n \leq q, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ & \leq \frac{1}{\ell_{(s-1)(t-1)}^\beta} \left| \left\{ (m, n): m \leq j_s, n \leq k_t, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ & = \frac{1}{\ell_{(s-1)(t-1)}^\beta} \{ \kappa_{11} + \kappa_{12} + \kappa_{21} + \kappa_{22} + \dots + \kappa_{s_0 t_0} + \dots + \kappa_{st} \} \\ & \leq \frac{s_0 t_0}{\ell_{(s-1)(t-1)}^\beta} \left(\max_{\substack{1 \leq m \leq s_0 \\ 1 \leq n \leq t_0}} \{ \kappa_{mn} \} \right) + \frac{1}{\ell_{(s-1)(t-1)}^\beta} \left\{ h_{s_0(t_0+1)}^\beta \frac{\kappa_{s_0(t_0+1)}}{h_{s_0(t_0+1)}^\beta} \right. \\ & \quad \left. + h_{(s_0+1)t_0}^\beta \frac{\kappa_{(s_0+1)t_0}}{h_{(s_0+1)t_0}^\beta} + h_{(s_0+1)(t_0+1)}^\beta \frac{\kappa_{(s_0+1)(t_0+1)}}{h_{(s_0+1)(t_0+1)}^\beta} + \dots + h_{st}^\beta \frac{\kappa_{st}}{h_{st}^\beta} \right\} \\ & \leq \frac{s_0 t_0 \gamma}{\ell_{(s-1)(t-1)}^\beta} + \frac{1}{\ell_{(s-1)(t-1)}^\beta} \left(\sup_{\substack{s > s_0 \\ t > t_0}} \frac{\kappa_{st}}{h_{st}^\beta} \right) \left(\sum_{m, n \geq s_0, t_0}^{s, t} h_{mn}^\beta \right) \\ & \leq \frac{s_0 t_0 \gamma}{\ell_{(s-1)(t-1)}^\beta} + \frac{1}{\ell_{(s-1)(t-1)}^\beta} \left(\sup_{\substack{s > s_0 \\ t > t_0}} \frac{\kappa_{st}}{h_{st}^\beta} \right) \left(\sum_{m, n \geq s_0, t_0}^{s, t} h_{mn} \right) \\ & \leq \frac{s_0 t_0 \gamma}{\ell_{(s-1)(t-1)}^\beta} + \xi \frac{(j_s - j_{s_0})(k_t - k_{t_0})}{\ell_{(s-1)(t-1)}} \\ & \leq \frac{s_0 t_0 \gamma}{\ell_{(s-1)(t-1)}^\beta} + \xi q_s q_t \\ & \leq \frac{s_0 t_0 \gamma}{\ell_{(s-1)(t-1)}^\beta} + \xi MN, \end{aligned}$$

for all u, v . Since $j_{s-1}, k_{t-1} \rightarrow \infty$ as $p, q \rightarrow \infty$, it follows that for each $y \in Y$

$$\frac{1}{(pq)^\beta} \left| \left\{ (m, n): m \leq p, n \leq q, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \rightarrow 0$$

uniformly in u, v . Thus, we get $C_{mn} \underset{w_2^\lambda(s_\sigma^\beta)}{\sim} D_{mn}$.

Theorem 3. If

$$1 < \liminf_s q_s^\beta \leq \limsup_s q_s < \infty \text{ and } 1 < \liminf_t q_t^\beta \leq \limsup_t q_t < \infty,$$

where $0 < \beta \leq 1$, then

$$C_{mn} \overset{W_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} D_{mn} \Leftrightarrow C_{mn} \overset{W_2^\lambda(S_\sigma^\beta)}{\sim} D_{mn}.$$

Proof. This can be obtained from Theorem 1 and Theorem 2, immediately.

Theorem 4. If $\liminf_{s,t \rightarrow \infty} \frac{h_{st}^\beta}{\ell_{st}} > 0$ where $0 < \beta \leq 1$, then $\{W_2^\lambda(S_\sigma)\} \subseteq \{W_2^\lambda(S_{\sigma\theta}^\beta)\}$.

Proof. For every $\xi > 0$ and each $y \in Y$, it is obvious that

$$\begin{aligned} \left\{ (m, n): m \leq j_s, n \leq k_t, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \\ \supset \left\{ (m, n) \in I_{st}: \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \frac{1}{\ell_{st}} \left| \left\{ (m, n): m \leq j_s, n \leq k_t, \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ \geq \frac{1}{\ell_{st}} \left| \left\{ (m, n) \in I_{st}: \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \\ = \frac{h_{st}^\beta}{\ell_{st} h_{st}^\beta} \left| \left\{ (m, n) \in I_{st}: \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right|, \end{aligned}$$

for all u, v . If $C_{mn} \overset{W_2^\lambda(S_\sigma)}{\sim} C$, then for each $y \in Y$ the term on the left side of the above inequality convergent to 0 and this implies that

$$\frac{1}{h_{st}^\beta} \left| \left\{ (m, n) \in I_{st}: \left| \rho_y \left(\frac{C_{\sigma^m(u)\sigma^n(v)}}{D_{\sigma^m(u)\sigma^n(v)}} \right) - \lambda \right| \geq \xi \right\} \right| \rightarrow 0$$

uniformly in u, v . Thus, we get $C_{mn} \overset{W_2^\lambda(S_{\sigma\theta}^\beta)}{\sim} C$. Consequently, $\{W_2^\lambda(S_\sigma)\} \subseteq \{W_2^\lambda(S_{\sigma\theta}^\beta)\}$.

References

- [1] Pringsheim, A., *Zur theorie der zweifach unendlichen Zahlenfolgen*, *Mathematische Annalen*, 53, 289-321, 1900.
- [2] Mursaleen, M., Edely, O.H.H., *Statistical convergence of double sequences*, *Journal of Mathematical Analysis and Applications*, 288, 223-231, 2003.
- [3] Patterson, R.F., Savaş, E., *Lacunary statistical convergence of double sequences*, *Mathematical Communications*, 10, 55-61, 2005.
- [4] Savaş, E., Patterson, R.F., *Double σ -convergence lacunary statistical sequences*, *Journal of Computational Analysis and Applications*, 11, 610-615, 2009.

- [5] Savaş, E., *Double almost statistical convergence of order α* , *Advances in Difference Equations*, 62, 1-9, 2013.
- [6] Savaş, E., *Double almost lacunary statistical convergence of order α* , *Advances in Difference Equations*, 254, 1-10, 2013.
- [7] Patterson, R.F., *Rates of convergence for double sequences*, *Southeast Asian Bulletin of Mathematics*, 26, 469-478, 2003.
- [8] Baronti, M., Papini, P., *Convergence of sequences of sets, in: methods of functional analysis in approximation theory*, Birkhäuser, Basel, 133-155, 1986.
- [9] Beer, G., *Wijsman convergence: A survey*, *Set-Valued Analysis*, 2, 77-94, 1994.
- [10] Nuray, F., Ulus, U., Dündar, E., *Lacunary statistical convergence of double sequences of sets*, *Soft Computing*, 20, 2883-2888, 2016.
- [11] Nuray, F., Ulus, U., *Lacunary invariant statistical convergence of double sequences of sets*, *Creative Mathematics and Informatics*, 28, 143-150, 2019.
- [12] Nuray, F., Dündar, E., Ulus, U., *Wijsman statistical convergence of double sequences of sets*, *Iranian Journal of Mathematical Sciences and Informatics*, 16, 55-64, 2021.
- [13] Nuray, F., Patterson, R.F., Dündar, E., *Asymptotically lacunary statistical equivalence of double sequences of sets*, *Demonstratio Mathematica*, 49, 183-196, 2016.
- [14] Çolak, R., *Statistical convergence of order α* , In: *Modern Methods in Analysis and Its Applications*, Anamaya Publishers, New Delhi, 121-129, 2010.
- [15] Gülle, E., *Double Wijsman asymptotically statistical equivalence of order α* , *Journal of Intelligent & Fuzzy Systems*, 38, 2081-2087, 2020.
- [16] Pancaroğlu, N., Nuray, F., *On invariant statistically convergence and lacunary invariant statistical convergence of sequences of sets*, *Progress in Applied Mathematics*, 5, 23-29, 2013.
- [17] Pancaroğlu, N., Nuray, F., Savaş, E., *On asymptotically lacunary invariant statistical equivalent set sequences*, *AIP Conference Proceedings*, 1558, 780-781, 2013.
- [18] Savaş, E., Nuray, F., *On σ -statistically convergence and lacunary σ -statistically convergence*, *Mathematica Slovaca*, 43, 309-315, 1993.
- [19] Savaş, E., *Asymptotically I -lacunary statistical equivalent of order α for sequences of sets*, *Journal of Nonlinear Science and its Applications*, 10, 2860-2867, 2017.
- [20] Şengül, H., Et, M., *On lacunary statistical convergence of order α* , *Acta Mathematica Scientia. Series B*, 34, 473-482, 2014.
- [21] Şengül, H., *On Wijsman I -lacunary statistical equivalence of order (η, μ)* , *Journal of Inequalities and Special Functions*, 9, 92-101, 2018.
- [22] Ulus, U., Nuray, F., *On asymptotically lacunary statistical equivalent set sequences*, *Journal of Mathematics*, 310438, 1-5, 2013.
- [23] Ulus, U., Gülle, E., *Some statistical convergence types of order α for double set sequences*, *Facta Universitatis, Series: Mathematics and Informatics*, 35, 595-603, 2020.
- [24] Ulus, U., Nuray, F., *Lacunary I -invariant convergence*, *Cumhuriyet Science Journal*, 41, 617-624, 2020.
- [25] Ulus, U., Dündar, E., Pancaroğlu Akın, N., *Lacunary invariant statistical equivalence for double set sequences*, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, (Accepted).



An Educational Higgs Study with CMS Open Data

Bora AKGUN^{1,2,*}

¹*Boğaziçi University, Department of Physics, Istanbul, Türkiye*

²*Feza Gürsey Center for Physics and Mathematics, Istanbul, Türkiye*

bora.akgun@boun.edu.tr, ORCID: 0000-0001-8888-3562

Received: 01.10.2021

Accepted: 03.12.2021

Published: 30.12.2021

Abstract

The Compact Muon Solenoid (CMS) Experiment is one of the particle detectors at CERN's Large Hadron Collider (LHC). The CMS Collaboration consists of more than 5000 scientists, engineers, technicians and students from more than 200 institutes and universities from more than 40 countries. The CMS collaboration has a wide physics program and published more than 1000 papers. The CMS has provided open access to 2 PB of its data recorded in proton-proton collisions for research and educational use. This paper gives a brief description on how to analyze the CMS open data and shows a simple Higgs Boson analysis with the data recorded in early 2011.

Keywords: CMS Open Data; Virtual Machine; Higgs.

CMS Açık Veri ile Eğitici Higgs Çalışması

Öz

Kompakt Müon Solenoid (CMS) Deneyi, CERN'in Büyük Hadron Çarpıştırıcısındaki (LHC) parçacık dedektörlerinden biridir. CMS İşbirliği, 40'tan fazla ülkeden 200'den fazla enstitü ve üniversiteden 5000'den fazla bilim insanı, mühendis, teknisyen ve öğrenciden oluşmaktadır. CMS İşbirliği geniş bir fizik programına sahiptir ve 1000'den fazla makale yayınlamıştır. CMS, araştırma ve eğitim amaçlı kullanım için proton-proton çarpışmalarında kaydedilen verilerinin 2 PB'sine açık erişim sağlamıştır. Bu makale, CMS açık verilerinin nasıl



analiz edileceğine dair kısa bir açıklama sunar ve 2011'in başlarında kaydedilen verilerle basit bir Higgs Boson analizini içerir.

Anahtar Kelimeler: CMS Açık Veri; Sanal Makine; Higgs.

1. Introduction

The Compact Muon Solenoid (CMS) [1] is a general-purpose detector at the Large Hadron Collider (LHC) at CERN. The CMS has a broad physics programme ranging from studying the Standard Model to searching for extra dimensions and particles that could make up dark matter. The CMS collaboration has published more than one thousand papers, one of the most important publications was on the observation of the Higgs Boson [2] in 2012. The CMS Collaboration continues its efforts on studying the properties of the Higgs Boson [3-38] ever since the first observation.

Since 2014 the CMS Collaboration started to release recorded data into open datasets. The CMS data are unique and are of interest to the scientific community as well as to those in education. Several papers [39-47] focusing on new methods and physics results have been published with the CMS open data.

This paper has educational intentions and focuses on CMS Open Data from 2011 (7 TeV) for a study on the Higgs Boson. Section 2 describes how to set up a virtual machine, the datasets and the event selection, Section 3 displays the event topology of the Higgs Boson candidates and Section 4 shows the results.

2. The Datasets and Event Selection

One of the decay channels of the Higgs Boson is $H \rightarrow ZZ \rightarrow 4l$, where the final decay products can be four muons, four electrons or two muons and two electrons. For this study, DoubleMuon [48] and DoubleElectron [49] datasets from RunA of 2011 were used since the events stored in these datasets have at least two high-energy muons or electrons.

A virtual machineⁱ is installed and the CMS computing environment is set upⁱⁱ to analyze the CMS Open Data. The Higgs Boson candidates were selected by requiring to have two Z boson candidates in the event. The Z boson candidates selection relied on the presence of a pair of muons or electrons. The muon candidates are selected from muons identified as both Global and Tracker Muon with the $\chi^2/\text{d.o.f}$ of the global-muon track fit is less than 10 and the number of inner-tracker hits are more than 10. The muon candidates are also required to have transverse momentum (p_T) higher than 5 GeV/c, absolute value of the pseudo-rapidity ($|\eta|$) less than 2.4 and the ratio of the

scalar sum of the P_T of tracks in the inner tracker (ΣP_T) and the transverse energies (ΣE_T) in calorimeter cells (both in the Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL)) within a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ ⁱⁱⁱ centered on the direction vector of the candidate particle ($dR_{0.3}$) and the P_T of the muon candidate less than 0.15. The electron candidates are selected from electrons passing loose identification requirements. The electron candidates are also required to have P_T higher than 7 GeV/c, $|\eta|$ less than 2.5 and additional isolation criteria: in the barrel region ($|\eta| < 1.44$) the ratio of the ΣP_T within $dR_{0.3}$ and the P_T of the electron candidate less than 0.09, the ratio of the ΣE_T in the ECAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.07 and the ratio of the ΣE_T in the HCAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.10, in the forward region ($|\eta| > 1.57$) the ratio of the ΣP_T within $dR_{0.3}$ and the P_T of the electron candidate less than 0.04, the ratio of the ΣE_T in the ECAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.05 and the ratio of the ΣE_T in the HCAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.0025. The region between 1.44 and 1.57 is excluded for the electrons due to low electrons reconstruction efficiency. Events with single Z boson candidate are also studied.

3. Event Displays

The CMS collaboration has developed a tool for visualizing particle collisions. Figure 1 shows event displays of Higgs Boson candidates from 2011 and 2012 decaying to four muons, four electrons or two muons and two electrons. The muon tracks and the muon stations with a muon hit are shown in red, the electron tracks and the energy deposited in ECAL cells are shown in green, and the energy deposited in the HCAL cells are shown in blue in the event displays.

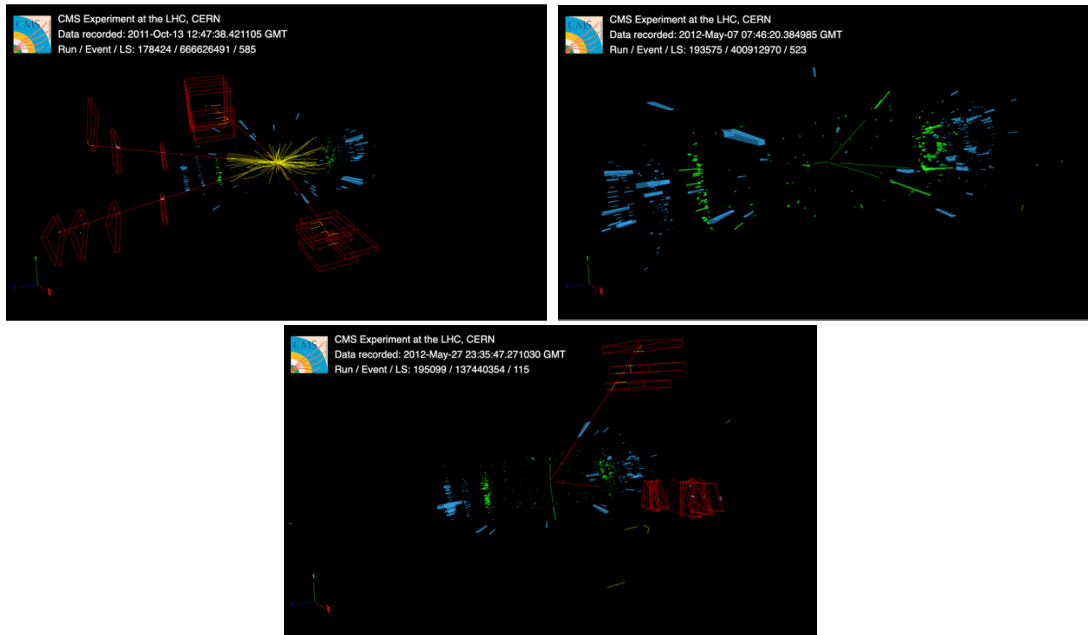


Figure 1: The event displays of the Higgs Boson candidates. The up-left display shows four muons, the up-right display shows four electrons, and the low-middle display shows two muons and two electrons. The particle tracks in the CMS tracker are shown in yellow in the upper-left display, for the other two displays visualization of the particle tracks are disabled to make the electron tracks (in green) visible to the reader

4. Results

One million events per dataset (DoubleMu and DoubleElectron) were analyzed for Z candidates. Figure 2 shows the kinematic properties of the muons used for Z candidate reconstruction. Figure 3 shows the kinematic properties and the mass of the Z candidates reconstructed by the muon pairs. Figure 4 shows the kinematic properties of the electrons used for Z candidate reconstruction. Figure 5 shows the kinematic properties and the mass of the Z candidates reconstructed by the electron pairs. Both the muons and electrons are mostly in the barrel region and their transverse momenta peaks at around a bit higher than $40 \text{ GeV}/c$. The Z candidates are mostly in the forward region. The reconstructed mass of the Z candidates centered at $91 \text{ GeV}/c^2$. The mass of the Z candidates have a narrower mass peak for the muon case than the electron case. This suggests that the muon resolution of the CMS detector is better than its electron resolution. The details of the muon reconstruction performance can be found at [50].

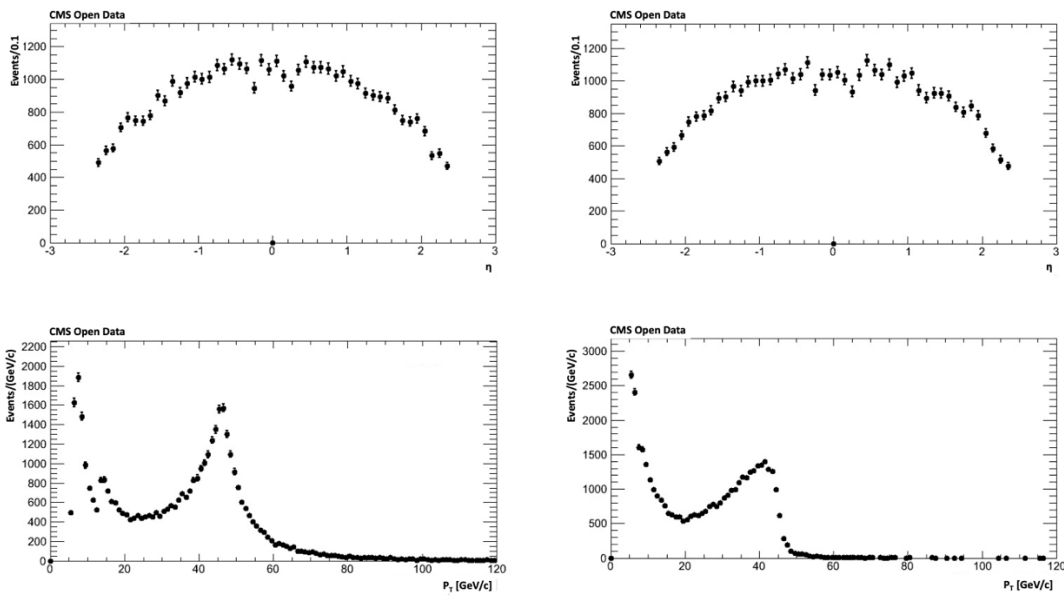


Figure 2: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the more energetic muon. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the less energetic muon

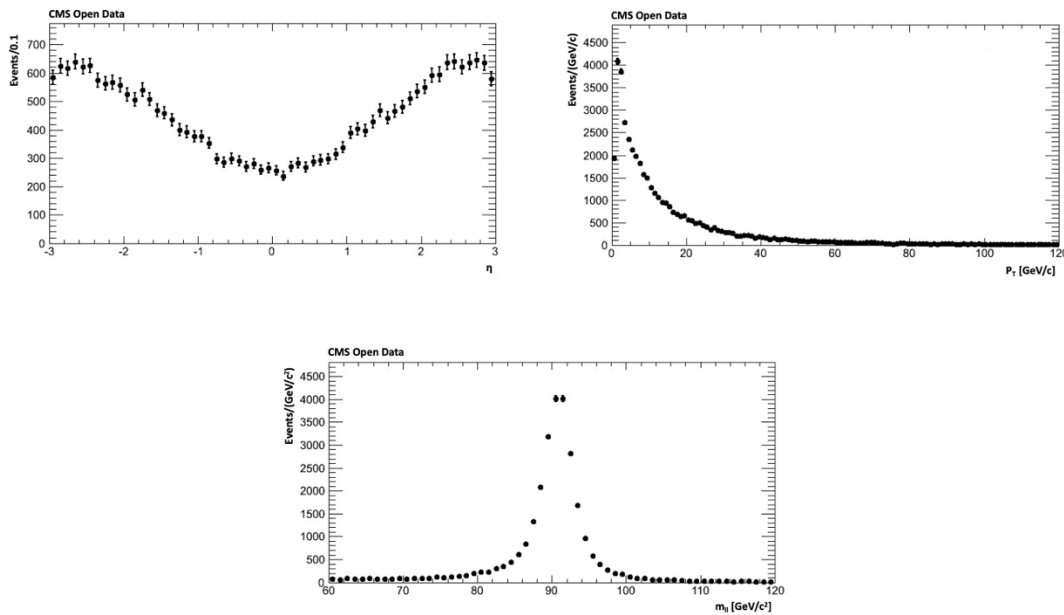


Figure 3: The pseudo-rapidity (up-left) and transverse momentum (up-right) distributions of the Z candidates reconstructed by the muon pairs. The mass distribution (low-middle) of the Z candidates reconstructed by the muon pairs

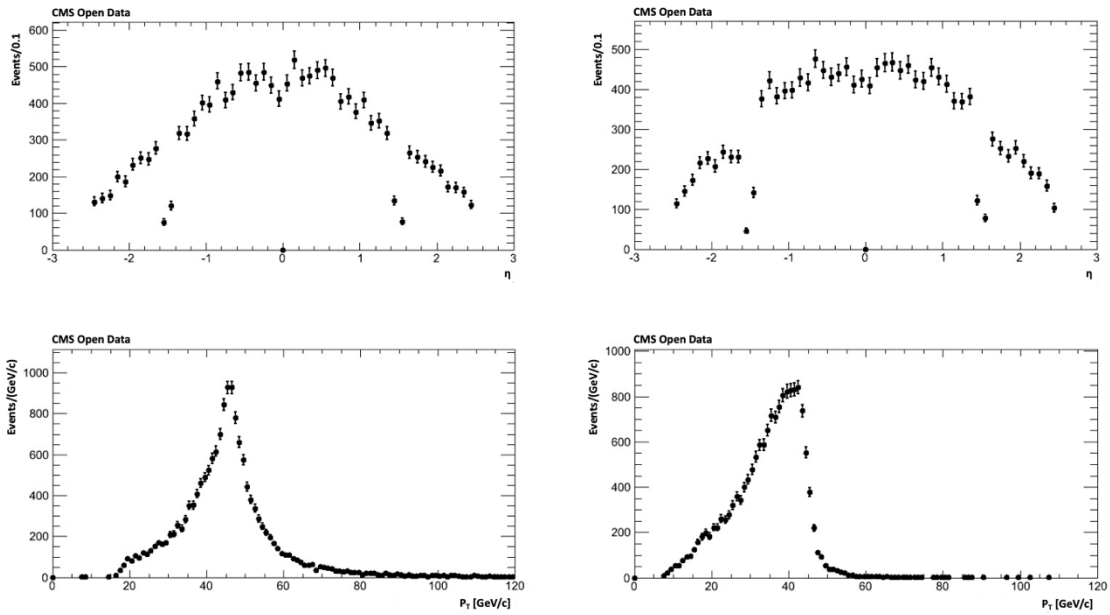


Figure 4: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the more energetic electron. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the less energetic electron

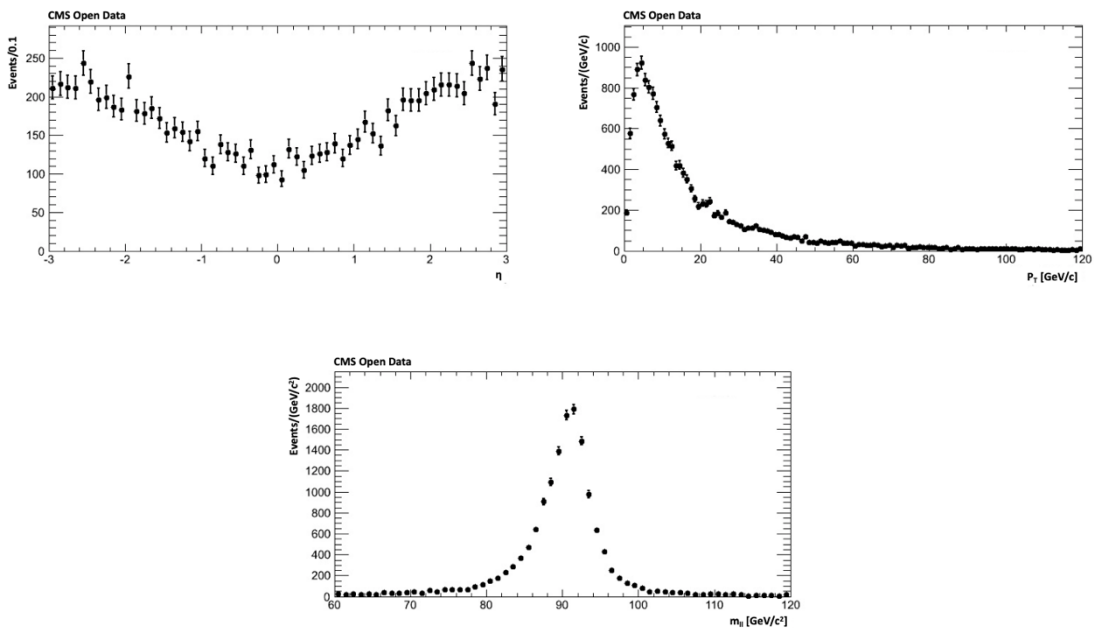


Figure 5: The pseudo-rapidity (up-left) and transverse momentum (up-right) distributions of the Z candidates reconstructed by the electron pairs. The mass distribution (low-middle) of the Z candidates reconstructed by the electron pairs.

All events were analyzed for Higgs Boson candidates; 35,329,695 events in DoubleMu dataset and 3,835,254 events in DoubleElectron dataset. Figure 6 shows the kinematic properties of the Z candidates used for Higgs Boson reconstruction by using the DoubleMu dataset. Figure

7 shows the kinematic properties and the mass of the Higgs Boson candidates by using the DoubleMu dataset. Figure 8 shows the kinematic properties of the Z candidates used for Higgs Boson reconstruction by using the DoubleElectron dataset. Figure 9 shows the kinematic properties and the mass of the Higgs Boson candidates by using the DoubleElectron dataset. The Z candidates are evenly distributed in pseudo-rapidity for both datasets. The Higgs candidates are mostly in the forward region for both datasets. Even though it is not possible to make a conclusion, the mass of the Higgs Boson candidates with the DoubleMu dataset shows a slight excess in the 120-130 GeV/c² range. A similar excess doesn't exist in the Higgs Boson candidate mass plot with the DoubleElectron dataset.

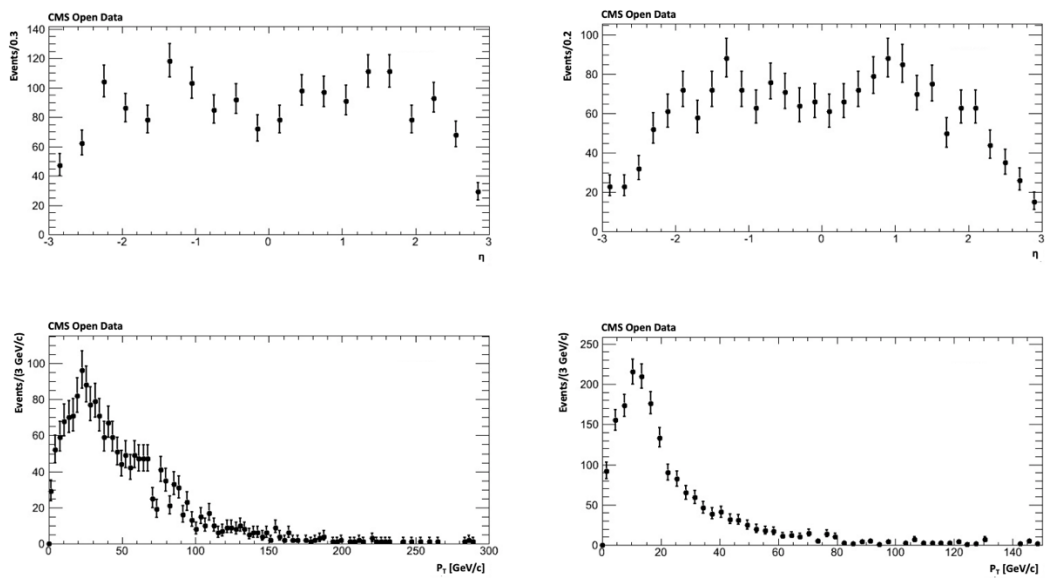


Figure 6: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the real Z candidates. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the virtual Z candidates. The plots are produced by using the DoubleMu dataset

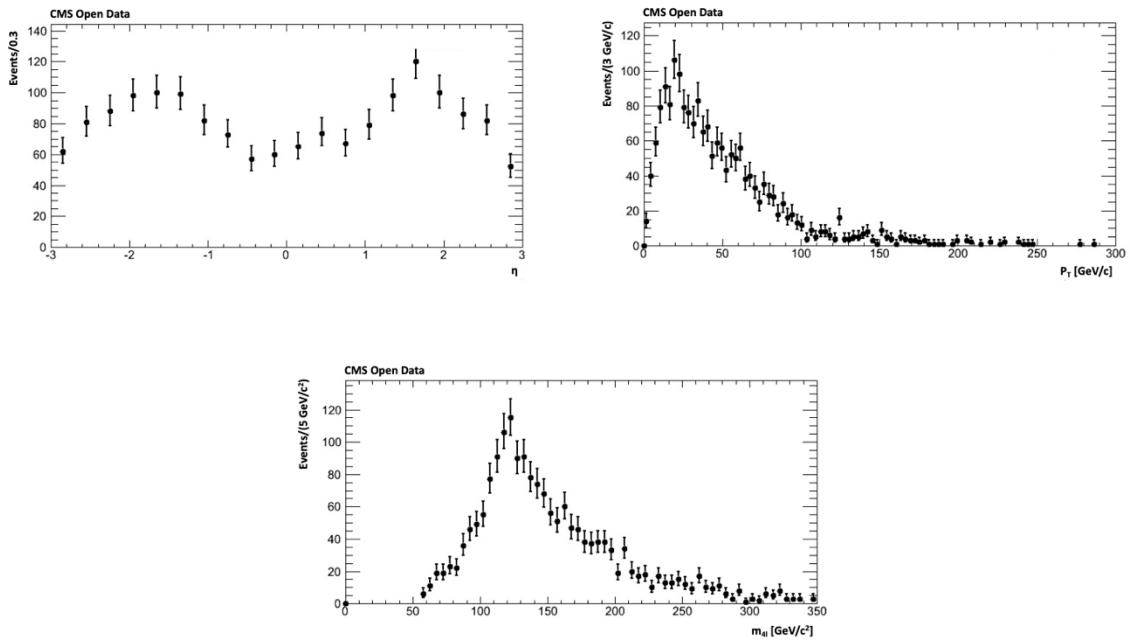


Figure 7: The pseudo-rapidity (up-left), transverse momentum (up-right) and the mass distribution (low-middle) of the Higgs candidates reconstructed by using the DoubleMu dataset

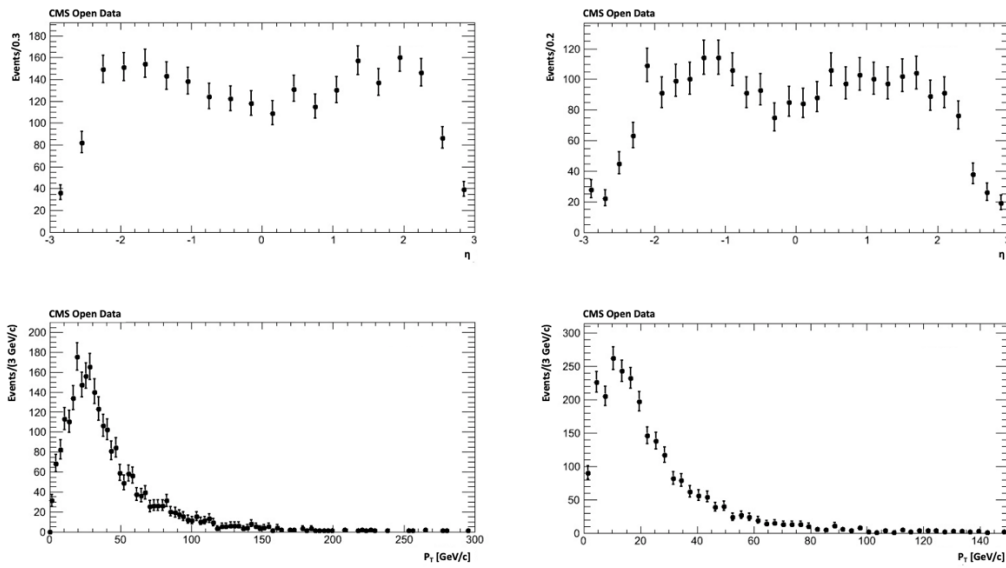


Figure 8: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the real Z candidates. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the virtual Z candidates. The plots are produced by using the DoubleElectron dataset

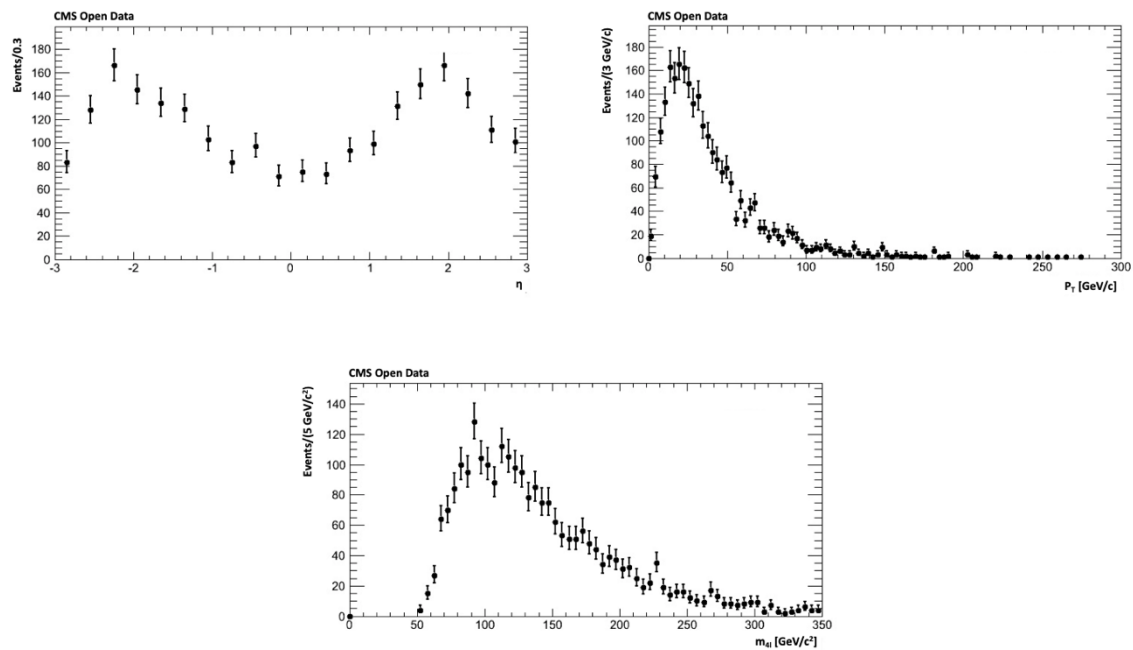


Figure 9: The pseudo-rapidity (up-left), transverse momentum (up-right) and the mass distribution (low-middle) of the Higgs candidates reconstructed by using the DoubleElectron dataset

5. Conclusion

The CMS Open Data collections are released for research and educational use. The data can be analyzed by installing a virtual machine and setting up the CMS computing environment. It is possible to visualize the collision data and make histograms. This paper aims to give pointers to guidelines on how to analyze the data and presents a simple Higgs Boson analysis with the data recorded by CMS experiment in early 2011.

References

- [1] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 S08004, 2008.
- [2] The CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP 06, 081 (2013), doi:10.1007/JHEP06(2013)081.
- [3] The CMS Collaboration, *Measurement of higgs boson production and properties in the ww decay channel with leptonic final states*, JHEP 01, 096 (2014), doi:10.1007/JHEP01(2014)096.
- [4] The CMS Collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, Phys. Rev. D 89, no.9, 092007 (2014), doi:10.1103/PhysRevD.89.092007.

- [5] The CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, *JHEP* 05, 104 (2014), doi:10.1007/JHEP05(2014)104.
- [6] The CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, *Nature Phys.* 10, 557-560 (2014), doi:10.1038/nphys3005.
- [7] The CMS Collaboration, *Observation of the diphoton decay of the Higgs boson and measurement of its properties*, *Eur. Phys. J. C* 74, no.10, 3076 (2014), doi:10.1140/epjc/s10052-014-3076-z.
- [8] The CMS Collaboration, *Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the atlas and cms experiments*, *Phys. Rev. Lett.* 114, 191803 (2015), doi:10.1103/PhysRevLett.114.191803.
- [9] The CMS Collaboration, *Limits on the Higgs boson lifetime and width from its decay to four charged leptons*, *Phys. Rev. D* 92, no.7, 072010 (2015), doi:10.1103/PhysRevD.92.072010.
- [10] The CMS Collaboration, *Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J. C* 76, no.1, 13 (2016), doi:10.1140/epjc/s10052-015-3853-3.
- [11] The CMS Collaboration, *Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* 04, 005 (2016), doi:10.1007/JHEP04(2016)005.
- [12] The CMS Collaboration, *Measurement of the transverse momentum spectrum of the Higgs boson produced in pp collisions at $s = \sqrt{8}$ TeV using $H \rightarrow WW$ decays*, *JHEP* 03, 032 (2017), doi:10.1007/JHEP03(2017)032.
- [13] The CMS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* 08, 045 (2016), doi:10.1007/JHEP08(2016)045.
- [14] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 11, 047 (2017), doi:10.1007/JHEP11(2017)047.
- [15] The CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*, *Phys. Lett. B* 779, 283-316 (2018), doi:10.1016/j.physletb.2018.02.004.
- [16] The CMS Collaboration, *Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair*, *Phys. Rev. Lett.* 120, no.7, 071802 (2018), doi:10.1103/PhysRevLett.120.071802.
- [17] The CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark-antiquark pair*, *Phys. Lett. B* 780, 501-532 (2018), doi:10.1016/j.physletb.2018.02.050.
- [18] The CMS Collaboration, *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV*, *JHEP* 08, 066 (2018), doi:10.1007/JHEP08(2018)066.

[19] The CMS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 11, 185 (2018), doi:10.1007/JHEP11(2018)185.

[20] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* 791, 96 (2019), doi:10.1016/j.physletb.2018.12.073.

[21] The CMS Collaboration, *Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 01, 183 (2019), doi:10.1007/JHEP01(2019)183.

[22] The CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*, *Phys. Rev. Lett.* 121, no.12, 121801 (2018), doi:10.1103/PhysRevLett.121.121801.

[23] The CMS Collaboration, *Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* 79, no.5, 421 (2019), doi:10.1140/epjc/s10052-019-6909-y.

[24] The CMS Collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* 122, no.12, 121803 (2019), doi:10.1103/PhysRevLett.122.121803.

[25] The CMS Collaboration, *Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* 792, 369-396 (2019), doi:10.1016/j.physletb.2019.03.059.

[26] The CMS Collaboration, *Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state*, *Phys. Rev. D* 99, no.11, 112003 (2019), doi:10.1103/PhysRevD.99.112003.

[27] The CMS Collaboration, *Constraints on anomalous HVV couplings from the production of Higgs bosons decaying to τ lepton pairs*, *Phys. Rev. D* 100, no.11, 112002 (2019), doi:10.1103/PhysRevD.100.112002.

[28] The CMS Collaboration, *Search for Higgs and Z boson decays to J/ψ or Y pairs in the four-muon final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* 797, 134811 (2019), doi:10.1016/j.physletb.2019.134811.

[29] The CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, *Phys. Lett. B* 805, 135425 (2020), doi:10.1016/j.physletb.2020.135425.

[30] The CMS Collaboration, *Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel*, *Phys. Rev. Lett.* 125, no.6, 061801 (2020), doi:10.1103/PhysRevLett.125.061801.

[31] The CMS Collaboration, *Search for resonant pair production of Higgs bosons in the $bbZZ$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* 102, no.3, 032003 (2020), doi:10.1103/PhysRevD.102.032003.

[32] The CMS Collaboration, *Inclusive search for highly boosted Higgs bosons decaying to bottom-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 12, 085 (2020), doi:10.1007/JHEP12(2020)085.

[33] The CMS Collaboration, *Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at $\sqrt{s} = 13$ TeV*, JHEP 03, 003 (2021), doi:10.1007/JHEP03(2021)003.

[34] The CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons*, JHEP 01, 148 (2021), doi:10.1007/JHEP01(2021)148.

[35] The CMS Collaboration, *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C 81, no.4, 378 (2021), doi:10.1140/epjc/s10052-021-09014-x.

[36] The CMS Collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP 03, 257 (2021), doi:10.1007/JHEP03(2021)257.

[37] The CMS Collaboration, *Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C 81, no.6, 488 (2021), doi:10.1140/epjc/s10052-021-09200-x.

[38] The CMS Collaboration, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV*, JHEP 07, 027 (2021), doi:10.1007/JHEP07(2021)027.

[39] Larkoski, A., Marzani, S., Thaler, J., Tripathy, A. and Xue, W., *Exposing the QCD Splitting Function with CMS Open Data*, Phys. Rev. Lett. 119 (2017) no.13, 132003, doi:10.1103/PhysRevLett.119.132003.

[40] Tripathy, A., Xue, W., Larkoski, A., Marzani, S. and Thaler, J., *Jet substructure studies with CMS Open Data*, Phys. Rev. D 96 (2017) no.7, 074003, doi:10.1103/PhysRevD.96.074003.

[41] Andrews, M., Paulini, M., Gleyzer, S. and Poczos, B., *End-to-end physics event classification with CMS Open Data: applying image-based deep learning to detector data for the direct classification of collision events at the LHC*, Comput. Softw. Big Sci. 4 (2020) no.1, 6, doi:10.1007/s41781-020-00038-8.

[42] Cesarotti, C., Soreq, Y., Strassler, M.J., Thaler, J., and Xue, W., *Searching in CMS open data for dimuon resonances with substantial transverse momentum*, Phys. Rev. D 100 (2019) no.1, 015021, doi:10.1103/PhysRevD.100.015021.

[43] Andrews, M., Alison, J., An, S., Bryant, P., Burkle, B., Gleyzer, S., Narain, M., Paulini, M., Poczos, B., and Usai, E., *End-to-end jet classification of quarks and gluons with the CMS Open Data*, Nucl. Instrum. Meth. A 977 (2020), 164304, doi:10.1016/j.nima.2020.164304.

[44] Apyan, A., Cuozzo, W., Klute, M., Saito, Y., Schott, M., and Sintayehu, B., *Opportunities and challenges of Standard Model production cross section measurements in proton-proton collisions at $\sqrt{s} = 8$ TeV using CMS Open Data*, JINST 15 (2020) no.01, P01009, doi:10.1088/1748-0221/15/01/P01009.

[45] Paktinat Mehdiabadi, S., and Fahim, A., *Explicit jet veto as a tool to purify the underlying event in the Drell-Yan process using CMS Open Data*, J. Phys. G 46 (2019) no.9, 095003, doi:10.1088/1361-6471/ab33a9.

[46] Komiske, P.T., Mastandrea, R., Metodiev, E. M., Naik, P. and Thaler, J., *Exploring the Space of Jets with CMS Open Data*, Phys. Rev. D 101 (2020) no.3, 034009, doi:10.1103/PhysRevD.101.034009.

[47] Knapp, O., Cerri, O., Dissertori, G., Nguyen, T.Q., Pierini, M. and Vlimant, J.R., *Adversarially learned anomaly detection on CMS Open Data: re-discovering the top quark*, Eur. Phys. J. Plus 136 (2021) no.2, 236, doi:10.1140/epjp/s13360-021-01109-4.

[48] The CMS Collaboration, *DoubleMu primary dataset in AOD format from RunA of 2011 (/DoubleMu/Run2011A-12Oct2013-v1/AOD)*. CERN Open Data Portal, (2016). doi:10.7483/OPENDATA.CMS.RZ34.QR6N.

[49] The CMS Collaboration, *DoubleElectron primary dataset in AOD format from RunA of 2011 (/DoubleElectron/Run2011A-12Oct2013-v1/AOD)*, CERN Open Data Portal, (2016). doi:10.7483/OPENDATA.CMS.MQXP.QNMB.

[50] The CMS Collaboration, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, JINST 7 (2012) P10002, doi:10.1088/1748-0221/7/10/P10002.

ⁱ VirtualBox version 6.1.18 and virtual machine image CMS-OpenData-1.5.3.ova are used. The links for the VirtualBox and the image can be found at: https://www.virtualbox.org/wiki/Download_Old_Builds_6_1 and <https://opendata.cern.ch/record/252>. The recipe can be found at: <https://github.com/cms-opendata-analyses/OutreachExercise2011>.

ⁱⁱ The recipe can be found at: <https://github.com/cms-opendata-analyses/OutreachExercise2011>.

ⁱⁱⁱ <http://opendata.cern.ch/visualise/events/cms>.



Exchange Bias Effect in NiMnSbB Ferromagnetic Shape Memory Alloys Depending on Mn Content

Gökhan KIRAT^{1,*}

¹*Inonu University, Scientific and Technological Research Center, 44280 Malatya, Türkiye,
gokhan.kirat@inonu.edu.tr, ORCID: 0000-0001-7357-2921*

Received: 25.08.2021

Accepted: 03.12.2021

Published: 31.12.2021

Abstract

In this study, exchange bias effect was investigated in boron added NiMnSb Heusler alloys. The samples were fabricated by arc melting method with $\text{Ni}_{49}\text{Mn}_{37.5}\text{Sb}_{13.5}+\text{B}_x$, $\text{Ni}_{49}\text{Mn}_{37.7}\text{Sb}_{13.3}+\text{B}_x$ and $\text{Ni}_{49}\text{Mn}_{37.9}\text{Sb}_{13.1}+\text{B}_x$ nominal compositions. XRD analyzes showed that the samples were in $L2_1$ crystal structure at room temperature. While the martensitic phase transition temperatures increased, the magnetic moment value decreased with the increase in Mn content. The decrease in magnetization is resulted from increase in antiferromagnetic interactions with the increase of Mn content. Under low field and zero field cooling M-T measurements revealed that antiferromagnetic interactions were more dominant at low temperature but ferromagnetic behavior was more effective in the system at high temperature region. The coexistence of ferromagnetism and antiferromagnetism in a magnetic material can cause the exchange bias effect. Therefore, the hysteresis loops were examined at 5 K of the samples cooled under 50 kOe magnetic field. The shift at the origin of the magnetization curve, which is attributed to the exchange bias effect, increased with increasing Mn amount.

Keywords: Ferromagnetic shape memory effect; NiMnSb Heusler alloys; Exchange bias effect, Antiferromagnetic interactions.



Bor Eklenmiş NiMnSb Ferromanyetik Şekil Hatırlamalı Alaşımlarda Exchange Bias Etkisi

Öz

Bu çalışmada bor eklenmiş NiMnSb Heusler alaşımlarında exchange bias etkisi araştırılmıştır. Numuneler, $Ni_{49}Mn_{37.5}Sb_{13.5}+B_x$, $Ni_{49}Mn_{37.7}Sb_{13.3}+B_x$ ve $Ni_{49}Mn_{37.9}Sb_{13.1}+B_x$ başlangıç kompozisyonunda ark ergitme yöntemiyle üretildi. XRD analizleri, örneklerin oda sıcaklığında $L2_1$ kristal yapısında olduğunu gösterdi. Mn içeriğinin artmasıyla martensitik faz geçiş sıcaklıkları artarken manyetik moment değeri azalmaktadır. Manyetizasyondaki azalma, Mn içeriğinin artmasıyla antiferromanyetik etkileşimlerdeki artıştan kaynaklanmaktadır. Düşük alan ve sıfır alan soğutması altındaki M-T ölçümleri, düşük sıcaklıkta antiferromanyetik etkileşimlerin daha baskın olduğunu, ancak yüksek sıcaklık bölgesinde sistemde ferromanyetik davranışın daha etkili olduğunu ortaya koydu. Manyetik bir malzemede ferromanyetizma ve antiferromanyetizmanın bir arada bulunması, exchange bias etkisine neden olabilir. Bu nedenle, 50 kOe manyetik alan altında soğutulan numunelerin 5 K'de histerezis döngüleri incelenmiştir. Exchange bias etkisine atfedilen manyetizasyon eğrisinin orijinindeki kayma, artan Mn miktarı ile artmıştır.

Anahtar Kelimeler: Ferromanyetik şekil hatırlama etkisi; NiMnSb Heusler alaşımları; Exchange bias etkisi; Antiferromanyetik etkileşimler.

1. Introduction

Shape memory alloys (SMAs) are extensively studied by scientists and industrial organizations due to their useful mechanical properties. As a result of these studies, nowadays SMAs are used in many technological applications such as orthodontic braces and cardiovascular stents used to open occluded or narrowed vessels, as well as industrial applications such as sensitive thermal sensors, medical devices, electronic devices, spacecraft, superelastic eyeglass frames and building materials [1]. While the shape memory effect can be activated by heat in classical SMAs, in ferromagnetic shaped alloys (FSMAs), in addition to heat, it can be achieved with an external magnetic field. The fact that the shape memory mechanism is faster than classical SMAs increases the interest in FSMAs. FSMAs, which can be activated sensitively by magnetic field, have various usage areas such as sensors, magnetic coolers and actuators [2].

The exchange bias (EB) term is used to describe the shift of the hysteresis loop along the magnetic field axis. The origin of this phenomenon is the exchange interaction between antiferromagnetic (AFM) and ferromagnetic (FM) spin structures. EB occurs when a FM and an

AFM spin structures are in contact, and the system is cooled from high temperatures through the Néel temperature of the AFM in an external magnetic field [3].

FSMAs are very convenient to be used in practical applications as it exhibits the shape memory effect and strong magnetization simultaneously. Ferromagnetic Ni_2MnZ ($Z = \text{In, Sn, Sb}$) Heusler alloys (HAs) have been studied extensively in recent years due to their multifunctional physical properties and magnetic behavior in martensitic transformation. Ni-Mn based HAs have attracted considerable attention in the family of shape memory alloys since they exhibit useful physical properties such as magnetic shape memory effect, magnetocaloric effect, magnetoresistance, elastocaloric effect [4].

During the martensitic transformation, the high temperature phase called austenite transforms into a low temperature phase called martensite. It has been reported that the crystal structure transforms from $L2_1$ (high temperature) to $10M$, $14M$ or $L1_0$ (low temperature) with starting martensitic transformation [5, 6]. Since the martensite phase has less symmetry, its magnetic state is not very clear [7]. Generally, the magnetization of HAs is caused by Mn atoms. The nearest neighbor distance between Mn atoms in Ni_2MnGa alloys is about 0.4119 nm. The Ruderman-Kittel-Kaeya-Yo exchange interaction mediated by the conduction electrons gives rise to a ferromagnetic ordering. In Mn-rich HAs, extra Mn atoms occupy Z sites, therefore the distance between Mn atoms suddenly decreases (for $\text{Ni}_2\text{Mn}_{1+x}\text{Ga}_{1-x}$ $x > 0$ about 0.2912 nm) and at such small distances, the magnetic moments preferably align as antiparallel in both high and low temperatures phases of the alloy [8]. With the increase of AFM interactions, which can be expected that the EB property will improve.

In this study, boron-doped NiMnSb SMAs were produced in ingot form by arc melting method. The martensite transition temperatures, hysteresis properties and magnetic properties of the produced samples were revealed by magnetization measurements depending on temperature and magnetic field. As explained above, AFM interactions are expected to increase by decreasing the distance between Mn atoms due to the rich Mn content. It is aimed to improve the EB property by increasing AFM interactions. Therefore, the EB properties of the samples will be investigated in the last stage of the study.

2. Experimental

Polycrystalline $\text{Ni}_{49}\text{Mn}_{37.5}\text{Sb}_{13.5}+\text{B}_x$, $\text{Ni}_{49}\text{Mn}_{37.7}\text{Sb}_{13.3}+\text{B}_x$ and $\text{Ni}_{49}\text{Mn}_{37.9}\text{Sb}_{13.1}+\text{B}_x$ ingots were produced by arc melting method in argon atmosphere. Samples were remelted several times to ensure homogenization. Mn evaporation at the production process makes it difficult to fabricate HAs at the desired stoichiometry [9, 10]. We reported that Mn evaporation was significantly

inhibited in boron added Ni-Mn-Sn HAs [2, 4]. It can be expected that Boron addition exhibited similar results in Ni-Mn-Sb alloys. Therefore, 3 moles boron were added to each sample and it is also known that excess B in the ingots can lead to fragility of the alloys. The ingot samples were heat treated in vacuumed quartz ampoules at 1173 K for 48 h, and then quenched in ice-water. The samples were marked as Sb-13.5, Sb-13.3 and Sb-13.1, respectively.

X-ray diffraction (XRD) for the structural characterization was performed with CuK α radiation in the Rigaku Miniflex 600 device.

Martensitic phase transition temperatures were found by temperature dependent magnetization (M-T) measurements. M-T measurements were performed under constant 100 Oe and 10 kOe magnetic fields between 5 K and 350 K, both heating and cooling. Magnetization dependent on magnetic field (M-H) measurements of ingots were measured at 5 K and 325 K up to the magnetic field of ± 9 T. To determine the EB properties, the samples were cooled to 5 K under 50 kOe magnetic field. The shifts in the origin of the magnetization curve were investigated by the M-H measurements of the samples cooled under the field. All magnetic moment measurements were carried out with VSM attachment of Quantum Design PPMS-9T system.

3. Results and Discussions

The XRD patterns of the Sb-13.5, Sb-13.3 and Sb-13.1 samples are shown in Fig. 1. It was determined that the samples were in the bcc L 2_1 crystal structure, so the samples used in the study are expected to be in the austenite phase at room temperature. The lattice parameter, for Sb-13.5, Sb-13.3 and Sb-13.1 samples were calculated to be 0.5955 nm, 0.5952 nm and 0.5948 nm, respectively, using Rietveld refinement. The lattice parameter decreased by decreasing the amount of Sb and increasing the Mn atoms with smaller radius simultaneously. Consequently, unit cell of austenite phase shrinks with increasing Mn content. The obtained results reveal that the interatomic distance between Mn-Mn decreases by replacing Sb with Mn. Although the intensity of the main peak (220) was obtained in the highest Sb-13.5 sample, no significant difference was observed in the XRD patterns of the samples. In addition, any peak attributable to impurity was not found. Ni-Mn based alloys that exhibit the shape memory effect generally have a cubic L 2_1 crystal structure in the austenite phase. Therefore, the samples can be expected to be in the austenite phase at room temperature.

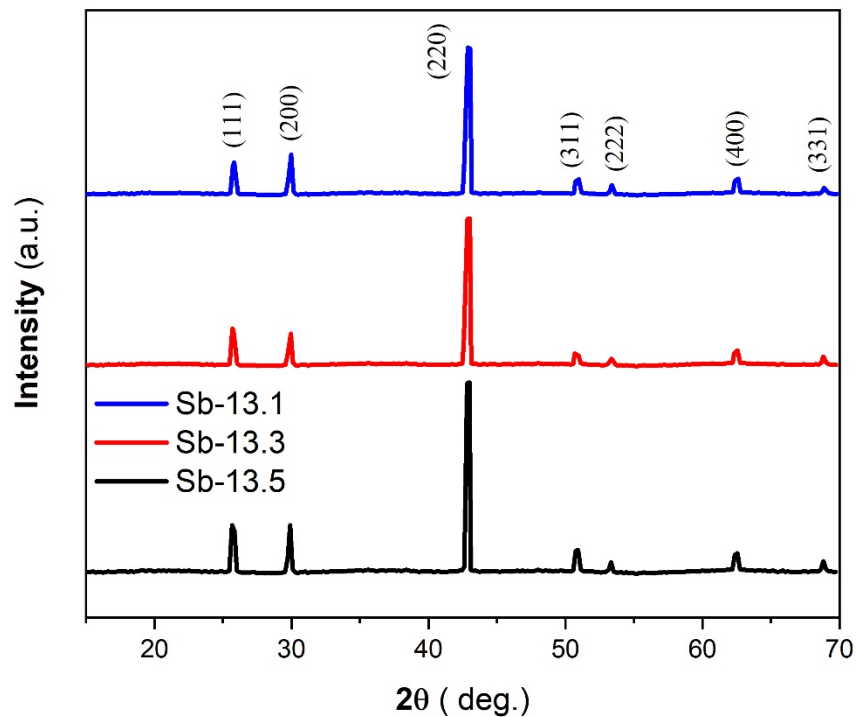


Figure 1: XRD pattern of Sb-13.5, Sb-13.3 and Sb-13.1

During the martensitic transformation, the austenite phase with high crystallographic symmetry (usually cubic) transforms into a martensite phase (like tetragonal and orthorhombic) with a lower order parameter [11]. When austenite to martensite phase transformation occurs in FSMAs, a sharp decrease in magnetization occurs due to the decrease in the order parameter in the crystal structure and vice versa. Additionally, a thermal hysteresis occurs during the phase transition because of the twin boundaries formed in the martensite phase and friction at the twin boundaries [12]. Temperature dependent magnetization curves (M-T) of the Sb-13.5, Sb-13.3 and Sb-13.1 samples are given in Fig. 2. According to the obtained results, all samples exhibited martensitic transformation. It was determined that all samples exhibited martensitic transformation and the phase transition temperature increased with increasing Mn content. The changing of the martensitic transition temperatures is usually attributed to the number of valence electrons per atom (e/a) and the unit-cell volume of the austenite phase. The phase transition temperature is usually directly proportional to e/a . If the e/a effect is analyzed in terms of the band model: Austenite with $L2_1$ crystal structure is stabilized since the Fermi surface just touches the (110) Brillouin zone. By increasing e/a , the Fermi surface and (110) Brillouin regions overlap; and as a result, excessive increase in system energy leads to lattice distortion to minimize free energy, i.e. martensite formation. The second factor affecting the phase transition temperatures is the atomic size effect. It is widely accepted that unit cell shrinkage of austenite at high temperature will promote martensitic transformation due to changes in the relative positions between the Fermi

surface and the Brillouin zone. As a result, increasing the Mn content raises the e/a ratio, thus causing an increase in the phase transition temperature [13].

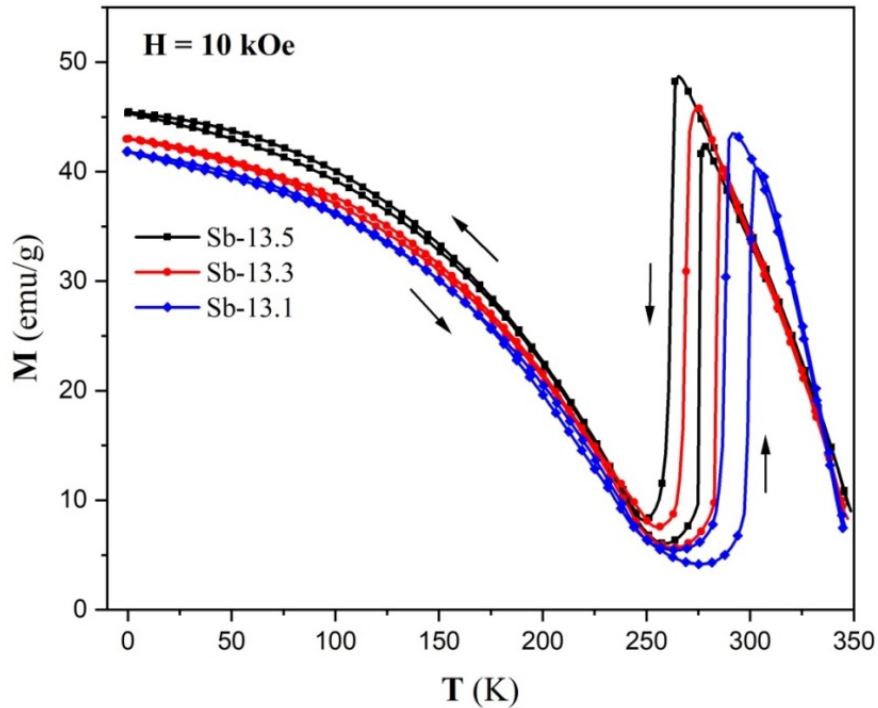


Figure 2: M-T curves of the Sb-13.5, Sb-13.3 and Sb-13.1 samples at 10 kOe

Another notable point in Fig. 2 is reducing in magnetic moment depend on Mn content. The maximum magnetic moment value decreased with the increase of Mn content. The contribution of Ni to the total magnetization is quite low since Ni atoms in Ni-Mn-X ($X=\text{Sn, Sb, In}$) alloys have an antiparallel spin arrangement. The total magnetic moment is mostly calculated from the spin moments of Mn atoms [14, 15]. Since the exchange interactions in HAs depend on the distance between Mn and Mn, the magnetization and phase transformation temperatures are highly dependent on the chemical composition [16]. The magnetic properties of HAs can be modified by changing the distance between Mn atoms. If the Mn content of HAs ($\text{Ni}_{50}\text{Mn}_{25+x}\text{Sb}_{25-x}$) is greater than 25, extra Mn atoms will occupy the Sb site. The distance between Mn atoms gets shorter and its spin moment will be antiparallel to that of other Mn atoms at Mn sites. An increase in antiparallel interactions causes a decrease in the maximum magnetic moment value in the FSMAs [15].

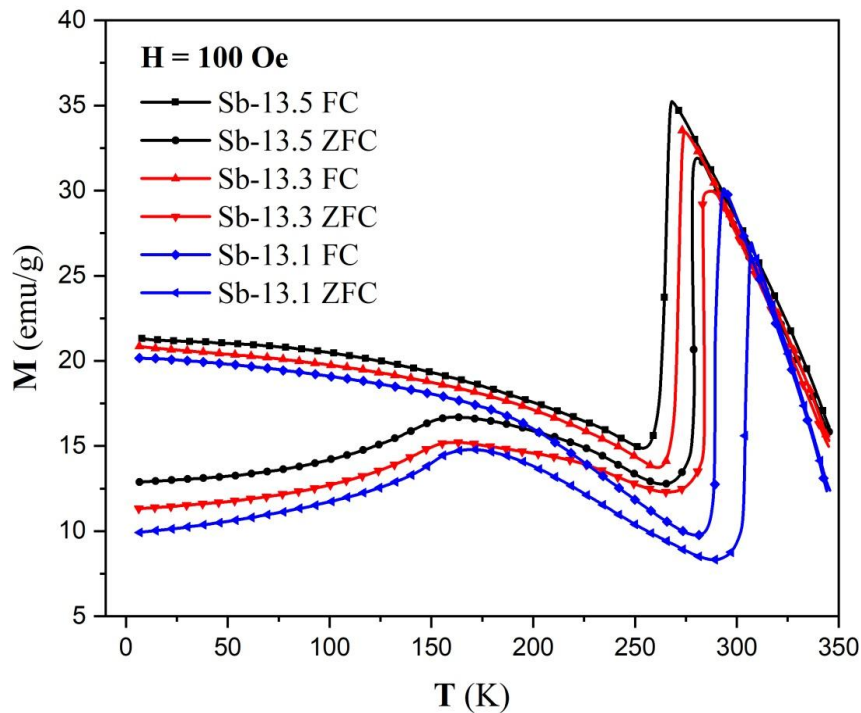
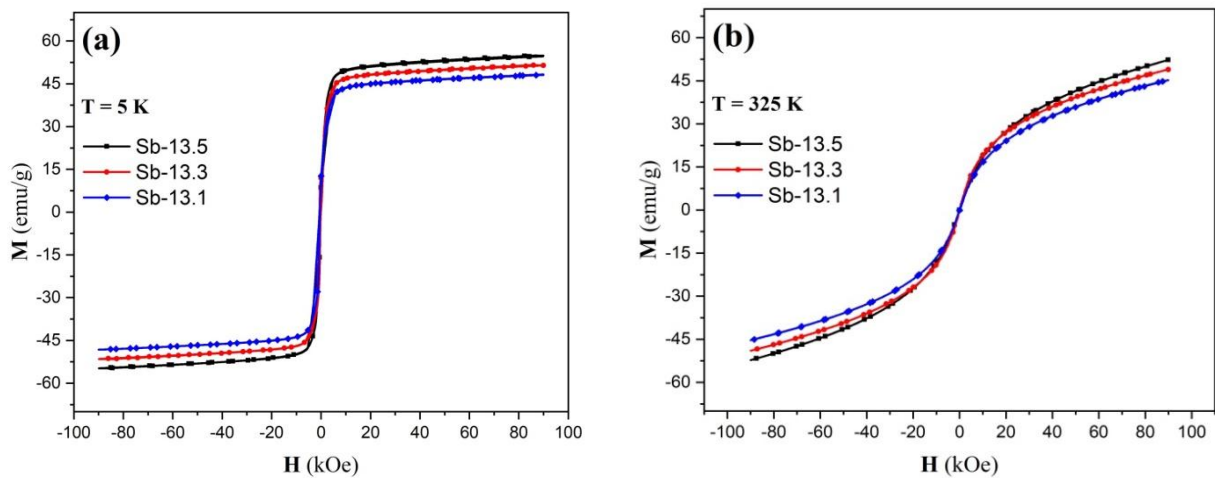


Figure 3: ZFC and FC M-T curves of the Sb-13.5, Sb-13.3 and Sb-13.1 samples at 100 Oe

Although it is expected that antiferromagnetic interactions will increase with increasing Mn amount, there is no data that can be direct evidence of the existence in antiferromagnetic interactions in Fig. 2 because the spin alignment of the materials is strongly affected by the 10 kOe magnetic field. Therefore, temperature-dependent magnetization measurements were carried out in both ZFC (Zero Field Cooling) and FC (Field Cooling) regimes at a relatively low magnetic field. Firstly, the samples were cooled to 5 K without a field, then the measurements were carried out by applying the field (100 Oe). In the final step, the magnetizations of the cooled samples were measured under 100 Oe field. The obtained results are shown in Fig. 3. It is clear that magnetization increases with increasing temperature in the low temperature region (between 5K and ~175K) in all samples. This increase is a common characteristic behavior of AFM. When the temperature rises above about 175 K, magnetization decreases. It shows that ferromagnetic interaction is more dominant in this region. With the further increase in temperature, phase transitions occurred at different temperatures for each sample and a sharp increase in magnetization occurred. In cooling M-T (FC) analyses, magnetization increases continuously after the completion of the martensitic phase transition due to permanent magnetization. When M-T curves in high magnetic field (10 kOe Fig. 2) and low magnetic field (100 Oe Fig. 3) are compared, phase transition temperatures decrease with increasing field. Martensitic phase transition temperatures are listed in Table 1.

Table 1: Martensitic phase transition temperatures of the samples

Sample	Magnetic Field	A_s (K)	A_f (K)	M_s (K)	M_f (K)
Sb-13.5	100 Oe	270.41	279.56	268.17	255.77
	10 kOe	267.16	277.14	264.83	249.40
Sb-13.3	100 Oe	278.42	287.14	274.22	264.23
	10 kOe	274.97	286.03	273.55	257.96
Sb-13.1	100 Oe	297.39	307.23	292.32	284.90
	10 kOe	292.97	302.38	290.51	279.38

**Figure 4:** M-H results of the samples (a) at 5 K (b) at 325 K

Magnetization vs applied magnetic field (M-H) curves of the samples at 5 K and 325 K are presented in Fig. 4a and 4b, respectively. While the magnetization saturates quickly at 5 K, they did not reach saturation at 325 K. Consistent with the M-T analysis, the highest magnetization value was obtained in the Sb-13.5 sample. As the Mn content increases, a contraction occurs in the unit cell and distance between Mn atoms becomes shorter. Mn atoms nearing each other more than the critical distance prefer antiparallel spin alignment. Therefore, a decrease occurs in total magnetization.

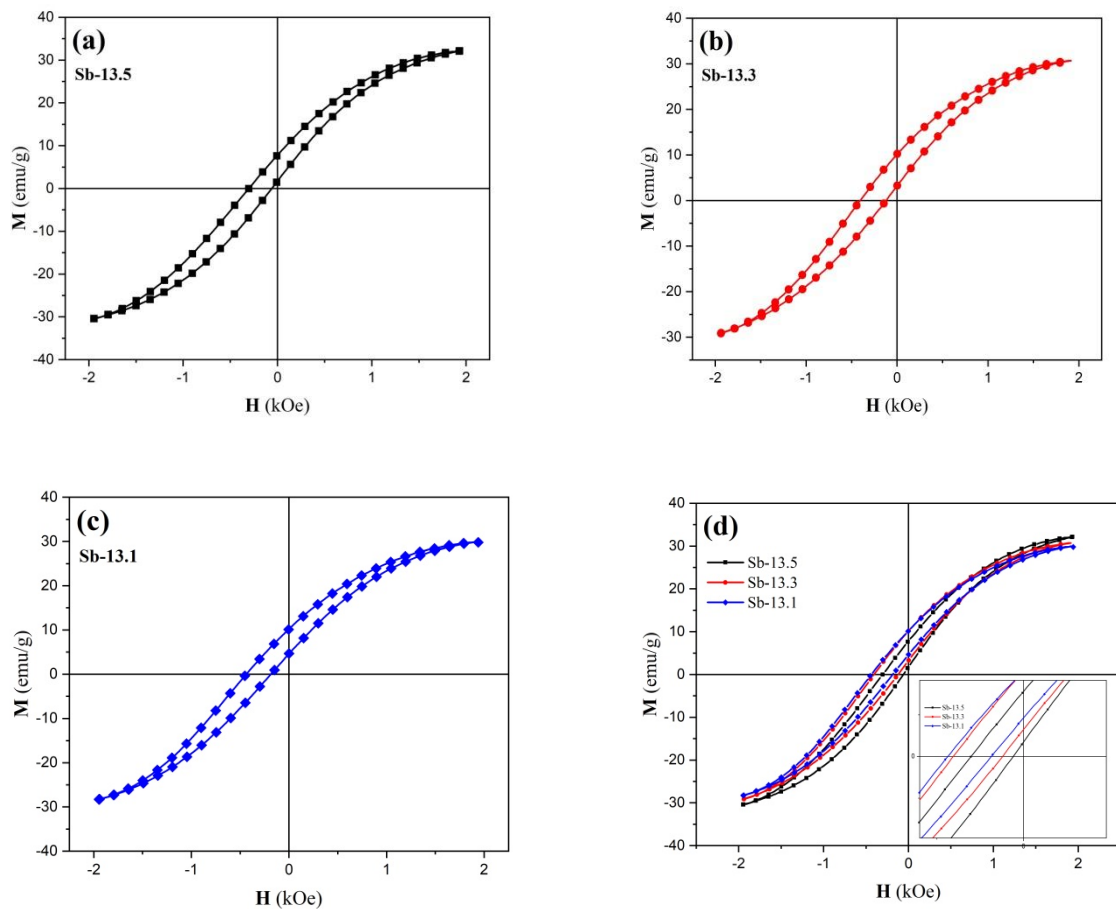


Figure 5: Hysteresis loops of the (a) Sb-13.5 (b) Sb-13.3 (c) Sb-13.1 and (d) overlapped for all samples at 5 K

The origin of the EB effect is fundamentally the exchange interaction that occurs at the interface of the antiferromagnetic and ferromagnetic phases. In the absence of an external magnetic field, the AFM has no net magnetization or is very small. With the application of the external magnetic field, the spins of the soft FM phase rapidly rotate in the same direction as the field. The spins of the AFM phase close to the AFM-FM interface are affected by this strong magnetization and they become the same alignment with FM phase. If the added energy is enough to create a Néel domain wall within the AFM, it implies a shift in the switching field of the FM. As a result, a shift occurs in the origin of the magnetization curve [17]. For the EB effect to occur, T_N (Néel temperature) should be lower than T_C (Curie temperature) [7]. To examine the EB effect, the samples were cooled under 50 kOe magnetic field and M-H measurements were carried out at 5 K. The obtained hysteresis loops are given separately for each sample (Fig. 5a-c) and overlapped for all samples (Fig. 5d). The shift to the centers of origins attributed to the EB effect is clearly visible from Fig. 5. The obtained data from the XRD revealed that the unit cell narrowed by decreasing the Sb content and increasing the Mn. As discussed above, with decreasing the distance between atoms, the AFM interactions increase between Mn atoms. With increasing AFM

interactions, the EB effect and hence the shift in the origin of the hysteresis increases. EB is mainly connected to the coexistence of FM-AFM interfaces [18]. Therefore, the presence of EB in HAs is a proof that AFM interactions are present in the system. This indicates that there is an AFM interaction inside the FM matrix at low temperatures. Currently, the effect of EB in FSMA is not completely understood as T_N is not clearly determined in these materials [19]. Besides, the EB field (H_{EB}) can be determined from Fig. 5 as a function of Mn content. The H_{EB} is defined as $H_{EB} = -(H_1 + H_2)/2$. Here, H_1 and H_2 are the left and right coercive fields, respectively [20–23]. H_{EB} values of Sb-13.5, Sb-13.3 and Sb-13.1 samples were calculated as 340 Oe, 600 Oe and 820 Oe, respectively.

4. Conclusion

In this study, the $Ni_{49}Mn_{37.5}Sb_{13.5}+B_x$, $Ni_{49}Mn_{37.7}Sb_{13.3}+B_x$ and $Ni_{49}Mn_{37.9}Sb_{13.1}+B_x$ ingots were fabricated by arc melting method. It was determined that the produced samples were in cubic $L2_1$ structure at room temperature with XRD analysis. Mn evaporation complicates the production of HAs. Boron was added to the nominal compositions to prevent Mn evaporation (3 mol). M-T analyzes at 10 kOe showed that all samples exhibited martensitic transformation. The phase transition temperature increased with increasing the content of Mn in accordance with the e/a rule. In HAs with excess Mn, the distance between Mn atoms becomes shorter and Mn atoms prefer an AFM alignment. ZFC and FC M-T measurements were performed at low magnetic field (100 Oe) to determine the antiferromagnetic behavior of the samples. ZFC measurements revealed that antiferromagnetism was dominant in the 5-175 K range. In the high temperature region, ferromagnetism exhibited a more dominant behavior. Magnetic field dependent magnetization measurements showed that Mn content and magnetization were inversely proportional. This is due to the increase of antiferromagnetic interactions in the ferromagnetic matrix with increasing Mn content. The EB effect, which is important for technological applications, was investigated in the last part. It was concluded that the EB effect was improved with the increase of Mn content. Besides, these results are evidence for the existence of antiferromagnetic interactions in HAs, especially at low temperatures.

References

- [1] Asai, M., Suzuki, Y., *Applications of shape memory alloys in Japan*, Materials Science Forum, 327–328, 17–22, 2000.
- [2] Kirat, G., Kizilaslan, O., Aksan, M.A., *Magnetoresistance properties of magnetic Ni-Mn-Sn-B shape memory ribbons and magnetic field sensor aspects operating at room temperature*, Journal of Magnetism and Magnetic Materials, 477, 366–371, 2019.
- [3] Blachowicz, T., Ehrmann, A., *Exchange bias in thin films—an update*, Coatings, 11, 1–

21, 2021.

[4] Kirat, G., Aksan, M.A., Aydogdu, Y., *Magnetic field induced martensitic transition in Fe doped Ni-Mn-Sn-B shape memory ribbons*, *Intermetallics*, 111, 106493, 2019.

[5] Sánchez Llamazares, J.L., Sanchez, T., Santos, J.D., Pérez, M.J., Sanchez, M.L., Hernando, B., Escoda, L., Suñol, J.J., Varga, R., *Martensitic phase transformation in rapidly solidified Mn₅₀Ni₄₀In₁₀ alloy ribbons*, *Applied Physics Letters*, 92, 8–11, 2008.

[6] Bachagha, T., Zhang, J., Sunol, J.J., Khitouni, M., *Martensitic transformation and magnetic behavior in Mn-rich Heusler MnNiIn shape memory alloys*, *IOP Conference Series Materials Science and Engineering*, 504, 012009, 2019.

[7] Wang, B.M., Liu, Y., Wang, L., Huang, S.L., Zhao, Y., Yang, Y., Zhang, H., *Exchange bias and its training effect in the martensitic state of bulk polycrystalline Ni_{49.5}Mn_{34.5}In₁₆*, *Journal of Applied Physics*, 104, 043916, 2008.

[8] Yu, S.Y., Cao, Z.X., Ma, L., Liu, G.D., Chen, J.L., Wu, G.H., *Realization of magnetic field-induced reversible martensitic transformation in NiCoMnGa alloys*, *Applied Physics Letters*, 91, 102507, 2007.

[9] Maziarz, W., Czaja, P., Wójcik, A., Dutkiewicz, J., Przewoźnik, J., Cesari, E., *Microstructure and martensitic transformation of Ni₅₀Mn_{37.5}Sn_{12.5-x}Ge_x (X=0, 1, 2, 3) Heusler alloys produced by various Technologies*, *Materials Today: Proceedings*, 2(3), S523–S528, 2015.

[10] Yu, J.K., Li, H.W., Zhai, Q.J., Fu, J.X., Luo, Z.P., Zheng, H.X., *Crystal structure and formation mechanism of the secondary phase in Heusler Ni-Mn-Sn-Co materials*. *Advances in Manufacturing*, 2, 353–357, 2014.

[11] Zelený, M., Straka, L., Sozinov, A., Heczko, O., *Transformation paths from cubic to low-symmetry structures in Heusler Ni₂MnGa compound*, *Scientific Reports*, 8, 7275, 2018.

[12] Wang, W.-H., Chen, J.-L., Liu, Z., Wu, G.-H., Zhan, W.-S., *Thermal hysteresis and friction of phase boundary motion in ferromagnetic Ni₅₂Mn₂₃Ga₂₅ single crystals*, *Physical Review B*, 65, 012416, 2001.

[13] Zheng, H., Wang, W., Xue, S., Zhai, Q., Frenzel, J., Luo, Z., *Composition-dependent crystal structure and martensitic transformation in Heusler Ni–Mn–Sn alloys*, *Acta Materialia*, 61, 4648–4656, 2013.

[14] Kübler, J., William, A.R., Sommers, C.B., *Formation and coupling of magnetic moments in Heusler alloys*, *Physical Review B*, 28, 1745–1755, 1983.

[15] Luo, H., Meng, F., Jiang, Q., Liu, H., Liu, E., Wu, G., Wang, Y., *Effect of boron on the martensitic transformation and magnetic properties of Ni₅₀Mn_{36.5}Sb_{13.5-x}B_x alloys*, *Scripta Materialia*, 63(6), 569–572, 2010.

[16] Hernando, B., Llamazares, J.L.S., Santos, J.D., Sánchez, M.L., Escoda, L., Suñol, J.J., Varga, R., García, C., González, J., *Grain oriented NiMnSn and NiMnIn Heusler alloys ribbons produced by melt spinning: Martensitic transformation and magnetic properties*, *Journal of Magnetism and Magnetic Materials*, 321(7), 763–768, 2009.

[17] Stamps, R.L., *Mechanisms for exchange bias*, *Journal of Physics D: Applied Physics*, 33(23), R247–R268, 2000.

[18] Czaja, P., Przewoźnik, J., Fitta, M., Bałanda, M., Chrobak, A., Kania, B., Zackiewicz,

P., Wójcik, A., Szezynger, M., Maziarz, W., *Effect of ball milling and thermal treatment on exchange bias and magnetocaloric properties of $Ni_{48}Mn_{39.5}Sn_{10.5}Al_2$ ribbons*, *Journal of Magnetism and Magnetic Materials*, 401, 223–230, 2016.

[19] Acet, M., Mañosa, L., Planes, A., *Magnetic-Field-Induced effects in martensitic Heusler-based magnetic shape memory alloys*, *Handbook of Magnetic Materials*, 19, 231–289, 2011.

[20] Li, S., Wang, S., Lu, Y., Zhang, C., Yang, X., Gao, J., Li, D., Zhu, Y., Liu, W., *Exchange bias effect in hybrid improper ferroelectricity $Ca_{2.94}Na_{0.06}Mn_2O_7$* , *AIP Advances*, 8, 015009, 2018.

[21] Wang, R.L., Lee, M.K., Xu, L.S., Sun, Z.G., Marchenkov, V. V., Tien, C., Huang, J.C.A., Yang, C.P., *Effect of thermal cycle on the interfacial antiferromagnetic spin configuration and exchange bias in Ni-Mn-Sb alloy*, *AIP Advances*, 2, 032181, 2012.

[22] Fertman, E.L., Fedorchenko, A. V., Desnenko, V.A., Shvartsman, V. V., Lupascu, D.C., Salamon, S., Wende, H., Vaisburd, A.I., Stanulis, A., Ramanaukas, R., Olekhovich, N.M., Pushkarev, A. V., Radyush, Y. V., Khalyavin, D.D., Salak, A.N., *Exchange bias effect in bulk multiferroic $BiFe_{0.5}Sc_{0.5}O_3$* , *AIP Advances*, 10, 045102, 2020.

[23] Abulyazied, D.E., Abomostafa, H.M., El komy, G.M., *Magnetic structured nickel core-shell @ silica/PMMA nanocomposites from synthesis to applications*, *Journal of Inorganic and Organometallic Polymers and Materials*, 30, 2335–2346, 2020.



A Comparative Study of DFT/B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++G(d,p) and HSEH1PBE/6-31G(d,p) Methods Applied to Molecular Geometry and Electronic properties of C_s-C₆₀Cl₆ Molecule

Ebru KARAKAŞ SARIKAYA^{1,*}, Ömer DERELİ², Semiha BAHÇELİ³

¹*Necmettin Erbakan of University, Engineering Faculty, Department of Basic Sciences, Konya, Türkiye*
ebrukarakas_84@hotmail.com, ORCID: 0000-0003-2149-9341

²*Necmettin Erbakan of University, Faculty of A. K. Education, Department of Physics, Konya, Türkiye*
odereli@erbakan.edu.tr, ORCID: 0000-0002-9031-8092

³*Emeritus Professor of Atomic and Molecular Physics, 06000, Ankara, Türkiye*
sbahceli@thk.edu.tr, ORCID: 0000-0002-5614-325X

Received: 16.05.2021

Accepted: 06.12.2021

Published: 31.12.2021

Abstract

In this study, four different levels, B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++G(d,p) and HSEH1PBE/6-31G(d,p) of the DFT quantum chemical calculation method have been applied to the molecular structure of the C_s-C₆₀Cl₆ molecule as a halogenated fullerene. Additionally, the molecular structure of pure C₆₀ fullerene was presented as complementary and supportive work. Furthermore, the simulated FT-IR, Raman and UV-Vis (in cyclohexane solvent) spectra, HOMO-LUMO analysis, the molecular electrostatic potential (MEP) map, the ¹³C NMR chemical shift values in both gas phase and tetrachloromethane with deuterated chloroform solvent and the thermodynamics properties at the mentioned levels of the C_s-C₆₀Cl₆ molecule were reported. Fullerene has many physical and electrochemical properties, which can be utilized in several medical fields. Especially, it can fit inside the hydrophobic cavity of HIV proteases, restricting the get into substrates to the catalytic site of the enzyme. Hence, it is utilizable as an antioxidant and radical scavenger.



Keywords: Fullerenes; $C_s-C_{60}Cl_6$; DFT; HOMO-LUMO analysis; The ^{13}C NMR chemical shifts; MEP.

$C_s-C_{60}Cl_6$ Molekülünün Moleküler Geometri ve Elektronik Özelliklerine Uygulanan Metot ve Baz Seti, DFT/B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++ G(d,p) ve HSEH1PBE/6-31G(d,p), Yöntemlerinin Karşılaştırmalı İncelemesi

Öz

Bu çalışmada, DFT kuantum kimyasal hesaplama yönteminin, B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++G(d,p) ve HSEH1PBE/6-31G(d,p) olmak üzere dört farklı düzeyinde, halojenleşmiş bir fullerene olan $C_s-C_{60}Cl_6$ molekülünün moleküler yapısına uygulanmıştır. Ek olarak, saf C_{60} fullerenin molekül yapısı, tamamlayıcı ve destekleyici bir çalışma olarak sunulmuştur. Ayrıca, simüle edilmiş FT-IR, Raman ve UV-Vis (sikloheksan çözücüsünde) spektrumları, HOMO-LUMO analizi, moleküler elektrostatik potansiyel (MEP) haritası, döteryumlanmış kloroform çözücü ile hem gaz fazında hem de tetraklorometanda ^{13}C NMR kimyasal kayma değerleri ve $C_s-C_{60}Cl_6$ molekülünün belirtilen düzeylerdeki termodinamik özellikleri rapor edilmiştir. Fulleren, çeşitli tıbbi alanlarda kullanılabilir birçok fiziksel ve elektrokimyasal özelliğe sahiptir. Özellikle, HIV proteazlarının hidrofobik boşluğunun içine sığabilir ve substratlara enzimin katalitik bölgesine girmesini kısıtlayabilir. Bu nedenle, bir antioksidan ve radikal temizleyici olarak kullanılabilir.

Anahtar Kelimeler: Fulleren; $C_s-C_{60}Cl_6$; DFT; HOMO-LUMO analizi; ^{13}C NMR kimyasal kayma; MEP.

1. Introduction

As the best example for the monotype molecules, the fullerene or Buckminster fullerene C_{60} , which derives its name from the geodesic dome designed by the architect Buckminster Fuller, is any of a series of hollow carbon molecules [1-4]. Over 40 years after its discovery, the fullerenes still attract the attention of many researchers in various scientific and technological fields such as nanotechnology, material science, medicine, chemistry, and physics [5-10]. Although the fullerene C_{60} is the most popularly applied in a large range of fields, the other fullerenes with a different number of carbon atoms were also discovered. Within this framework, the sizes of hollow carbon polyhedra called fullerenes vary from C_{20} to lattices of more than 400 carbon atoms [11-12].

Furthermore, the derivatives of the various fullerene complexes with heavy metals were also found the interesting applications in medicinal chemistry, material science, and nanotechnology [13-17].

On the other hand, it is well-known that halogen bonding occurred in a substance with a halogen atom provides a nucleophilic region which plays a crucial role in pharmacology, materials sciences and biology [18]. In this context, the halogenated fullerenes have also been investigated intensively during the last two decades and demonstrate encouraging characteristics toward material science utilizations since they are flexible predecessors for the synthesis of diverse intricate derivatives. In this framework, the chloro fullerenes are very useful for the halogen derivatives with optical and biomedical properties [19-22]. Likewise, sizable completion has been accomplished in the development and structural specification related to fullerene bromides [19] and fluorides [23]. Regardless of countless statements on chlorination of C_{60} via distinct reagents, solely two characteristic compounds have been seal off density functional theory (DFT) for fullerene derivatives has been broadly experienced in describe their structure and molecular properties [24-27]. In this context, the use of the B3LYP level with the 6-31G* basis set of the theory can be mentioned for applicability to large molecules [28].

DFT is used for calculating molecular geometry parameters of compounds, spectroscopic properties such as Raman, 1H and ^{13}C NMR chemical shift values and thermodynamic characteristics, UV-Vis spectra, IR, electronic features like HOMO, MEP, and LUMO of the compounds [29]. Take into consideration, the satisfactory electronic correlation influences and selecting capable basis sets in this technique, the trustworthy vibrational frequencies and optimized geometric coefficients of molecules be able to foresee [30-33]. At the beginning of the 1990s in DFT, the Becke-3-Lee Yang Parr (B3LYP) hybrid approach and the Becke's three parameters incorporating Perdew and Wang's 1991 (B3PW91) approach are labeled first-generation methods, and the Heyd-Scuseria-Ernzerhof hybrid combined with Perdew, Burke, and Ernzerhof's exchange and correlation functions (HSEH1PBE) approach, which is also mentioned as the HSE06 approach, is a second-generation method [33-37].

Our study is settled on research carried out by Kuvychko et al. [21] in where the Infrared and Raman spectra of $C_s-C_{60}Cl_6$ were extensively investigated by both experimental studies and the vibrational frequency computations toward the PBE/6-31G* grade of the principles. However, the molecular structure and electronic properties of the $C_s-C_{60}Cl_6$ molecule were not studied properly in the mentioned work. Therefore, even if it is an absence of experimental information considering DFT calculations are a forceful quantum chemical instrument for the identification

of the electronic frame related to molecules, we found it worth focusing on investigating the molecular frame and electronic features of $C_s-C_{60}Cl_6$ compounds in detailed data [38, 39].

In this framework, the present work reports the results of the quantum chemical calculations of the optimized molecular construction, the FT-IR, UV-Vis (in cyclohexane solvent) and FT-Raman by using the B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++G(d,p), and HSEH1PBE/6-31G(d,p) levels of the theory [34,40,41]. In this article, abbreviations will be used BS-1 for B3LYP/6-31G(d,p), BS-2 for RM062X/6-31G(d,p), BS-3 for B3LYP/6-311++G(d,p), BS-4 for HSEH1PBE/6-31G(d,p), respectively. Furthermore, the NMR chemical shifts (1H and ^{13}C) both in tetrachloromethane solutions and in the gas-phase; the UV spectrum in cyclohexane (the upper limit electronic suction wavelengths given that λ_{max}); LUMO-HOMO energy spaces and thermodynamics characteristics of the $C_s-C_{60}Cl_6$ molecule are presented.

2. Materials and Methods

2.1. Computational details

In this study, optimization calculations in the ground state were made in the Gaussian 09W [42] program for all conformations. The output files were contemplated via Gauss View software. [43]. Also, calculations were performed by Dell Rack server. The structural properties, which is electronic properties and molecular geometry of the $C_s-C_{60}Cl_6$ fullerene molecule, were estimate at the BS-1, BS-2, BS-3 and BS-4 levels of the DFT. However, the simulated FT-Raman and FT-IR vibrational spectra of the $C_s-C_{60}Cl_6$ molecule were presented by computing at the mentioned levels [34, 40, 41]. Furthermore, the optimized structure of the C_{60} fullerene was calculated with only at the BS-1 method.

Nevertheless, the stable structure of the $C_s-C_{60}Cl_6$ fullerene molecule was initially obtained with the BS-3 method by way of CPCM (the conductor-like polarizable continuum technique). After that, NMR values were determined using GIAO (the gauge-including atomic orbital) method attribute to optimization at the mentioned levels in tetrachloromethane with deuterated chloroform and in the gas phase ($\epsilon = (2.24$ and $4.8)$) [38, 45]. The UV-Vis spectra (in ethanol and in the gas phase) were acquired utilizing the time-dependent DFT (TD-DFT) theory at the BS-1, BS-2 and BS-3 levels [46, 47]. Furthermore, the FMO and MEP maps that figure out the title compound were performed through the BS-1, BS-2, BS-3 and BS-4 levels of the theory additionally, their 3D plots were confirmed at the mentioned level.

3. Results and Discussion

3.1. Molecular structures

The stable structures of the fullerene C_{60} molecule at the BS-1 and the $C_s-C_{60}Cl_6$ fullerene molecule at the BS-3 levels, respectively, and their atom numbers are exhibited in Fig. 1.

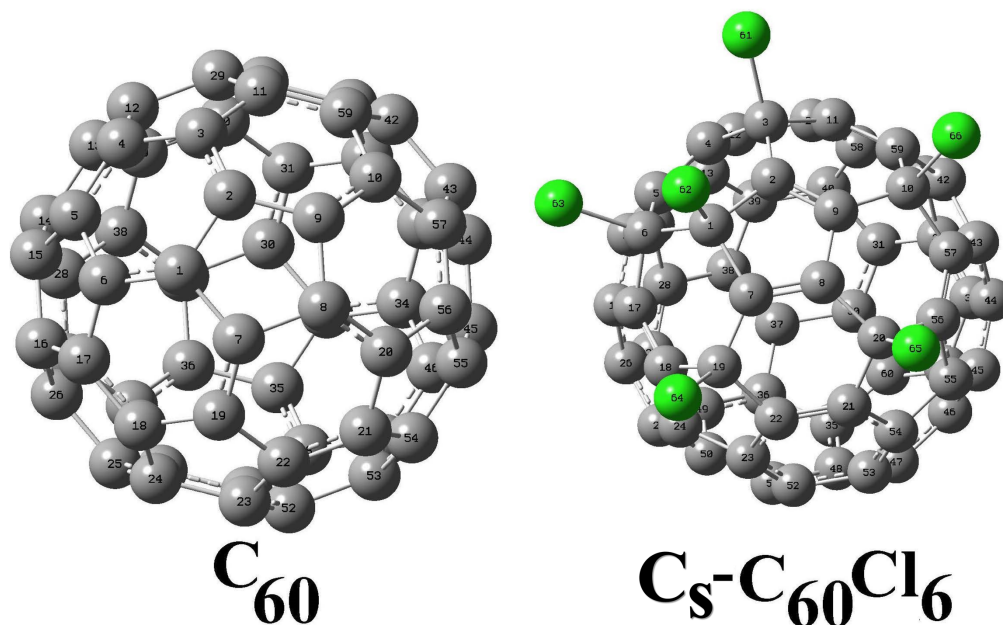


Figure 1: The stable optimized molecular structures of the C_{60} fullerene at the BS-1 level on the left and the $C_s-C_{60}Cl_6$ molecule at the BS-3 level on the right with their atom numberings

Table 1: The calculated most stable state energy values of C_{60} and $C_s-C_{60}Cl_6$ fullerenes

	Method/Basis set	E (Hartree)	Dip. Mom. (D)
C_{60}	(BS-1)	-2286.174135	0.000000
$C_s-C_{60}Cl_6$	(BS-1)	-5047.325711	6.476442
$C_s-C_{60}Cl_6$	(BS-2)	-5046.447915	5.267343
$C_s-C_{60}Cl_6$	(BS-3)	-5044.200344	5.527418
$C_s-C_{60}Cl_6$	(BS-4)	-5047.944625	5.238593

Furthermore, the calculated energy values of the fullerene C_{60} and the $C_s-C_{60}Cl_6$ compound for their most stable states are summed up in Table 1. Nonetheless, the stable geometry parameters of the C_{60} fullerene calculated with the BS-1 method are stated in Table S1 as the supplementary material. Similarly, the optimized molecular geometry parameters calculated at the BS-1, BS-2, BS-3 and BS-4 levels of the $C_s-C_{60}Cl_6$ compound are given in Table S2.

As far as we know, the researchers have not been yet performed X-ray spectroscopy work on the $C_s-C_{60}Cl_6$ sample. However, the chlorination with ICl produces the $C_s-C_{60}Cl_6$ compound which is isostructural with $Cl_{60}Br_6$ [48-49]. For this reason, we used some crystallographic values

of the $\text{Cl}_{60}\text{Br}_6$ compound [48] in order to compare with the bond lengths calculated with the BS-1 of $\text{C}_s\text{-C}_{60}\text{Cl}_6$ fullerene. In this context, the $\text{C}_s\text{-C}_{60}\text{Cl}_6$ molecular has a non-crystallographic C_s symmetry, a monoclinic crystal system, and a P21/c space group [48-49]. Therefore, the experimental values of the C-C(Cl) single bonds, the C-Cl bond lengths, and the remaining C-C bonds in the central pentagons of the $\text{C}_s\text{-C}_{60}\text{Cl}_6$ fullerene should be averaged 1.96 Å, 1.53 Å and 1.45 Å while the remaining inter pentagonal C-C bonds should be averaged 1.38 Å [49-50]. According to our calculations which are shown in Table S2, the C-Cl bond lengths in the pentagons can be averaged 1.85 Å at the BS-1 and BS-3 methods, 1.81 Å at the BS-3 method and 1.82 Å at the BS-4 method. The dissimilarity among the experimental C-Br and theoretical C-Cl bond lengths which is about 0.11 Å can be attributed to the effect of chlorine halogen atom on the bond with C atom as less electronegative than Br atoms [50]. Similarly, the C1-C2, C1-C7 and C8-C9 single bond lengths in the pentagon of the $\text{C}_s\text{-C}_{60}\text{Cl}_6$ fullerene were calculated as 1.54 Å at the BS-1, BS-2 and BS-3 levels, and as 1.53 Å at the BS-4 level which is excellent correspondence with the experimental results [48]. Likewise, the C7=C8 and C2=C29 double bonds in the pentagon were reckoned as 1.35 Å and 1.34 Å at the BS-1 level and also worked out BS-2, BS-3 and BS-4 levels, respectively, for both double bonds which are also in perfect match with the experimental values [51], while the C1-C6 and C8-C20 single bonds in inter pentagon were found as 1.58 Å at the BS-1 and BS-3 levels, and as 1.57 Å at the BS-2 and BS-4 levels.

Moreover, by considering Fig. 1 and Table S1 for C_{60} fullerene the calculated C-C single bonds at the BS-1 level can be averaged almost as 1.45 Å and the double C=C bonds can be also averaged as 1.39 Å. These calculated data are in excellent correspondence including the values in the literature [52]. Therefore, by considering the given results we can state that the chlorination of C_{60} fullerene exhibits a significant distortion on the fullerene.

Furthermore, as seen in Table S2, we can declare that the dihedral angles and bond of the $\text{C}_s\text{-C}_{60}\text{Cl}_6$ molecule computed at the mentioned four different levels can be formed a database for the studies about the halogenated fullerenes in the future.

3.2. Vibrational frequencies

While we pointed out in the section introduction, the experimental and theoretical determined at the PBE/TZ2P level Infrared and Raman vibrational frequency values of the $\text{C}_s\text{-C}_{60}\text{Cl}_6$ molecule were given in the work which was verified by Kuvychko et al. [21]. Furthermore, the experimental infrared vibrational frequencies [53] and the assignments of the infrared and Raman vibrational band modes were verified in the early studies [19, 21, 54]. For this reason, we

presented here, only the simulated IR and Raman vibrational spectra of the $C_s-C_{60}Cl_6$ molecule at the BS-1, BS-2, BS-3 and BS-4 levels are shown in Figs. 2 and 3, respectively.

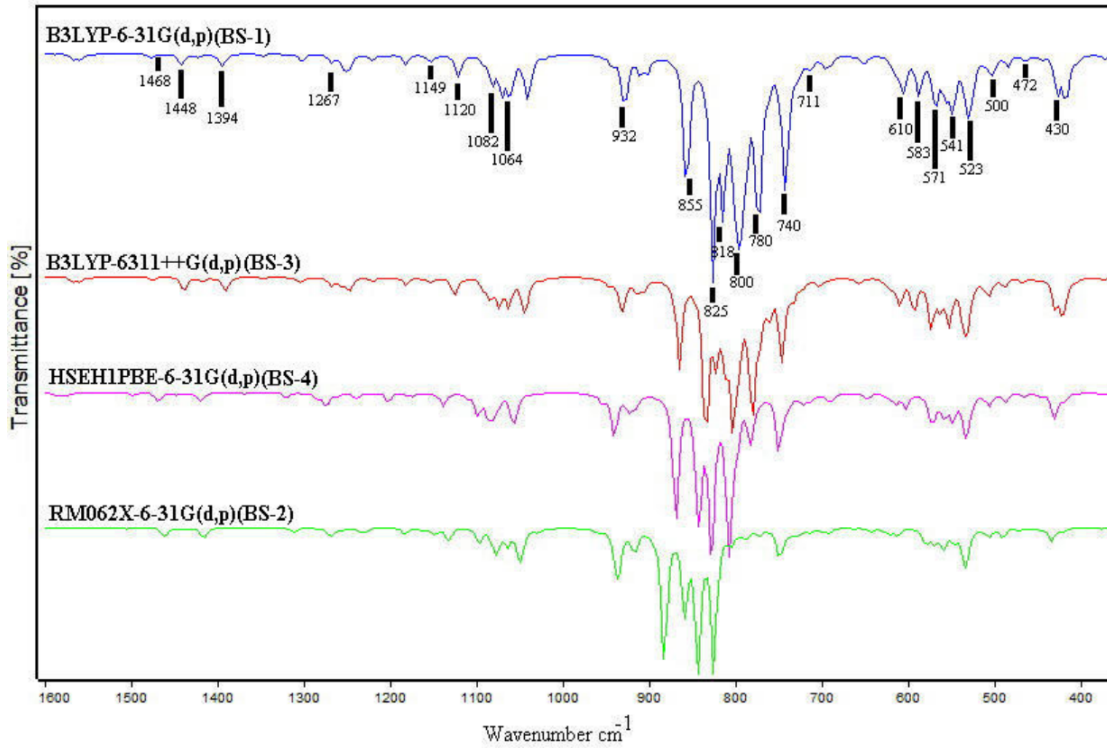


Figure 2: The simulated IR spectra at the BS-1, BS-2, BS-3 and BS-4 levels for the $C_s-C_{60}Cl_6$ molecule

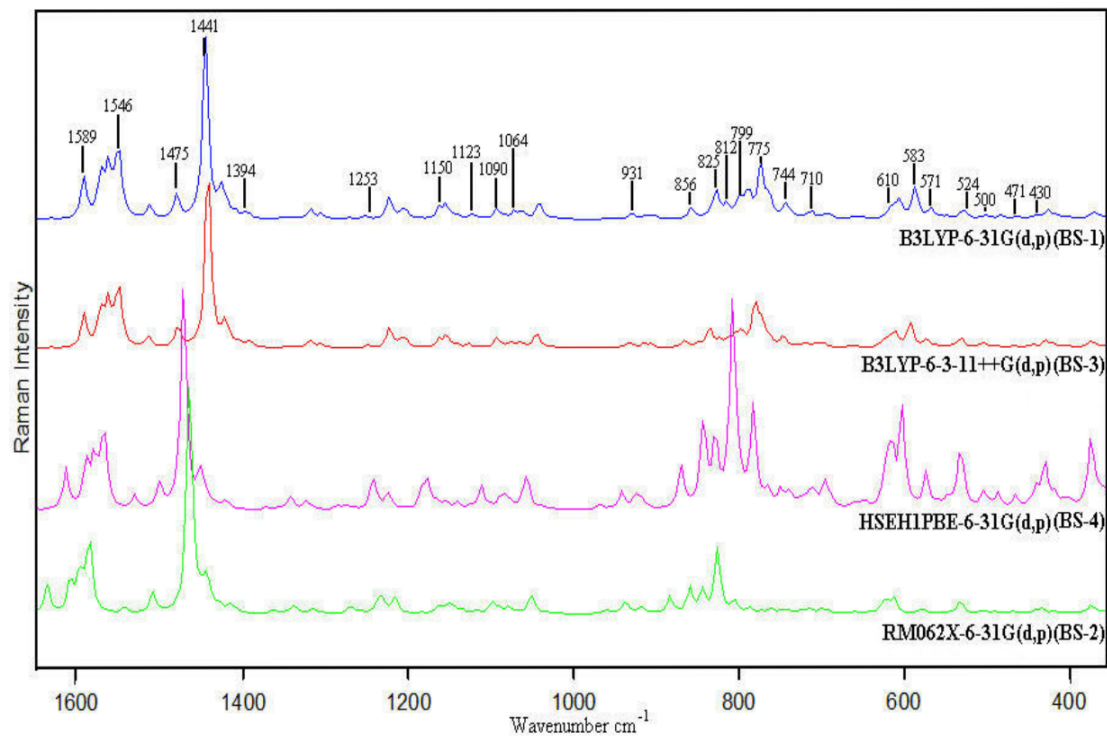


Figure 3: The simulated Raman spectra at the BS-1, BS-2, BS-3 and BS-4 levels for the $C_s-C_{60}Cl_6$ molecule

By considering Table S3 the experimental/calculated values in tetrachloromethane with deuterated chloroform solvent at the BS-1, BS-2, BS-3 and BS-4 levels for C1, C3 and C20 which are connected to the chlorine atoms with high electronegativity of the $C_s-C_{60}Cl_6$ molecule were found at 69.4 ppm/ 76.5, 70.8, 67.8, 89.4 ppm, 55.4 ppm/63.1, 55.2, 54.8, 75.5 ppm and 54.9 ppm/ 62.2, 54.5, 53,9, 74,6 ppm, respectively [48]. Therefore, we can state that the experimental and computed with the BS-4 method for C^{13} NMR features of the cited carbon atoms are in good agreement. Otherwise, the highest C^{13} NMR chemical shift values as experimentally/ theoretically in the mentioned chemical solution were found in the regions 152.8-146.9 ppm/ 169.8- 161.3 ppm at the BS-3 level as seen in Table S3.

3.4. UV-Visible spectroscopy and HOMO-LUMO analysis

The simulated at the BS-1, BS-2 and BS-4 levels of the theorem UV-Vis spectra in the cyclohexane solvent and gas-phase of the $C_s-C_{60}Cl_6$ molecule are handed over in Fig. 5. Furthermore, the calculated λ_{max} worth, f oscillator strengths (vacuum/cyclohexane) and also excitation energies are presented in Table 2. By considering Table 2 and Fig. 5, the resemblance between the experimental and computed λ_{max} values at the cited levels demonstrates that the nearest to experimental values valid for the values calculated with BS-2 method [53].

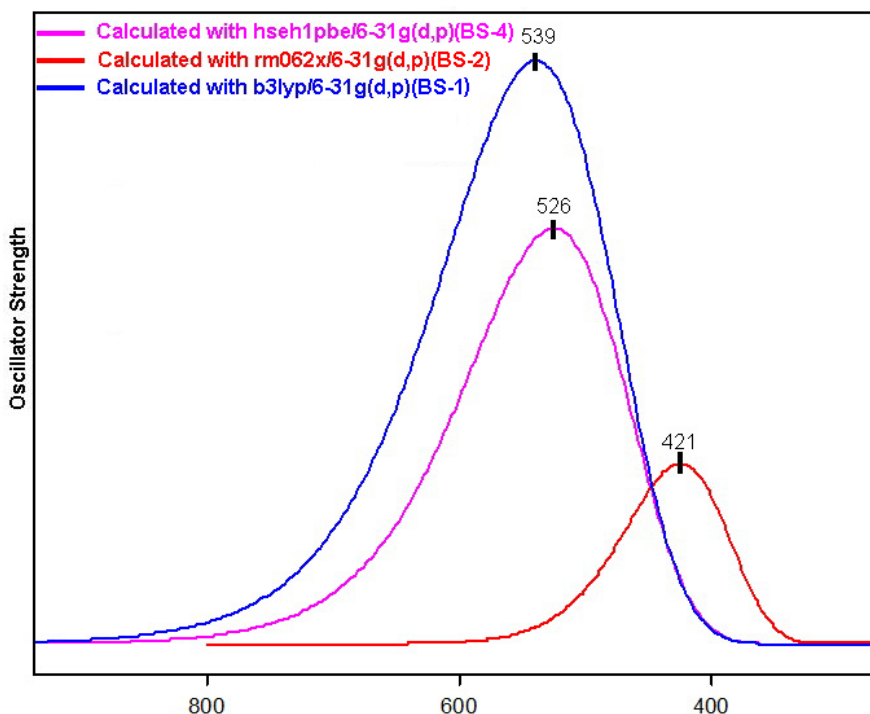


Figure 5: The simulated UV-Vis spectra in the cyclohexane solvent and gas phase at the BS-1, BS-2 and BS-4 levels for the $C_s-C_{60}Cl_6$ molecule

Therefore, the recorded electronic absorption bands at 380 nm, 280 nm and 257 nm in the cyclohexane solvent can correspond to the computed wavelengths at 427.26/428.22 nm, 421.42/421.43 nm and 418.25/418.30 nm in vacuum/cyclohexane at the BS-2 level, respectively which are attributed to the transitions $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$, respectively [53, 55].

However, the deviation among the experimental and computed λ_{\max} worth for the transition $n \rightarrow \pi^*$ is 48.22 nm while this difference values become 141.43 nm and 161.3 nm for the transitions $\pi \rightarrow \pi^*$ which are highly large. For this reason, the UV-Vis spectroscopy for fullerenes derivatives was found as an unreliable method in the past [56]. These differences in the λ_{\max} worth computed at the BS-1 and BS-4 levels of the theory for the C_s - $C_{60}Cl_6$ molecule were found larger than those worth worked out at the BS-2 level. The calculated data can see by looking at Table 2 and Fig. 5.

Table 2: The calculated at the BS-1, BS-2 and BS-4 levels of the theory UV-Vis parameters in the cyclohexane solvent and gas phase of the C_s - $C_{60}Cl_6$ molecule

Experimental*	(BS-1)	(BS-2)	(BS-3)
λ_{\max} (nm) (in cyclohexane)	λ_{\max} (nm) (vacuum/ cyclohexane)	λ_{\max} (nm) (vacuum/ cyclohexane)	λ_{\max} (nm) (vacuum/ cyclohexane)
386	545.40/ 553.69	427.26/ 428.22	532.67/ 537.61
280	534.55/ 539.45	421.42/ 421.43	524.40/ 526.80
257	518.72/ 518.52	418.25/ 418.30	505.28/ 512.82
211	Excitation energies (eV) (vacuum/ cyclohexane)	Excitation energies (eV) (vacuum/ cyclohexane)	Excitation energies (eV) (vacuum/ cyclohexane)
	2.2733/ 2.2392	2.9018/ 2.8953	2.3276/ 2.3062
	2.3194/ 2.2984	2.9420/ 2.9420	2.3643/ 2.3535
	2.3902/ 2.3911	2.9643/ 2.9640	2.4538/ 2.4177
	f (oscillator strengths) (vacuum/ cyclohexane)	f (oscillator strengths) (vacuum/ cyclohexane)	f (oscillator strengths) (vacuum/ cyclohexane)
	0.0002/ 0.0006	0.0001/ 0.0002	0.0001/ 0.0003
	0.0000/ 0.0000	0.0000/ 0.0000	0.0000/ 0.0001
	0.0002/ 0.0004	0.0001/ 0.0001	0.0004/ 0.0003

Taken from ref. [53].

On the other hand, the lowest occupied and highest unoccupied molecular orbitals (LUMO and HOMO) are also called the frontier molecular orbitals (or FMOs) and are played an important role in chemical reaction [57, 58].

The formed energy space among LUMO and HOMO molecular orbitals, which is a crucial parameter, indicates the molecular chemical stability. Furthermore, molecular properties such as electronegativity, kinetic stability, chemical reactivity, polarizability, softness, and chemical hardness, and aromaticity can find out via utilizing this energy space [57, 59].

Figure 6 shows the 3D plots of LUMO-HOMO molecular orbitals for the C_s - $C_{60}Cl_6$ molecule obtained at the BS-1, BS-2, BS-3 and BS-4 levels in the result of our calculation.

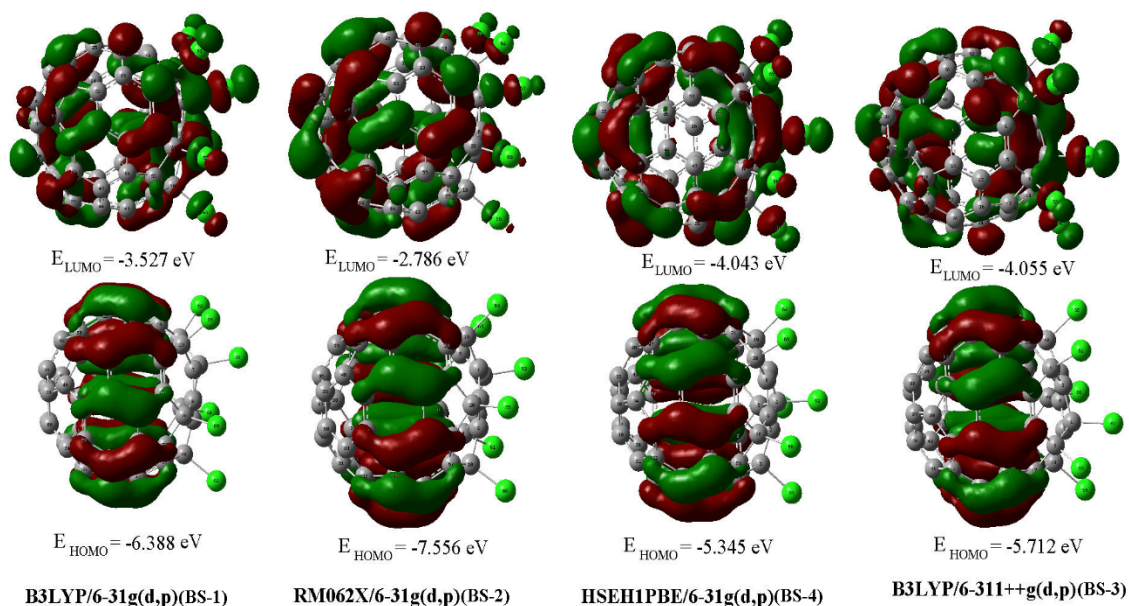


Figure 6: The 3D plots of HOMO-LUMO of the C_s - $C_{60}Cl_6$ molecule obtained at the BS-1, BS-2, BS-3 and BS-4 levels

Table 3: Some molecular properties of C_s - $C_{60}Cl_6$ molecules

Method/Basis set	BS-1	BS-2	BS-3	BS-4	
Electronic Energy (EE)	-5047.325711	-5046.447915	-5044.200344	-5047.944625	Hartree
Zero-point Energy Correction	0.383416	0.391142	0.389969	0.382543	Hartree
Thermal Correction to Energy	0.414375	0.421606	0.420520	0.413441	Hartree
Thermal Correction to Enthalpy	0.415319	0.422550	0.421464	0.414385	Hartree
Thermal Correction to Free Energy	0.328937	0.336738	0.335767	0.328075	Hartree
EE + Zero-point Energy	-5046.942295	-5046.056772	-5043.810375	-5047.562082	Hartree
EE + Thermal Energy Correction	-5046.911337	-5046.026309	-5043.779824	-5047.531184	Hartree
EE + Thermal Enthalpy Correction	-5046.910392	-5046.025365	-5043.778880	-5047.530240	Hartree
EE + Thermal Free Energy Correction	-5046.996775	-5046.111176	-5043.864577	-5047.616550	Hartree
E (Thermal)	260.024	264.562	263.880	259.438	kcal/mol
Heat Capacity (C_v)	146.280	143.598	144.230	146.103	cal/mol-K
Entropy (S)	181.806	180.606	180.365	181.654	cal/mol-K

By using the calculated energy gap values between the HOMO–LUMO levels given in Fig. 6, some molecular properties of the C_s - $C_{60}Cl_6$ compound were indexed in Table 3 [60]. Therefore, the calculated energy gap value at the BS-2 level was found as 4.77 eV as seen in Table 3. However, this value is not in excellent correspondence including the calculated energy value at the BS-2 level of about 3 eV for the transition $n \rightarrow \pi^*$ as can be understood from Table 2. However, these differences are even greater in the other three levels. These results support the view that the UV-Vis technique is not reliable for the fullerene derivatives [56].

3.5. The molecular electrostatic potential (MEP)

So as to understand the molecular affects in structure, the molecular electrostatic potential (MEP) is well-known as a significant instrument. Also, MEP mapping is a valuable tool to study molecular properties, like electronegativity, chemical reactivity, dipole moments, and partial charges of molecules. On the other hand, the MEP has also described as the affect energy among a unit positive test charge and the charge distribution of the molecule [61,62]. Generally, the electrostatic potentials on surfaces are shown in different colors. Consequently, the blue and red painted sections illustrate the zones of positive and negative electrostatic potential while green colored sections demonstrate the zones with zero potential. The negative zones of molecular electrostatic potential are associated with electrophilic reactivity and also the positive zone are associated with nucleophilic reactivity.

In our calculations, we present the 3D plots of the MEP for the C_s - $C_{60}Cl_6$ molecule by using the optimized molecular structures at the BS-1, BS-2, BS-3 and BS-4 levels in Fig. 7.

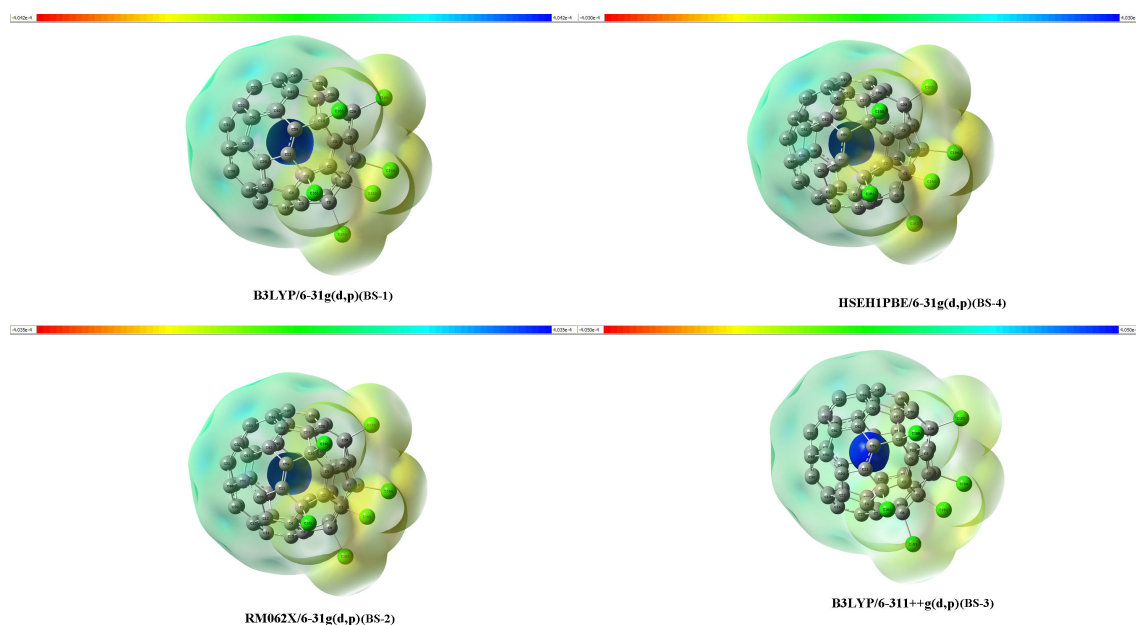


Figure 7: Molecular electrostatic potential (MEP) surface of the C_s - $C_{60}Cl_6$ molecule obtained at the BS-1, BS-2, BS-3 and BS-4 levels.

By considering Fig. 7, the negative regions (with red color or seems to be yellow here) of the MEP plane are confined on the Cl atoms in the C_s - $C_{60}Cl_6$ molecule while the dark blue and blue sites are localized on the carbon atoms that are in the core of the title compound. On the other hand, the green regions are confined to the surroundings of the fullerene cage.

3.6. Thermodynamic parameters

The total energy of molecular structure is the sum of its vibrational, rotational, electronic and translational energies ($E = E_v + E_r + E_e + E_t$). Furthermore, the thermodynamic parameters like rotational constants (GHz), entropy, S , heat capacity, C_v , thermal energy, E , zero-point vibration energy, $ZPVE$, dipole moment can be calculated by using DFT computational methods [63-65]. In the gas phase, the computed thermodynamic parameters of the $C_s-C_{60}Cl_6$ molecule at the BS-1, BS-2, BS-3 and BS-4 levels are donated in Table 4.

According to Table 4, the minimum energy of the $C_s-C_{60}Cl_6$ was found as -5047.944625 Hartree DFT calculations at the BS-3 method. Meanwhile sum calculated thermal energies were realized to be 260.024, 264.552, 263.880, and 259.438 kcal/mol at the BS-1, BS-2, BS-3 and BS-4 levels, respectively.

Table 4: The calculated thermodynamic parameters (in gas phase) of $C_s-C_{60}Cl_6$ molecule

Method/Basis set	BS-1	BS-2	BS-3	BS-4	
Electronic Energy (EE)	-5047.325711	-5046.447915	-5047.944625	-5044.200344	Hartree
Zero-point Energy Correction	0.383416	0.391142	0.382543	0.389969	Hartree
Thermal Correction to Energy	0.414375	0.421606	0.413441	0.420520	Hartree
Thermal Correction to Enthalpy	0.415319	0.422550	0.414385	0.421464	Hartree
Thermal Correction to Free Energy	0.328937	0.336738	0.328075	0.335767	Hartree
EE + Zero-point Energy	-5046.942295	-5046.056772	-5047.562082	-5043.810375	Hartree
EE + Thermal Energy Correction	-5046.911337	-5046.026309	-5047.531184	-5043.779824	Hartree
EE + Thermal Enthalpy Correction	-5046.910392	-5046.025365	-5047.530240	-5043.778880	Hartree
EE + Thermal Free Energy Correction	-5046.996775	-5046.111176	-5047.616550	-5043.864577	Hartree
E (Thermal)	260.024	264.562	259.438	263.880	kcal/mol
Heat Capacity (C_v)	146.280	143.598	146.103	144.230	cal/mol-K
Entropy (S)	181.806	180.606	181.654	180.365	cal/mol-K
Taken from ref. [48].					

4. Conclusion

In the current study, the examination of the stable structure of the $C_s-C_{60}Cl_6$ molecule at the BS-1, BS-2, BS-3 and BS-4 levels of the theory and the stable molecular structure of C_{60} fullerene at the BS-1 level was presented as a supportive work. Similarly, the simulated Raman and Infrared spectra of the $C_s-C_{60}Cl_6$ at the noticed level are in excellent correspondence including the experimental data. The ^{13}C NMR chemical shift values of the $C_s-C_{60}Cl_6$ compound in tetrachloromethane with deuterated chloroform solution and gas-phase were verified at the noticed levels.

The computed chemical shift values in tetrachloromethane with deuterated chloroform solvent at the BS-4 level of the carbon atoms which are connected to the Cl halogen atoms of the $C_s-C_{60}Cl_6$ molecule are in excellent correspondence including the experimental data. The simulated UV-Vis spectra of the $C_s-C_{60}Cl_6$ molecule in the cyclohexane solvent and gas-phase

were computed at the BS-1, BS-2 and BS-4. Furthermore, the MEP and HOMO–LUMO analyses and thermodynamic parameters of the $C_5-C_{60}Cl_6$ molecule at the mentioned levels were also presented.

References

- [1] Kroto, H.W., Heath, J.R., O'Brien, S.C., Curl, R.F. Smalley, R.E., *C₆₀:Buckminsterfullerene*, Nature, 318(6042), 162-163, 1985.
- [2] Ward, J., ed., *The artifacts of R. buckminster fuller: A comprehensive collection of his designs and drawings in four volumes*, New York: Garland, 1984.
- [3] Krätschmer, W., Fostiropoulos, K., Huffman, D.R., *The infrared and ultraviolet absorption spectra of laboratory-produced carbon dust: evidence for the presence of the C₆₀ molecule*, Chemical Physics Letters, 170(2-3), 167-170, 1990.
- [4] Kroto, H.W., Allaf, A.W., Balm, S.P., *C₆₀:Buckminsterfullerene*, Chemical Reviews, 91(6), 1213-1235, 1991.
- [5] Lin, T., Zhang, W.D., Huang, J., He, C.A., *DFT study of the amination of fullerenes and carbon nanotubes: reactivity and curvature*, The Journal of Physical Chemistry B, 109(28), 13755-13760, 2005.
- [6] Yang, C.C., Shen, J.Y., *Well-defined sensing property of ZnO: Al relative humidity sensor with selected buffer layer*, Vacuum, 118, 118-124, 2015.
- [7] Nalwa, H.S., ed. *Handbook of advanced electronic and photonic materials and devices*, ten-volume set, Academic Press, 1 ed., United States, 2000.
- [8] Omacrsawa, E., *Perspectives of fullerene nanotechnology*, Kluwer Academic Publisher, Dordrecht-Boston-London, 2002.
- [9] Bianco A., Da Ros T., Langa. F., Nierengarten J.F., *Biological applications of fullerenes, in fullerenes principles and applications*, Royal Society of Chemistry, 10, 301-328, 2008.
- [10] Hudhomme P., Cousseau, J., Langa. F., Nierengarten J.F., *Plastic solar cells using fullerene derivatives in the photoactive layer, in fullerenes principles and applications*, Royal Society of Chemistry, 8, 221-265, 2008.
- [11] So, H.Y., Wilkins, C.L., *First observation of carbon aggregate ions $>C_{600}^+$ by laser desorption Fourier transform mass spectrometry*, The Journal of Physical Chemistry, 93(4), 1184-1187, 1989.
- [12] Rubin, Y., Kahr, M., Knobler, C.B., Diederich, F., Wilkins, C.L., *The higher oxides of carbon $C_{8n}O_{2n}$ ($n= 3-5$): synthesis, characterization, and X-ray crystal structure. Formation of cyclo $[n]$ carbon ions C_n^+ ($n= 18, 24$), C_n^- ($n= 18, 24, 30$), and higher carbon ions including C_{60}^+ in laser desorption Fourier transform mass spectrometric experiments*, Journal of the American Chemical Society, 113(2), 495-500, 1991.
- [13] Makurin, Y.N., Sofronov, A.A., Gusev, A.I., Ivanovsky, A.L., *Electronic structure and chemical stabilization of C_{28} fullerene*, Chemical Physics, 270(2), 293-308, 2001.

- [14] Lin, M., Chiu, Y.N., Lai, S.T., Xiao, J., Fu, M., *Theoretical study of metallofullerenes $M@C_{32}$* , Journal of Molecular Structure: Theochem, 422(1-3), 57-67, 1998.
- [15] Guo, T., Smalley, R.E., Scuseria, G.E., *Ab initio theoretical predictions of C_{28} , $C_{28}H_4$, $C_{28}F_4$, $(Ti@C_{28})H_4$, and $M@C_{28}$ ($M = Mg, Al, Si, S, Ca, Sc, Ti, Ge, Zr, \text{ and } Sn$)*, The Journal of Chemical Physics, 99(1), 352-359, 1993.
- [16] Kadish, K.M., Ruoff, R.S., *Fullerenes: chemistry, physics, and technology*, John Wiley & Sons, Newyork, 2000.
- [17] Prinzbach, H., Weiler, A., Landenberger, P., Wahl, F., Wörth, J., Scott, L.T., Issendorff, B.V., *Gas-phase production and photoelectron spectroscopy of the smallest fullerene, C_{20}* , Nature, 407(6800), 60-63, 2000.
- [18] Priimagi, A., Cavallo, G., Metrangolo, P., Resnati, G., *The halogen bond in the design of functional supramolecular materials: recent advances*, Accounts of Chemical Research, 46(11), 2686-2695, 2013.
- [19] Popov A.A., Senyavin V.M., and Granovsky A.A., *Vibrational spectra of chloro- and bromofullerenes*, Fullerenes, Nanotubes, And Carbon Nanostructures, Marcel Dekker, 12, 305–310, 2004.
- [20] Troshin, P.A., Lyubovskaya, R.N., Ioffe, I.N., Shustova, N.B., Kemnitz, E., Troyanov, S.I., *Synthesis and structure of the highly chlorinated [60]Fullerene $C_{60}Cl_{30}$ with a drum-shaped carbon cage*, Angewandte Chemie International Edition, 44(2), 234-237, 2005.
- [21] Kuvychko, I.V., Streletskii, A.V., Popov, A.A., Kotsiris, S.G., Drewello, T., Strauss, S.H., Boltalina, O.V., *Seven-minute synthesis of pure Cs- $C_{60}Cl_6$ from [60] Fullerene and iodine monochloride: First IR, Raman, and Mass spectra of 99 mol% $C_{60}Cl_6$* , Chemistry A European Journal, 11(18), 5426-5436, 2005.
- [22] Yan, Q.B., Zheng, Q.R., Su, G., *Theoretical study on the structures, properties and spectroscopies of fullerene derivatives $C_{66}X_4$ ($X = H, F, Cl$)*, Carbon, 45(9), 1821-1827, 2007.
- [23] Troyanov, S.I., Boltalina, O.V., Kouvytchko, I.V., Troshin, P.A., Kemnitz, E., Hitchcock, P.B., Taylor, R., *Molecular and crystal structure of the adducts of $C_{60}F_{18}$ with aromatic hydrocarbons*, Fullerenes, Nanotubes And Carbon Nanostructures, 10(3), 243-259, 2002.
- [24] Adjizian, J.J., Vlandas, A., Rio, J., Charlier, J.C., Ewels, C.P., *Ab initio infrared vibrational modes for neutral and charged small fullerenes (C_{20} , C_{24} , C_{26} , C_{28} , C_{30} and C_{60})*, Philosophical Transactions Of The Royal Society A: Mathematical, Physical And Engineering Sciences, 374(2076), 20150323, 2016.
- [25] Yang, T., Zhao, X., Nagase, S., *Di-lanthanide encapsulated into large fullerene C_{100} : a DFT survey*, Physical Chemistry Chemical Physics, 13(11), 5034-5037, 2011.
- [26] Carter, E.A., Rossky, P.J., *Computational and theoretical chemistry*, Accounts of Chemical Research, 39(2), 71-72, 2006.
- [27] Bauernschmitt, R., Ahlrichs, R., Hennrich, F.H., Kappes, M.M., *Experiment versus time dependent density functional theory prediction of fullerene electronic absorption*, Journal of the American Chemical Society, 120(20), 5052-5059, 1998.

- [28] Schettino, V., Pagliai, M., Cardini, G., *The infrared and Raman spectra of fullerene C₇₀. DFT calculations and correlation with C₆₀*, *The Journal of Physical Chemistry A*, 106(9), 1815-1823, 2002.
- [29] Jensen, F. *Introduction to computational chemistry*, John Wiley & Sons, Newyork, 664p., 3 ed., 1974.
- [30] Pulay, P., *Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. Theory*, *Molecular Physics*, 17(2), 197-204, 1969.
- [31] Perdew, J.P., "Electronic Structure of Solids", in *Proceeding of the 21st Annual International Symposium*, p11, 1991.
- [32] Perdew, J. P., Wang, Y., *Pair-distribution function and its coupling-constant average for the spin-polarized electron gas*, *Physical Review B*, 46(20), 12947, 1992.
- [33] Becke, A.D., *A new mixing of Hartree-Fock and local density-functional theories*, *The Journal of Chemical Physics*, 98(2), 1372-1377, 1993.
- [34] Lee, C., Yang, W., Parr, R.G., *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, *Physical Review B*, 37(2), 785, 1988.
- [35] Burke, K., Perdew, J.P., Wang, Y., *Derivation of a Generalized Gradient Approximation: The PW₉₁ Density Functional*, *Electronic Density Functional Theory*, Springer, Boston, MA., 81-111, 1998.
- [36] Zhao, Y., Pu, J., Lynch, B. J., Truhlar, D.G., *Tests of second-generation and third-generation density functionals for thermochemical kinetics*, *Physical Chemistry Chemical Physics*, 6(4), 673-676, 2004.
- [37] Heyd, J., Scuseria, G.E., Ernzerhof, M., *Hybrid functionals based on a screened Coulomb potential*, *The Journal of Chemical Physics*, 118(18), 8207-8215, 2006.
- [38] Wolinski, K., Hinton, J.F., Pulay, P., *Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations*, *Journal of the American Chemical Society*, 112(23), 8251-8260, 1990.
- [39] Avci, D., Dede, B., Bahceli, S., Varkal, D., *Spectroscopic and quantum chemical calculation study on 2-ethoxythiazole molecule*, *Journal of Molecular Structure*, 1138, 110-117, 2017.
- [40] Heyd, J., Scuseria, G.E., Ernzerhof, M., *Hybrid functionals based on a screened Coulomb potential*, *The Journal of Chemical Physics*, 118(18), 8207-8215, 2003.
- [41] Zhao, Y., Truhlar, D.G., *Comparative DFT study of van der Waals complexes: Rare-Gas Dimers, Alkaline-Earth Dimers, Zinc Dimer, And Zinc-Rare-Gas Dimers*, *The Journal of Physical Chemistry A*, 110(15), 5121-5129, 2006.
- [42] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., ... Fox, D.J., *Gaussian 09*, Gaussian, Inc., Wallingford CT, 2009.
- [43] Dennington, R., Keith, T., Millam, J., *GaussView Version 5*. Semichem Inc., Shawnee Mission, Kans, 2009.
- [44] Foresman, J.B., Frisch., E., *Exploring chemistry with electronic structure methods*, Gaussian Inc., Pittsburgh, Pa., USA, 1993.

- [45] London, F. *Théorie quantique des courants interatomiques dans les combinaisons aromatiques*, Journal de Physique et le Radium, 8(10), 397-409, 1937.
- [46] Bauernschmitt, R., Ahlrichs, R., *Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory*, Chemical Physics Letters, 256(4-5), 454-464, 1996.
- [47] Jamorski, C., Casida, M.E., Salahub, D.R., *Dynamic polarizabilities and excitation spectra from a molecular implementation of time-dependent density-functional response theory: N₂ as a case study*, The Journal of Chemical Physics, 104(13), 5134-5147, 1996.
- [48] Birkett, P.R., Hitchcock, P.B., Kroto, H.W., Taylor, R., Walton, D.R., *Preparation and characterization of C₆₀Br₆ and C₆₀Br₈*, Nature, 357(6378), 479-481, 1992.
- [49] Birkett, P.R., Avent, A.G., Darwish, A.D., Kroto, H.W., Taylor, R., Walton, D.R., *Holey fullerenes! a bis-lactone derivative of [70 fullerene with an eleven-atom orifice*, Journal of the Chemical Society, Chemical Communications, 18, 1869-1870, 1995.
- [50] Politzer, P., Murray, J.S., Clark, T., *Halogen bonding and other σ -hole interactions: a perspective*, Physical Chemistry Chemical Physics, 15(27), 11178-11189, 2013.
- [51] Troyanov, S.I., Troshin, P.A., Boltalina, O.V., Kemnitz, E., *Bromination of [60]Fullerene. II. Crystal and molecular structure of [60]Fullerene bromides, C₆₀Br₆, C₆₀Br₈, and C₆₀Br₂₄*, Fullerenes, Nanotubes and Carbon Nanostructures, 11(1), 61-77, 2003.
- [52] Fedurco, M., Olmstead, M.M., Fawcett, W.R., *Single-crystal X-ray structure of C₆₀•6SbPh₃. A well-ordered structure of C₆₀ and a new fullerene solvent*, Inorganic Chemistry, 34(1), 390-392, 1995.
- [53] Birkett, P.R., Avent, A.G., Darwish, A.D., Kroto, H.W., Taylor, R., Walton, D.R.M., *Preparation and ¹³C NMR spectroscopic characterisation of C₆₀Cl₆*, Journal of the Chemical Society, Chemical Communications, 15, 1230-1232, 1993.
- [54] Kuvychko, I.V., Streletskii, A.V., Shustova, N.B., Seppelt, K., Drewello, T., Popov, A.A., Boltalina, O.V., *Soluble Chlorofullerenes C₆₀Cl_{2,4,6,8,10}. Synthesis, purification, compositional analysis, stability, and experimental/theoretical structure elucidation, including the X-ray structure of C₁-C₆₀Cl₁₀*, Journal of the American Chemical Society, 132(18), 6443-6462, 2010.
- [55] Süleymanoğlu, N., Ustabaş, R., Alpaslan, Y.B., Eydurhan, F., İskeleli, N.O., *Experimental and theoretical investigation of the molecular and electronic structure of 3-ethoxy-4-isopropylaminocyclobut-3-ene-1,2-dione*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 96, 35-41, 2012.
- [56] Popov, A.A., Kareev, I.E., Shustova, N.B., Stukalin, E.B., Lebedkin, S.F., Seppelt, K., Dunsch, L., *Electrochemical, spectroscopic, and DFT study of C₆₀(CF₃)_n frontier orbitals (n= 2–18): the link between double bonds in pentagons and reduction potentials*, Journal of the American Chemical Society, 129(37), 11551-11568, 2007.
- [57] Pearson, R.G., *Absolute electronegativity and hardness correlated with molecular orbital theory*, Proceedings of the National Academy of Sciences, 83(22), 8440-8441, 1986.
- [58] Lee, C.K., Kim, Y.H. *Spectroscopic studies of conjugated uracil derivatives*, Bulletin of the Korean Chemical Society, 12(2), 207-210, 1991.

[59] Fukui, K., *Role of frontier orbitals in chemical reactions*, Science, 218(4574), 747-754, 1982.

[60] Chtita, S., Ghamali, M., Larif, M., Adad, A., Hmammouchi, R., Bouachrine, M., Lakhlifi, T., *Prediction of biological activity of imidazo [1,2-a] pyrazine derivatives by combining DFT and QSAR results*, International Journal of Innovative Research in Science, Engineering and Technology, 2(11), 7951-7962, 2013.

[61] Murray, J.S., Sen K. eds., *Molecular electrostatic potentials concepts and applications*, Elsevier Science BV, Amsterdam, The Netherlands, 1996.

[62] Pîrnău, A., Chiş, V., Oniga, O., Leopold, N., Szabo, L., Baias, M., nCozar, O., *Vibrational and DFT study of 5-(3-pyridyl-methylidene)-thiazolidine-2-thione-4-one*, Vibrational spectroscopy, 48(2), 289-296, 2008.

[63] Sarıkaya, E.K., Bahçeli, S., Varkal, D., Dereli, Ö., *FT-IR, micro-Raman and UV-vis spectroscopic and quantum chemical calculation studies on the 6-chloro-4-hydroxy-3-phenyl pyridazine compound*, Journal of Molecular Structure, 1141, 44-52, 2017.

[64] Yilmaz, M., Aydin, B., Dogan, O., Dereli, O., *Molecular structure and spectral investigations of 3,5-Di-tert-butyl-o-benzoquinone*, Journal of Molecular Structure, 1128, 345-354, 2017.

[65] Fankam, J.B., Ejuh, G.W., Tchangnwa N.F., Ndjaka, J.M.B., *Theoretical investigation of the molecular structure, vibrational spectra, thermodynamic and nonlinear optical properties of 4,5-dibromo-2,7-dinitro-fluorescein*, Optical and Quantum Electronics, 52, 1-23, 2020.

SUPPLEMENTARY INFORMATION FILE



Adıyaman University
Journal of Science

<https://dergipark.org.tr/en/pub/adyujsci>

DergiPark
AKADEMİK

ISSN 2147-1630
e-ISSN 2146-586X

A Comparative Study of DFT/B3LYP/6-31G(d,p), RM062X/6-31G(d,p), B3LYP/6-311++G(d,p) and HSEH1PBE/6-31G(d,p) Methods Applied to Molecular Geometry and Electronic properties of C_s - $C_{60}Cl_6$

Molecule

Ebru KARAKAŞ SARIKAYA^{1,*}, Ömer DERELİ², Semiha BAHÇELİ³

¹*Necmettin Erbakan of University, Engineering Faculty, Department of Basic Sciences, Konya, Türkiye
ebrukarakas_84@hotmail.com, ORCID: 0000-0003-2149-9341*

²*Necmettin Erbakan of University, Faculty of A. K. Education, Department of Physics, Konya, Türkiye
odereli@erbakan.edu.tr, ORCID: 0000-0002-9031-8092*

³*Emeritus Professor of Atomic and Molecular Physics, 06000, Ankara, Türkiye
sbahceli@thk.edu.tr, ORCID: 0000-0002-5614-325X*



Table S1. The calculated geometric parameters at the B3LYP/6-31G(d,p) level of C₆₀ fullerene molecule. Bond lengths in Angstrom (Å) and angles in Degrees (°).

Bond Length		Bond Angles		Bond Angles		Dihedral Angles		Dihedral Angles	
C(60)-C(34)	1.395	C(35)-C(60)-C(30)	107.996	C(12)-C(4)-C(3)	107.993	C(30)-C(60)-C(34)	138.191	C(13)-C(39)-C(38)	0.001
C(60)-C(30)	1.453	C(35)-C(60)-C(34)	119.983	C(2)-C(3)-C(11)	119.998	C(30)-C(60)-C(34)	0.002	C(39)-C(38)-C(37)	0.007
C(60)-C(35)	1.453	C(30)-C(60)-C(34)	120.004	C(2)-C(3)-C(4)	120.007	C(35)-C(60)-C(34)	0.029	C(39)-C(38)-C(37)	138.211
C(59)-C(42)	1.453	C(10)-C(59)-C(11)	120.002	C(11)-C(3)-C(4)	108.017	C(35)-C(60)-C(34)	-138.160	C(28)-C(38)-C(37)	-138.181
C(59)-C(11)	1.396	C(10)-C(59)-C(42)	108.008	C(3)-C(2)-C(9)	120.006	C(34)-C(60)-C(30)	-142.614	C(28)-C(38)-C(37)	0.023
C(59)-C(10)	1.453	C(11)-C(59)-C(42)	120.006	C(3)-C(2)-C(1)	120.008	C(34)-C(60)-C(30)	0.028	C(39)-C(38)-C(28)	-0.007
C(58)-C(41)	1.453	C(29)-C(58)-C(40)	119.991	C(9)-C(2)-C(1)	108.010	C(35)-C(60)-C(30)	-0.021	C(39)-C(38)-C(28)	-142.663
C(58)-C(40)	1.453	C(29)-C(58)-C(41)	120.002	C(6)-C(1)-C(7)	120.004	C(35)-C(60)-C(30)	142.621	C(37)-C(38)-C(28)	142.612
C(58)-C(29)	1.396	C(40)-C(58)-C(41)	108.003	C(6)-C(1)-C(2)	119.983	C(34)-C(60)-C(35)	142.633	C(37)-C(38)-C(28)	-0.044
C(57)-C(43)	1.453	C(56)-C(57)-C(43)	119.996	C(7)-C(1)-C(2)	107.996	C(34)-C(60)-C(35)	-0.024	C(38)-C(37)-C(30)	142.630
C(57)-C(56)	1.396	C(56)-C(57)-C(10)	119.985	C(33)-C(34)-C(46)	108.001	C(30)-C(60)-C(35)	0.031	C(38)-C(37)-C(30)	-0.012
C(57)-C(10)	1.453	C(43)-C(57)-C(10)	108.002	C(33)-C(34)-C(60)	119.989	C(30)-C(60)-C(35)	-142.626	C(36)-C(37)-C(30)	0.004
C(56)-C(55)	1.453	C(20)-C(56)-C(55)	107.999	C(46)-C(34)-C(60)	120.009	C(11)-C(59)-C(42)	-142.647	C(36)-C(37)-C(30)	-142.638
C(56)-C(20)	1.453	C(20)-C(56)-C(57)	120.010	C(32)-C(33)-C(34)	120.004	C(11)-C(59)-C(42)	-0.012	C(30)-C(37)-C(36)	0.015
C(55)-C(45)	1.395	C(55)-C(56)-C(57)	120.013	C(32)-C(33)-C(44)	120.016	C(10)-C(59)-C(42)	-0.004	C(30)-C(37)-C(36)	142.627
C(55)-C(54)	1.453	C(45)-C(55)-C(56)	119.993	C(34)-C(33)-C(44)	108.011	C(10)-C(59)-C(42)	142.632	C(38)-C(37)-C(36)	-142.617
C(54)-C(53)	1.396	C(45)-C(55)-C(54)	120.019	C(33)-C(32)-C(41)	119.999	C(42)-C(59)-C(11)	138.176	C(38)-C(37)-C(36)	-0.005
C(54)-C(21)	1.453	C(56)-C(55)-C(54)	108.010	C(33)-C(32)-C(31)	120.013	C(42)-C(59)-C(11)	0.019	C(49)-C(36)-C(35)	-142.644
C(53)-C(47)	1.453	C(21)-C(54)-C(53)	120.009	C(41)-C(32)-C(31)	108.012	C(10)-C(59)-C(11)	-0.039	C(49)-C(36)-C(35)	0.014
C(53)-C(52)	1.453	C(21)-C(54)-C(55)	107.993	C(30)-C(31)-C(40)	119.988	C(10)-C(59)-C(11)	-138.197	C(37)-C(36)-C(35)	-0.028
C(52)-C(51)	1.453	C(53)-C(54)-C(55)	119.999	C(30)-C(31)-C(32)	119.979	C(42)-C(59)-C(10)	-0.004	C(37)-C(36)-C(35)	142.629
C(52)-C(23)	1.396	C(52)-C(53)-C(47)	107.991	C(40)-C(31)-C(32)	107.994	C(42)-C(59)-C(10)	-142.623	C(60)-C(34)-C(33)	142.657
C(51)-C(48)	1.453	C(52)-C(53)-C(54)	119.992	C(31)-C(30)-C(30)	120.011	C(11)-C(59)-C(10)	142.642	C(60)-C(34)-C(33)	-0.013
C(51)-C(50)	1.396	C(47)-C(53)-C(54)	119.979	C(31)-C(30)-C(67)	120.010	C(11)-C(59)-C(10)	0.022	C(46)-C(34)-C(33)	0.031
C(50)-C(49)	1.453	C(23)-C(52)-C(51)	120.001	C(60)-C(30)-C(37)	107.992	C(40)-C(58)-C(41)	-142.613	C(46)-C(34)-C(33)	-142.639
C(50)-C(25)	1.453	C(23)-C(52)-C(53)	120.007	C(12)-C(29)-C(11)	108.014	C(40)-C(58)-C(41)	-0.018	C(44)-C(33)-C(32)	-138.245
C(49)-C(36)	1.396	C(51)-C(52)-C(53)	108.014	C(12)-C(29)-C(58)	120.007	C(29)-C(58)-C(41)	0.001	C(44)-C(33)-C(32)	-0.015
C(49)-C(27)	1.453	C(50)-C(51)-C(48)	120.001	C(11)-C(29)-C(58)	120.001	C(29)-C(58)-C(41)	142.596	C(34)-C(33)-C(32)	-0.005
C(48)-C(47)	1.453	C(50)-C(51)-C(52)	119.995	C(27)-C(28)-C(38)	120.010	C(41)-C(58)-C(40)	-0.015	C(34)-C(33)-C(32)	138.226
C(48)-C(35)	1.395	C(48)-C(51)-C(52)	107.986	C(27)-C(28)-C(14)	120.013	C(41)-C(58)-C(40)	142.630	C(41)-C(32)-C(31)	-0.053
C(47)-C(46)	1.396	C(25)-C(50)-C(49)	108.007	C(38)-C(28)-C(14)	107.999	C(29)-C(58)-C(40)	-142.633	C(41)-C(32)-C(31)	-142.620
C(46)-C(45)	1.453	C(25)-C(50)-C(51)	120.006	C(26)-C(27)-C(28)	119.996	C(29)-C(58)-C(40)	0.012	C(33)-C(32)-C(31)	142.602
C(46)-C(34)	1.453	C(49)-C(50)-C(51)	120.002	C(26)-C(27)-C(49)	108.002	C(41)-C(58)-C(29)	0.006	C(33)-C(32)-C(31)	0.034
C(45)-C(44)	1.453	C(36)-C(49)-C(50)	119.994	C(28)-C(27)-C(49)	119.985	C(41)-C(58)-C(29)	-138.222	C(32)-C(31)-C(30)	138.159
C(44)-C(33)	1.453	C(36)-C(49)-C(27)	120.003	C(16)-C(26)-C(27)	119.990	C(40)-C(58)-C(29)	138.190	C(32)-C(31)-C(30)	-0.046
C(44)-C(43)	1.396	C(50)-C(49)-C(27)	108.000	C(16)-C(26)-C(25)	119.994	C(40)-C(58)-C(29)	-0.038	C(40)-C(31)-C(30)	0.030
C(43)-C(42)	1.453	C(35)-C(48)-C(51)	119.999	C(27)-C(26)-C(25)	107.996	C(10)-C(57)-C(43)	-0.012	C(40)-C(31)-C(30)	-138.175
C(42)-C(41)	1.395	C(35)-C(48)-C(47)	120.007	C(24)-C(25)-C(50)	119.995	C(10)-C(57)-C(43)	-142.605	C(58)-C(29)-C(11)	-142.615
C(41)-C(32)	1.453	C(51)-C(48)-C(47)	108.017	C(24)-C(25)-C(26)	120.015	C(56)-C(57)-C(43)	142.582	C(58)-C(29)-C(11)	-0.016
C(40)-C(31)	1.453	C(46)-C(47)-C(53)	120.020	C(50)-C(25)-C(26)	107.995	C(56)-C(57)-C(43)	-0.011	C(12)-C(29)-C(11)	0.040
C(40)-C(39)	1.395	C(46)-C(47)-C(48)	119.986	C(18)-C(24)-C(23)	107.993	C(10)-C(57)-C(56)	138.177	C(12)-C(29)-C(11)	142.639
C(39)-C(13)	1.453	C(53)-C(47)-C(48)	107.993	C(18)-C(24)-C(25)	119.988	C(10)-C(57)-C(56)	-0.045	C(58)-C(29)-C(12)	0.055
C(39)-C(38)	1.453	C(34)-C(46)-C(45)	107.992	C(23)-C(24)-C(25)	120.001	C(43)-C(57)-C(56)	0.012	C(58)-C(29)-C(12)	142.618
C(38)-C(37)	1.395	C(34)-C(46)-C(47)	120.006	C(24)-C(23)-C(22)	120.002	C(43)-C(57)-C(56)	-138.210	C(11)-C(29)-C(12)	-142.597
C(38)-C(28)	1.453	C(45)-C(46)-C(47)	119.995	C(24)-C(23)-C(52)	108.003	C(43)-C(57)-C(10)	0.010	C(11)-C(29)-C(12)	-0.034
C(37)-C(30)	1.453	C(46)-C(45)-C(55)	119.988	C(52)-C(23)-C(22)	119.991	C(43)-C(57)-C(10)	142.625	C(38)-C(28)-C(14)	0.010
C(37)-C(36)	1.453	C(46)-C(45)-C(44)	108.013	C(19)-C(22)-C(21)	120.005	C(56)-C(57)-C(10)	-142.589	C(38)-C(28)-C(14)	-142.650
C(36)-C(35)	1.453	C(55)-C(45)-C(44)	120.001	C(19)-C(22)-C(23)	107.998	C(56)-C(57)-C(10)	0.026	C(27)-C(28)-C(14)	142.665
C(34)-C(33)	1.453	C(43)-C(44)-C(33)	119.988	C(21)-C(22)-C(23)	120.013	C(57)-C(56)-C(55)	-142.665	C(27)-C(28)-C(14)	0.005
C(33)-C(32)	1.395	C(43)-C(44)-C(45)	120.007	C(22)-C(21)-C(54)	119.988	C(57)-C(56)-C(55)	-0.005	C(14)-C(28)-C(27)	-138.177
C(32)-C(31)	1.453	C(33)-C(44)-C(45)	107.983	C(22)-C(21)-C(20)	120.001	C(20)-C(56)-C(55)	-0.010	C(14)-C(28)-C(27)	-0.012
C(31)-C(30)	1.396	C(44)-C(43)-C(57)	119.990	C(54)-C(21)-C(20)	108.000	C(20)-C(56)-C(55)	142.650	C(38)-C(28)-C(27)	0.045
C(29)-C(11)	1.453	C(44)-C(43)-C(42)	119.993	C(8)-C(20)-C(56)	120.006	C(57)-C(56)-C(20)	142.663	C(38)-C(28)-C(27)	138.210
C(29)-C(12)	1.453	C(57)-C(43)-C(42)	107.996	C(8)-C(20)-C(21)	119.995	C(57)-C(56)-C(20)	0.044	C(49)-C(27)-C(26)	0.012
C(28)-C(14)	1.453	C(41)-C(42)-C(59)	119.995	C(56)-C(20)-C(21)	107.998	C(55)-C(56)-C(20)	0.007	C(49)-C(27)-C(26)	142.605
C(28)-C(27)	1.396	C(41)-C(42)-C(43)	120.015	C(7)-C(19)-C(22)	119.988	C(55)-C(56)-C(20)	-142.612	C(28)-C(27)-C(26)	-142.582
C(27)-C(26)	1.453	C(59)-C(42)-C(43)	107.995	C(7)-C(19)-C(18)	119.979	C(54)-C(55)-C(45)	138.223	C(28)-C(27)-C(26)	0.011
C(26)-C(16)	1.396	C(32)-C(41)-C(42)	119.988	C(22)-C(19)-C(18)	107.994	C(54)-C(55)-C(45)	0.025	C(25)-C(26)-C(16)	138.155
C(26)-C(25)	1.453	C(32)-C(41)-C(58)	107.993	C(17)-C(18)-C(24)	119.999	C(56)-C(55)-C(45)	-0.004	C(25)-C(26)-C(16)	0.006
C(25)-C(24)	1.395	C(42)-C(41)-C(58)	120.001	C(17)-C(18)-C(19)	120.014	C(56)-C(55)-C(45)	-138.202	C(27)-C(26)-C(16)	-0.002
C(24)-C(23)	1.453	C(39)-C(40)-C(31)	120.005	C(24)-C(18)-C(19)	108.011	C(56)-C(55)-C(54)	142.636	C(27)-C(26)-C(16)	-138.151
C(24)-C(18)	1.453	C(39)-C(40)-C(58)	120.013	C(6)-C(17)-C(18)	108.011	C(56)-C(55)-C(54)	0.010	C(27)-C(26)-C(25)	-0.010
C(23)-C(22)	1.453	C(31)-C(40)-C(58)	107.998	C(6)-C(17)-C(18)	120.004	C(45)-C(55)-C(54)	-0.013	C(27)-C(26)-C(25)	142.617
C(22)-C(21)	1.395	C(38)-C(39)-C(13)	108.000	C(16)-C(17)-C(18)	120.016	C(45)-C(55)-C(54)	-142.639	C(16)-C(26)-C(25)	-142.602
C(22)-C(19)	1.453	C(38)-C(39)-C(40)	120.001	C(17)-C(16)-C(26)	119.988	C(55)-C(54)-C(53)	-0.014	C(16)-C(26)-C(25)	0.025
C(21)-C(20)	1.453	C(13)-C(39)-C(40)	119.988	C(17)-C(16)-C(15)	107.983	C(55)-C(54)-C(53)	-138.143	C(26)-C(25)-C(24)	-138.198
C(20)-C(8)	1.395	C(28)-C(38)-C(37)	120.006	C(26)-C(16)-C(15)	120.007	C(21)-C(54)-C(53)	138.173	C(26)-C(25)-C(24)	-0.036
C(19)-C(7)	1.396	C(28)-C(38)-C(39)	107.998	C(5)-C(15)-C(14)	119.988	C(21)-C(54)-C(53)	0.045	C(50)-C(25)-C(24)	-0.002
C(19)-C(18)	1.453	C(37)-C(38)-C(39)	119.995	C(5)-C(15)-C(16)	108.013	C(55)-C(54)-C(21)	-0.005	C(50)-C(25)-C(24)	138.159
C(18)-C(17)	1.395	C(36)-C(37)-C(38)	119.989	C(14)-C(15)-C(16)	120.001	C(55)-C(54)-C(21)	142.603	C(25)-C(24)-C(23)	142.613
C(17)-C(16)	1.453	C(36)-C(37)-C(30)	108.015	C(15)-C(14)-C(28)	119.993	C(53)-C(54)-C(21)	-142.627	C(25)-C(24)-C(23)	-0.001
C(16)-C(15)	1.453	C(38)-C(37)-C(30)	120.001	C(15)-C(14)-C(13)	120.019	C(53)-C(54)-C(21)	-0.019	C(18)-C(24)-C(23)	0.018
C(15)-C(14)	1.395	C(37)-C(36)-C(49)	120.007	C(28)-C(14)-C(13)	108.010	C(54)-C(53)-C(47)	-142.584	C(18)-C(24)-C(23)	-142.596
C(15)-C(5)	1.453	C(37)-C(36)-C(35)	107.987	C(12)-C(13)-C(39)	120.009	C(54)-C(53)-C(47)	0.029	C(25)-C(24)-C(18)	-142.644
C(14)-C(13)	1.453	C(49)-C(36)-C(35)	119.999	C(12)-C(13)-C(14)	119.998	C(52)-C(53)-C(47)	-0.015	C(25)-C(24)-C(18)	0.017
C(13)-C(12)	1.396	C(48)-C(35)-C(60)	120.008	C(39)-C(13)-C(14)	107.993	C(52)-C(53)-C(47)	142.598	C(23)-C(24)-C(18)	-0.044
C(12)-C(4)	1.453	C(48)-C(35)-C(36)	120.006	C(4)-C(12)-C(13)	119.979	C(54)-C(53)-C(52)	142.597	C(23)-C(24)-C(18)	142.617
C(11)-C(3)	1.453	C(60)-C(35)-C(36)	108.010	C(4)-C(12)-C(22)	107.991	C(54)-C(53)-C(52)	-0.055	C(52)-C(23)-C(22)	-0.012
C(10)-C(9)	1.396	C(6)-C(5)-C(4)-C(3)	0.010	C(52)-C(23)-C(22)	142.633	C(49)-C(50)-C(25)	-142.632	C(7)-C(19)-C(18)	142.620
C(9)-C(8)	1.453	C(6)-C(5)-C(4)-C(12)	138.194	C(24)-C(23)-C(22)	-142.630	C(27)-C(49)-C(36)	138.180	C(7)-C(19)-C(18)	-0.034
C(9)-C(2)	1.453	C(15)-C(5)-C(4)-C(3)	-138.167	C(24)-C(23)-C(21)	0.015	C(27)-C(49)-C(36)	0.006	C(19)-C(18)-C(17)	138.245
C(8)-C(7)	1.453	C(15)-C(5)-C(4)-C(12)	0.017	C(23)-C(22)-C(21)	138.174	C(27)-C(49)-C(36)	-0.005	C(19)-C(18)-C(17)	0.005
C(7)-C(1)	1.453	C(12)-C(4)-C(3)-C(11)	0.009	C(23)-C(22)-C(21)	0.002	C(50)-C(49)-C(36)	-138.179	C(24)-C(18)-C(17)	0.015
		C(12)-C(4)-C(3)-C(2)	-142.						

C(6)-C(1)	1.395	C(5)-C(4)-C(3)-C(11)	142.638	C(19)-C(22)-C(21)-	-138.211	C(50)-C(49)-C(27)-	-0.010	C(18)-C(17)-C(16)-	-142.621
C(6)-C(5)	1.453	C(5)-C(4)-C(3)-C(2)	-0.015	C(23)-C(22)-C(19)-	-0.042	C(36)-C(49)-C(27)-	-0.026	C(18)-C(17)-C(16)-	-0.026
C(5)-C(4)	1.396	C(4)-C(3)-C(2)-C(1)	-0.001	C(23)-C(22)-C(19)-	-142.605	C(36)-C(49)-C(27)-	-142.625	C(6)-C(17)-C(16)-	0.044
C(4)-C(3)	1.453	C(4)-C(3)-C(2)-C(9)	138.228	C(21)-C(22)-C(19)-	142.607	C(51)-C(48)-C(47)-	-0.009	C(6)-C(17)-C(16)-	142.639
C(3)-C(2)	1.395	C(11)-C(3)-C(2)-C(1)	-138.233	C(21)-C(22)-C(19)-	0.044	C(51)-C(48)-C(47)-	-142.638	C(18)-C(17)-C(6)-	142.639
C(2)-C(1)	1.453	C(11)-C(3)-C(2)-C(9)	-0.003	C(54)-C(21)-C(20)-	-0.001	C(35)-C(48)-C(47)-	142.644	C(18)-C(17)-C(6)-	0.013
C(13)-C(12)-C(29)	119.992	C(9)-C(2)-C(1)-C(7)	-0.031	C(54)-C(21)-C(20)-	142.623	C(35)-C(48)-C(47)-	0.015	C(16)-C(17)-C(6)-	-0.031
C(29)-C(11)-C(59)	119.995	C(9)-C(2)-C(1)-C(6)	-142.633	C(22)-C(21)-C(20)-	-142.603	C(47)-C(48)-C(35)-	-138.228	C(16)-C(17)-C(6)-	-142.657
C(29)-C(11)-C(3)	107.986	C(3)-C(2)-C(1)-C(7)	142.626	C(22)-C(21)-C(20)-	0.021	C(47)-C(48)-C(35)-	0.001	C(26)-C(16)-C(15)-	-0.005
C(59)-C(11)-C(3)	120.001	C(3)-C(2)-C(1)-C(6)	0.024	C(21)-C(20)-C(8)-	-0.007	C(51)-C(48)-C(35)-	0.003	C(26)-C(16)-C(15)-	-142.627
C(9)-C(10)-C(59)	119.994	C(47)-C(53)-C(52)-C(51)	0.034	C(21)-C(20)-C(8)-	-138.211	C(51)-C(48)-C(35)-	138.232	C(17)-C(16)-C(15)-	142.582
C(9)-C(10)-C(57)	120.003	C(47)-C(53)-C(52)-C(23)	-142.618	C(56)-C(20)-C(8)-	138.181	C(48)-C(47)-C(46)-	138.167	C(17)-C(16)-C(15)-	-0.040
C(59)-C(10)-C(57)	107.999	C(53)-C(52)-C(51)-C(48)	-0.040	C(56)-C(20)-C(8)-	-0.023	C(48)-C(47)-C(46)-	-0.010	C(16)-C(15)-C(14)-	-138.223
C(2)-C(9)-C(8)	107.987	C(53)-C(52)-C(51)-C(50)	-142.639	C(18)-C(19)-C(7)-	-138.159	C(53)-C(47)-C(46)-	-0.017	C(16)-C(15)-C(14)-	0.004
C(2)-C(9)-C(10)	119.999	C(23)-C(52)-C(51)-C(48)	142.615	C(18)-C(19)-C(7)-	0.045	C(53)-C(47)-C(46)-	-138.194	C(5)-C(15)-C(14)-	-0.025
C(8)-C(9)-C(10)	120.007	C(23)-C(52)-C(51)-C(50)	0.016	C(22)-C(19)-C(7)-	-0.030	C(47)-C(46)-C(45)-	-142.637	C(5)-C(15)-C(14)-	138.202
C(9)-C(8)-C(20)	119.989	C(53)-C(52)-C(23)-C(22)	0.038	C(22)-C(19)-C(7)-	138.175	C(47)-C(46)-C(45)-	-0.010	C(16)-C(15)-C(5)-	0.020
C(9)-C(8)-C(7)	108.015	C(53)-C(52)-C(23)-C(24)	138.222	C(22)-C(19)-C(18)-	0.053	C(34)-C(46)-C(45)-	-0.020	C(16)-C(15)-C(5)-	142.637
C(20)-C(8)-C(7)	120.001	C(51)-C(52)-C(23)-C(22)	-138.190	C(22)-C(19)-C(18)-	-142.602	C(34)-C(46)-C(45)-	142.607	C(14)-C(15)-C(5)-	-142.607
C(1)-C(7)-C(19)	120.011	C(51)-C(52)-C(23)-C(24)	-0.006	C(10)-C(9)-C(2)-	142.644	C(47)-C(46)-C(34)-	-0.012	C(14)-C(15)-C(5)-	0.010
C(1)-C(7)-C(8)	107.992	C(52)-C(51)-C(48)-C(47)	0.030	C(10)-C(9)-C(2)-	-0.014	C(47)-C(46)-C(34)-	142.605	C(28)-C(14)-C(13)-	-0.010
C(19)-C(7)-C(8)	120.010	C(52)-C(51)-C(48)-C(35)	-142.627	C(8)-C(9)-C(2)-	0.028	C(45)-C(46)-C(34)-	-142.624	C(28)-C(14)-C(13)-	-142.636
C(1)-C(6)-C(17)	119.989	C(50)-C(51)-C(48)-C(47)	142.627	C(8)-C(9)-C(2)-	-142.629	C(45)-C(46)-C(34)-	-0.007	C(15)-C(14)-C(13)-	142.639
C(1)-C(6)-C(5)	120.009	C(50)-C(51)-C(48)-C(35)	-0.030	C(20)-C(8)-C(7)-	0.012	C(55)-C(45)-C(44)-	-142.581	C(15)-C(14)-C(13)-	0.012
C(17)-C(6)-C(5)	108.001	C(52)-C(51)-C(50)-C(49)	138.197	C(20)-C(8)-C(7)-	-142.630	C(55)-C(45)-C(44)-	0.006	C(14)-C(13)-C(12)-	138.143
C(4)-C(5)-C(15)	119.995	C(52)-C(51)-C(50)-C(25)	-0.019	C(9)-C(8)-C(7)-	142.638	C(46)-C(45)-C(44)-	0.040	C(14)-C(13)-C(12)-	0.015
C(4)-C(5)-C(6)	120.006	C(48)-C(51)-C(50)-C(49)	0.039	C(9)-C(8)-C(7)-	-0.004	C(46)-C(45)-C(44)-	142.627	C(39)-C(13)-C(12)-	-0.045
C(15)-C(5)-C(6)	107.992	C(48)-C(51)-C(50)-C(25)	-138.176	C(8)-C(7)-C(1)-	0.021	C(45)-C(44)-C(33)-	-0.044	C(39)-C(13)-C(12)-	-138.173
C(5)-C(4)-C(12)	120.020	C(51)-C(50)-C(49)-C(27)	-142.641	C(8)-C(7)-C(1)-	142.614	C(45)-C(44)-C(33)-	142.621	C(29)-C(12)-C(4)-	0.015
C(5)-C(4)-C(3)	119.986	C(51)-C(50)-C(49)-C(36)	-0.022	C(19)-C(7)-C(1)-	-142.621	C(43)-C(44)-C(33)-	-142.639	C(29)-C(12)-C(4)-	-142.598
		C(25)-C(50)-C(49)-C(27)	0.004	C(19)-C(7)-C(1)-	-0.028	C(43)-C(44)-C(33)-	0.026	C(13)-C(12)-C(4)-	142.584
		C(25)-C(50)-C(49)-C(36)	142.623	C(5)-C(6)-C(1)-	-0.029	C(45)-C(44)-C(43)-	-138.155	C(13)-C(12)-C(4)-	-0.029
		C(51)-C(50)-C(25)-C(26)	142.647	C(5)-C(6)-C(1)-	-138.192	C(45)-C(44)-C(43)-	0.002	C(59)-C(11)-C(3)-	-142.627
		C(51)-C(50)-C(25)-C(24)	0.012	C(17)-C(6)-C(1)-	138.160	C(33)-C(44)-C(43)-	-0.006	C(59)-C(11)-C(3)-	0.030
		C(49)-C(50)-C(25)-C(26)	0.004	C(17)-C(6)-C(1)-	-0.002	C(33)-C(44)-C(43)-	138.151	C(29)-C(11)-C(3)-	-0.030
		C(58)-C(41)-C(32)-C(33)	-142.617	C(17)-C(6)-C(5)-	0.007	C(57)-C(43)-C(42)-	0.010	C(29)-C(11)-C(3)-	142.627
		C(42)-C(41)-C(32)-C(31)	142.644	C(17)-C(6)-C(5)-	-142.605	C(57)-C(43)-C(42)-	-142.617	C(57)-C(10)-C(9)-	-0.006
		C(42)-C(41)-C(32)-C(33)	-0.017	C(31)-C(40)-C(39)-	138.211	C(44)-C(43)-C(42)-	142.602	C(57)-C(10)-C(9)-	-138.180
		C(58)-C(40)-C(31)-C(32)	0.042	C(31)-C(40)-C(39)-	0.039	C(44)-C(43)-C(42)-	-0.025	C(59)-C(10)-C(9)-	138.179
		C(58)-C(40)-C(31)-C(30)	142.605	C(40)-C(39)-C(13)-	-142.603	C(43)-C(42)-C(41)-	138.198	C(59)-C(10)-C(9)-	0.005
		C(39)-C(40)-C(31)-C(32)	-142.607	C(40)-C(39)-C(13)-	0.019	C(43)-C(42)-C(41)-	0.037	C(10)-C(9)-C(8)-	-142.627
		C(39)-C(40)-C(31)-C(30)	-0.044	C(38)-C(39)-C(13)-	0.005	C(59)-C(42)-C(41)-	0.002	C(10)-C(9)-C(8)-	0.005
		C(58)-C(40)-C(39)-C(13)	-0.002	C(38)-C(39)-C(13)-	142.627	C(59)-C(42)-C(41)-	-138.159	C(2)-C(9)-C(8)-C(7)	-0.015
		C(58)-C(40)-C(39)-C(38)	-138.174	C(40)-C(39)-C(38)-	-0.021	C(58)-C(41)-C(32)-	0.044	C(2)-C(9)-C(8)-	142.617
		C(1)-C(6)-C(5)-C(4)	0.012	C(40)-C(39)-C(38)-	142.603	C(13)-C(39)-C(38)-	-142.623	C(1)-C(6)-C(5)-	142.624

Table S2. The calculated geometric parameters at the four different levels of the C₆₀Cl₆ molecule. Bond lengths in Angstrom (Å) and angles in Degrees (°).

Bond Length	Method/Basis set				Bond Angles	Method/Basis set				Dihedral Angles	Method/Basis set			
	B3LYP/6-31G(d,p)	RM062X/6-31G(d,p)	HSEH1PBE/6-31G(d,p)	B3LYP/6-311++G(d,p)		B3LYP/6-31G(d,p)	RM062X/6-31G(d,p)	HSEH1PBE/6-31G(d,p)	B3LYP/6-311++G(d,p)		B3LYP/6-31G(d,p)	RM062X/6-31G(d,p)	HSEH1PBE/6-31G(d,p)	B3LYP/6-311++G(d,p)
(C1,C2)	1.54	1.54	1.53	1.54	(C2,C1,C6)	114.5	114.4	114.4	114.5	(C6,C1,C2,C3)	25.2	26.3	25.8	24.9
(C1,C6)	1.58	1.57	1.57	1.58	(C2,C1,C7)	100.9	100.6	100.9	100.9	(C6,C1,C2,C9)	-137.5	-137.9	-137.8	-137.5
(C1,C7)	1.54	1.54	1.53	1.54	(C2,C1,C162)	106.4	106.7	106.5	106.3	(C7,C1,C2,C3)	148.6	149.4	149.2	148.4
(C1,C162)	1.84	1.81	1.81	1.84	(C6,C1,C7)	114.5	114.4	114.4	114.5	(C7,C1,C2,C9)	-14.1	-14.8	-14.4	-13.9
(C2,C3)	1.49	1.49	1.49	1.49	(C6,C1,C162)	113.1	113.0	113.0	113.1	(Cl62,C1,C2,C3)	-100.5	-99.4	-99.8	-100.8
(C2,C9)	1.35	1.34	1.34	1.34	(C7,C1,C162)	106.4	106.7	106.5	106.3	(Cl62,C1,C2,C9)	96.8	96.4	96.7	96.8
(C3,C4)	1.53	1.52	1.52	1.52	(C1,C2,C3)	125.5	125.6	125.7	125.4	(C2,C1,C6,C5)	1.1	1.2	1.3	1.1
(C3,C11)	1.54	1.54	1.53	1.54	(C1,C2,C9)	108.9	109.1	108.9	108.9	(C2,C1,C6,C17)	114.7	114.2	114.5	114.8
(C3,C161)	1.85	1.81	1.82	1.85	(C3,C2,C9)	123.2	123.4	123.2	123.2	(C2,C1,C6,C163)	-122.1	-122.3	-122.1	-122.0
(C4,C5)	1.37	1.37	1.37	1.37	(C2,C3,C4)	108.5	107.7	107.9	108.6	(C7,C1,C6,C5)	-114.7	-114.2	-114.5	-114.8
(C4,C12)	1.44	1.44	1.44	1.44	(C2,C3,C11)	109.8	109.2	109.4	109.9	(C7,C1,C6,C17)	-1.1	-1.2	-1.3	-1.1
(C5,C6)	1.53	1.53	1.52	1.53	(C2,C3,C161)	113.3	113.1	113.6	113.2	(C7,C1,C6,C163)	122.1	122.3	122.1	122.0
(C5,C15)	1.44	1.44	1.43	1.44	(C4,C3,C11)	101.2	100.9	101.0	101.3	(Cl62,C1,C6,C5)	123.2	123.5	123.4	123.1
(C6,C17)	1.53	1.53	1.52	1.53	(C4,C3,C161)	111.7	112.3	112.0	111.5	(Cl62,C1,C6,C17)	-123.2	-123.5	-123.4	-123.1
(C6,C163)	1.84	1.81	1.82	1.84	(C11,C3,C161)	111.7	112.8	112.1	111.6	(Cl62,C1,C6,C163)	0.0	0.0	0.0	0.0
(C7,C8)	1.35	1.34	1.34	1.34	(C3,C4,C5)	123.8	123.8	123.8	123.8	(C2,C1,C7,C8)	14.1	14.8	14.4	13.9
(C7,C19)	1.49	1.49	1.49	1.49	(C3,C4,C12)	109.8	109.9	109.9	109.7	(C2,C1,C7,C19)	-148.6	-149.4	-149.2	-148.4
(C8,C9)	1.47	1.47	1.46	1.47	(C5,C4,C12)	119.5	119.6	119.5	119.6	(C6,C1,C7,C8)	137.5	137.9	137.8	137.5
(C8,C20)	1.50	1.50	1.49	1.49	(C4,C5,C6)	126.2	126.4	126.3	126.2	(C6,C1,C7,C19)	-25.2	-26.3	-25.8	-24.9
(C9,C10)	1.50	1.50	1.49	1.49	(C4,C5,C15)	119.7	119.7	119.7	119.7	(Cl62,C1,C7,C8)	-96.8	-96.4	-96.7	-96.8
(C10,C57)	1.52	1.52	1.52	1.52	(C6,C5,C15)	109.0	109.1	109.1	109.0	(Cl62,C1,C7,C19)	100.5	99.4	99.8	100.8
(C10,C59)	1.53	1.52	1.52	1.52	(C1,C6,C5)	112.8	112.6	112.7	112.9	(C2,C1,C162,C163)	126.5	126.6	126.4	126.5
(C10,C166)	1.85	1.81	1.82	1.85	(C1,C6,C17)	112.8	112.6	112.7	112.9	(C6,C1,C162,C163)	0.0	0.0	0.0	0.0
(C11,C29)	1.44	1.44	1.43	1.44	(C1,C6,C163)	114.4	114.4	114.4	114.4	(C7,C1,C162,C163)	-126.5	-126.6	-126.4	-126.5
(C11,C59)	1.37	1.37	1.37	1.37	(C5,C6,C17)	100.9	100.7	100.8	101.0	(C1,C2,C3,C4)	-35.1	-36.5	-36.0	-34.7
(C12,C13)	1.40	1.39	1.39	1.39	(C5,C6,C163)	107.4	107.8	107.7	107.3	(C1,C2,C3,C11)	-144.8	-145.3	-145.0	-144.6
(C12,C29)	1.45	1.45	1.44	1.45	(C17,C6,C163)	107.4	107.8	107.7	107.3	(C1,C2,C3,C161)	89.5	88.2	88.8	89.8
(C13,C14)	1.44	1.44	1.44	1.44	(C1,C7,C8)	108.9	109.1	108.9	108.9	(C9,C2,C3,C4)	125.3	125.5	125.3	125.3
(C13,C39)	1.45	1.45	1.45	1.45	(C1,C7,C19)	125.5	125.6	125.7	125.4	(C9,C2,C3,C11)	15.6	16.7	16.3	15.3
(C14,C15)	1.39	1.39	1.39	1.39	(C8,C7,C19)	123.2	123.4	123.2	123.2	(C9,C2,C3,C161)	-110.1	-109.7	-109.8	-110.2
(C14,C28)	1.45	1.45	1.44	1.45	(C7,C8,C9)	109.6	109.4	109.5	109.6	(C1,C2,C9,C8)	9.5	9.9	9.7	9.4
(C15,C16)	1.45	1.45	1.44	1.44	(C7,C8,C20)	125.3	125.5	125.4	125.3	(C1,C2,C9,C10)	170.9	171.8	171.7	170.5
(C16,C17)	1.44	1.44	1.43	1.44	(C9,C8,C20)	122.4	122.4	122.5	122.3	(C3,C2,C9,C8)	-153.7	-154.7	-154.4	-153.5
(C16,C26)	1.39	1.39	1.39	1.39	(C2,C9,C8)	109.6	109.4	109.5	109.6	(C3,C2,C9,C10)	7.7	7.2	7.6	7.6
(C17,C18)	1.37	1.37	1.37	1.37	(C2,C9,C10)	125.3	125.5	125.4	125.3	(C2,C3,C4,C5)	20.5	21.3	21.1	20.3
(C18,C19)	1.53	1.52	1.52	1.52	(C8,C9,C10)	122.4	122.4	122.5	122.3	(C2,C3,C4,C12)	-130.1	-129.6	-129.8	-130.3
(C18,C24)	1.44	1.44	1.44	1.44	(C9,C10,C57)	109.5	109.0	109.1	109.7	(C11,C3,C4,C5)	136.0	135.7	135.8	136.0
(C19,C22)	1.54	1.54	1.53	1.54	(C9,C10,C59)	108.9	108.3	108.5	109.1	(C11,C3,C4,C12)	-14.7	-15.2	-15.1	-14.6
(C19,C164)	1.85	1.81	1.82	1.85	(C9,C10,C166)	111.0	111.2	111.3	110.9	(Cl61,C3,C4,C5)	-105.1	-103.9	-104.7	-105.2
(C20,C21)	1.53	1.52	1.52	1.52	(C57,C10,C59)	102.3	102.1	102.2	102.4	(Cl61,C3,C4,C12)	104.3	105.1	104.4	104.3
(C20,C56)	1.52	1.52	1.52	1.52	(C57,C10,C166)	111.9	112.3	112.2	111.7	(C2,C3,C11,C29)	128.8	128.2	128.4	129.0
(C20,C165)	1.85	1.81	1.82	1.85	(C59,C10,C166)	112.8	113.4	113.2	112.7	(C2,C3,C11,C59)	-20.0	-20.8	-20.7	-19.7
(C21,C22)	1.37	1.37	1.37	1.37	(C3,C11,C29)	109.6	109.6	109.7	109.5	(C4,C3,C11,C29)	14.4	14.9	14.7	14.3
(C21,C54)	1.44	1.44	1.43	1.44	(C3,C11,C59)	123.4	123.5	123.5	123.4	(C4,C3,C11,C59)	-134.4	-134.1	-134.3	-134.5
(C22,C23)	1.44	1.44	1.43	1.44	(C29,C11,C59)	119.3	119.2	119.2	119.3	(Cl61,C3,C11,C29)	-104.6	-105.1	-104.6	-104.5
(C23,C24)	1.45	1.45	1.44	1.45	(C4,C12,C13)	121.2	121.2	121.3	121.2	(Cl61,C3,C11,C59)	106.6	105.9	106.3	106.8
(C23,C52)	1.40	1.39	1.40	1.40	(C4,C12,C29)	108.5	108.4	108.4	108.5	(C3,C4,C5,C6)	2.9	2.9	2.8	2.9
(C24,C25)	1.40	1.39	1.39	1.39	(C13,C12,C29)	120.1	120.1	120.2	120.2	(C3,C4,C5,C15)	-149.0	-149.7	-149.4	-149.0
(C25,C26)	1.44	1.44	1.44	1.44	(C12,C13,C14)	119.1	119.1	119.0	119.1	(C12,C4,C5,C6)	150.9	151.3	151.1	150.8
(C25,C50)	1.45	1.45	1.45	1.45	(C12,C13,C39)	120.3	120.3	120.3	120.3	(C12,C4,C5,C15)	-1.1	-1.3	-1.2	-1.1
(C26,C27)	1.45	1.45	1.44	1.45	(C14,C13,C39)	108.0	108.0	108.0	108.0	(C3,C4,C12,C13)	155.4	155.6	155.6	155.3

(C27,C28)	1.40	1.39	1.39	1.39	(C13,C14,C15)	119.0	119.0	119.0	119.0	(C3,C4,C12,C29)	10.1	10.5	10.4	10.0
(C27,C49)	1.45	1.45	1.44	1.45	(C13,C14,C28)	108.4	108.4	108.4	108.4	(C5,C4,C12,C13)	3.3	3.3	3.3	3.4
(C28,C38)	1.45	1.45	1.44	1.45	(C15,C14,C28)	120.0	119.9	119.9	119.9	(C5,C4,C12,C29)	-142.0	-141.9	-142.0	-142.0
(C29,C58)	1.40	1.39	1.40	1.40	(C5,C15,C14)	121.4	121.3	121.4	121.3	(C4,C5,C6,C1)	-14.4	-15.0	-14.7	-14.3
(C30,C31)	1.40	1.39	1.40	1.40	(C5,C15,C16)	108.5	108.4	108.4	108.5	(C4,C5,C6,C17)	-135.0	-135.1	-135.0	-135.0
(C30,C37)	1.45	1.45	1.45	1.45	(C14,C15,C16)	120.1	120.1	120.1	120.1	(C4,C5,C6,C163)	112.6	112.1	112.3	112.7
(C30,C60)	1.45	1.45	1.45	1.45	(C15,C16,C17)	108.5	108.4	108.4	108.5	(C15,C5,C6,C1)	140.0	140.0	139.9	140.1
(C31,C32)	1.45	1.45	1.44	1.45	(C15,C16,C26)	120.1	120.1	120.1	120.1	(C15,C5,C6,C17)	19.4	19.9	19.6	19.3
(C31,C40)	1.45	1.45	1.45	1.45	(C17,C16,C26)	121.4	121.3	121.4	121.3	(C15,C5,C6,C163)	-93.0	-92.9	-93.1	-92.9
(C32,C33)	1.40	1.39	1.39	1.39	(C6,C17,C16)	109.0	109.1	109.1	109.0	(C4,C5,C15,C14)	-1.9	-1.5	-1.8	-1.9
(C32,C41)	1.45	1.45	1.45	1.45	(C6,C17,C18)	126.2	126.4	126.3	126.2	(C4,C5,C15,C16)	143.5	143.7	143.5	143.5
(C33,C34)	1.45	1.45	1.45	1.45	(C16,C17,C18)	119.7	119.7	119.7	119.7	(C6,C5,C15,C14)	-158.2	-158.4	-158.3	-158.2
(C33,C44)	1.45	1.45	1.45	1.45	(C17,C18,C19)	123.8	123.8	123.8	123.8	(C6,C5,C15,C16)	-12.9	-13.2	-13.0	-12.8
(C34,C46)	1.45	1.45	1.45	1.45	(C17,C18,C24)	119.5	119.6	119.5	119.6	(C1,C6,C17,C16)	-140.0	-140.0	-139.9	-140.1
(C34,C60)	1.40	1.39	1.39	1.40	(C19,C18,C24)	109.8	109.9	109.9	109.7	(C1,C6,C17,C18)	14.4	15.0	14.7	14.3
(C35,C36)	1.45	1.45	1.45	1.45	(C7,C19,C18)	108.5	107.7	107.9	108.6	(C5,C6,C17,C16)	-19.4	-19.9	-19.6	-19.3
(C35,C48)	1.40	1.39	1.40	1.40	(C7,C19,C22)	109.8	109.2	109.4	109.9	(C5,C6,C17,C18)	135.0	135.1	135.0	135.0
(C35,C60)	1.45	1.45	1.45	1.45	(C7,C19,C164)	113.3	113.1	113.6	113.2	(C163,C6,C17,C16)	93.0	92.9	93.1	92.9
(C36,C37)	1.45	1.45	1.45	1.45	(C18,C19,C22)	101.2	100.9	101.0	101.3	(C163,C6,C17,C18)	-112.6	-112.1	-112.3	-112.7
(C36,C49)	1.40	1.39	1.39	1.40	(C18,C19,C164)	111.7	112.3	112.0	111.5	(C1,C6,C163,C162)	0.0	0.0	0.0	0.0
(C37,C38)	1.40	1.39	1.39	1.40	(C22,C19,C164)	111.7	112.8	112.1	111.6	(C5,C6,C163,C162)	-126.1	-126.0	-126.0	-126.1
(C38,C39)	1.45	1.45	1.45	1.45	(C8,C20,C21)	108.9	108.3	108.5	109.1	(C17,C6,C163,C162)	126.1	126.0	126.0	126.1
(C39,C40)	1.39	1.39	1.39	1.39	(C8,C20,C56)	109.5	109.0	109.1	109.7	(C1,C7,C8,C9)	-9.5	-9.9	-9.7	-9.4
(C40,C58)	1.45	1.45	1.45	1.45	(C8,C20,C165)	111.0	111.2	111.3	110.9	(C1,C7,C8,C20)	-170.9	-171.8	-171.7	-170.5
(C41,C42)	1.40	1.39	1.39	1.39	(C21,C20,C56)	102.3	102.1	102.2	102.4	(C19,C7,C8,C9)	153.7	154.7	154.4	153.5
(C41,C58)	1.44	1.44	1.43	1.44	(C21,C20,C165)	112.8	113.4	113.2	112.7	(C19,C7,C8,C20)	-7.7	-7.2	-7.6	-7.6
(C42,C43)	1.45	1.45	1.45	1.45	(C56,C20,C165)	111.9	112.3	112.2	111.7	(C1,C7,C19,C18)	35.1	36.5	36.0	34.7
(C42,C59)	1.44	1.44	1.43	1.44	(C20,C21,C22)	122.2	122.4	122.3	122.2	(C1,C7,C19,C22)	144.8	145.3	145.0	144.6
(C43,C44)	1.40	1.39	1.39	1.39	(C20,C21,C54)	109.2	109.3	109.3	109.2	(C1,C7,C19,C164)	-89.5	-88.2	-88.8	-89.8
(C43,C57)	1.43	1.43	1.43	1.43	(C22,C21,C54)	120.2	120.2	120.2	120.2	(C8,C7,C19,C18)	-125.3	-125.5	-125.3	-125.3
(C44,C45)	1.44	1.44	1.43	1.44	(C19,C22,C21)	123.4	123.5	123.5	123.4	(C8,C7,C19,C22)	-15.6	-16.7	-16.3	-15.3
(C45,C46)	1.45	1.45	1.45	1.45	(C19,C22,C23)	109.6	109.6	109.7	109.5	(C8,C7,C19,C164)	110.1	109.7	109.8	110.2
(C45,C55)	1.40	1.39	1.39	1.39	(C21,C22,C23)	119.3	119.2	119.2	119.3	(C7,C8,C9,C2)	0.0	0.0	0.0	0.0
(C46,C47)	1.40	1.39	1.39	1.39	(C22,C23,C24)	108.6	108.6	108.6	108.7	(C7,C8,C9,C10)	-162.1	-162.5	-162.7	-161.8
(C47,C48)	1.45	1.45	1.44	1.45	(C22,C23,C52)	121.2	121.2	121.3	121.2	(C20,C8,C9,C2)	162.1	162.5	162.7	161.8
(C47,C53)	1.45	1.45	1.45	1.45	(C24,C23,C52)	119.6	119.5	119.5	119.6	(C20,C8,C9,C10)	0.0	0.0	0.0	0.0
(C48,C51)	1.45	1.45	1.45	1.45	(C18,C24,C23)	108.5	108.4	108.4	108.5	(C7,C8,C20,C21)	25.6	26.1	26.3	25.3
(C49,C50)	1.45	1.45	1.45	1.45	(C18,C24,C25)	121.2	121.2	121.3	121.2	(C7,C8,C20,C56)	136.8	136.5	136.8	136.7
(C50,C51)	1.39	1.39	1.39	1.39	(C23,C24,C25)	120.1	120.1	120.2	120.2	(C7,C8,C20,C165)	-99.2	-99.2	-98.9	-99.4
(C51,C52)	1.45	1.45	1.45	1.45	(C24,C25,C26)	119.1	119.1	119.0	119.1	(C9,C8,C20,C21)	-133.5	-133.5	-133.6	-133.5
(C52,C53)	1.44	1.44	1.43	1.44	(C24,C25,C50)	120.3	120.3	120.3	120.3	(C9,C8,C20,C56)	-22.4	-23.1	-23.0	-22.1
(C53,C54)	1.40	1.39	1.39	1.39	(C26,C25,C50)	108.0	108.0	108.0	108.0	(C9,C8,C20,C165)	101.7	101.2	101.3	101.7
(C54,C55)	1.45	1.45	1.45	1.45	(C16,C26,C25)	119.0	119.0	119.0	119.0	(C2,C9,C10,C57)	-136.8	-136.5	-136.8	-136.7
(C55,C56)	1.43	1.43	1.43	1.43	(C16,C26,C27)	120.0	119.9	119.9	119.9	(C2,C9,C10,C59)	-25.6	-26.1	-26.3	-25.3
(C56,C57)	1.38	1.37	1.37	1.37	(C25,C26,C27)	108.4	108.4	108.4	108.4	(C2,C9,C10,C166)	99.2	99.2	98.9	99.4
(C162,C163)	3.06	3.03	3.03	3.06	(C26,C27,C28)	119.9	120.0	120.0	120.0	(C8,C9,C10,C57)	22.4	23.1	23.0	22.1
					(C26,C27,C49)	107.8	107.8	107.8	107.8	(C8,C9,C10,C59)	133.5	133.5	133.6	133.5
					(C28,C27,C49)	120.1	120.1	120.1	120.1	(C8,C9,C10,C166)	-101.7	-101.2	-101.3	-101.7
					(C14,C28,C27)	119.9	120.0	120.0	120.0	(C9,C10,C57,C43)	127.9	127.5	127.5	128.1
					(C14,C28,C38)	107.8	107.8	107.8	107.8	(C9,C10,C57,C56)	-22.8	-23.6	-23.4	-22.5
					(C27,C28,C38)	120.1	120.1	120.1	120.1	(C59,C10,C57,C43)	12.4	13.0	12.7	12.4
					(C11,C29,C12)	108.6	108.6	108.6	108.7	(C59,C10,C57,C56)	-138.2	-138.0	-138.1	-138.3
					(C11,C29,C58)	121.2	121.2	121.3	121.2	(C166,C10,C57,C43)	-108.6	-108.8	-108.8	-108.5
					(C12,C29,C58)	119.6	119.5	119.5	119.6	(C166,C10,C57,C56)	100.8	100.1	100.3	100.9
					(C31,C30,C37)	119.9	119.9	119.9	119.9	(C9,C10,C59,C11)	20.3	21.1	21.0	20.1
					(C31,C30,C60)	120.0	120.0	120.0	120.0	(C9,C10,C59,C42)	-127.7	-127.3	-127.3	-127.9
					(C37,C30,C60)	107.9	108.0	108.0	107.9	(C57,C10,C59,C11)	136.2	136.1	136.1	136.2
					(C30,C31,C32)	120.0	120.0	120.0	120.0	(C57,C10,C59,C42)	-11.9	-12.4	-12.2	-11.8
					(C30,C31,C40)	120.2	120.2	120.2	120.1	(C166,C10,C59,C11)	-103.4	-102.9	-103.1	-103.6

(C32,C31,C40)	107.8	107.8	107.8	107.8	(Cl66,C10,C59,C42	108.5	108.7	108.7	108.4
(C31,C32,C33)	119.9	119.9	119.9	119.9	(C3,C11,C29,C12)	-9.3	-9.6	-9.5	-9.2
(C31,C32,C41)	107.9	107.9	107.9	107.9	(C3,C11,C29,C58)	-153.6	-153.8	-153.9	-153.6
(C33,C32,C41)	120.1	120.2	120.1	120.1	(C59,C11,C29,C12)	141.0	140.9	141.1	141.0
(C32,C33,C34)	120.2	120.2	120.2	120.2	(C59,C11,C29,C58)	-3.3	-3.3	-3.3	-3.4
(C32,C33,C44)	119.8	119.8	119.8	119.9	(C3,C11,C59,C10)	1.6	1.4	1.6	1.6
(C34,C33,C44)	107.8	107.8	107.8	107.8	(C3,C11,C59,C42)	146.3	146.6	146.5	146.2
(C33,C34,C46)	107.9	108.0	107.9	107.9	(C29,C11,C59,C10)	-144.4	-144.8	-144.7	-144.3
(C33,C34,C60)	119.9	119.9	120.0	119.9	(C29,C11,C59,C42)	0.3	0.4	0.3	0.3
(C46,C34,C60)	119.9	119.9	120.0	119.9	(C4,C12,C13,C14)	-2.6	-2.3	-2.5	-2.6
(C36,C35,C48)	119.9	119.9	119.9	119.9	(C4,C12,C13,C39)	-139.8	-139.6	-139.6	-139.7
(C36,C35,C60)	107.9	108.0	108.0	107.9	(C29,C12,C13,C14)	138.8	138.8	138.8	138.8
(C48,C35,C60)	120.0	120.0	120.0	120.0	(C29,C12,C13,C39)	1.6	1.5	1.7	1.7
(C35,C36,C37)	108.0	108.0	108.0	108.0	(C4,C12,C29,C11)	-0.4	-0.5	-0.5	-0.5
(C35,C36,C49)	119.9	119.8	119.9	119.9	(C4,C12,C29,C58)	144.6	144.4	144.6	144.6
(C37,C36,C49)	120.1	120.0	120.1	120.1	(C13,C12,C29,C11)	-146.2	-146.1	-146.2	-146.2
(C30,C37,C36)	108.0	108.0	108.0	108.0	(C13,C12,C29,C58)	-1.1	-1.1	-1.1	-1.1
(C30,C37,C38)	119.9	119.8	119.9	119.9	(C12,C13,C14,C15)	-0.4	-0.5	-0.4	-0.4
(C36,C37,C38)	120.1	120.0	120.1	120.1	(C12,C13,C14,C28)	-142.2	-142.3	-142.1	-142.1
(C28,C38,C37)	119.8	119.8	119.8	119.8	(C39,C13,C14,C15)	141.5	141.5	141.5	141.5
(C28,C38,C39)	107.9	107.9	107.9	107.9	(C39,C13,C14,C28)	-0.3	-0.3	-0.3	-0.2
(C37,C38,C39)	120.1	120.1	120.1	120.1	(C12,C13,C39,C38)	141.4	141.6	141.4	141.4
(C13,C39,C38)	107.9	107.9	107.9	107.9	(C12,C13,C39,C40)	-0.8	-0.8	-0.9	-0.9
(C13,C39,C40)	119.7	119.7	119.7	119.7	(C14,C13,C39,C38)	0.1	0.1	0.1	0.1
(C38,C39,C40)	120.1	120.1	120.1	120.1	(C14,C13,C39,C40)	-142.2	-142.2	-142.2	-142.2
(C31,C40,C39)	119.9	119.9	119.9	119.9	(C13,C14,C15,C5)	2.6	2.4	2.6	2.6
(C31,C40,C58)	108.1	108.1	108.1	108.1	(C13,C14,C15,C16)	-138.8	-138.8	-138.8	-138.8
(C39,C40,C58)	120.0	120.0	120.0	120.0	(C28,C14,C15,C5)	140.0	139.8	139.8	139.9
(C32,C41,C42)	120.1	120.1	120.1	120.1	(C28,C14,C15,C16)	-1.5	-1.5	-1.6	-1.6
(C32,C41,C58)	108.3	108.3	108.3	108.3	(C13,C14,C28,C27)	142.8	142.8	142.8	142.8
(C42,C41,C58)	119.0	119.0	118.9	119.0	(C13,C14,C28,C38)	0.3	0.3	0.3	0.3
(C41,C42,C43)	119.6	119.6	119.6	119.6	(C15,C14,C28,C27)	1.5	1.5	1.5	1.6
(C41,C42,C59)	121.0	121.0	121.1	121.0	(C15,C14,C28,C38)	-141.0	-141.0	-141.0	-141.0
(C43,C42,C59)	108.8	108.8	108.8	108.8	(C5,C15,C16,C17)	0.0	0.0	0.0	0.0
(C42,C43,C44)	120.3	120.3	120.3	120.3	(C5,C15,C16,C26)	-145.9	-145.7	-145.9	-145.9
(C42,C43,C57)	108.7	108.6	108.6	108.7	(C14,C15,C16,C17)	145.9	145.7	145.9	145.9
(C44,C43,C57)	120.8	120.8	120.9	120.8	(C14,C15,C16,C26)	0.0	0.0	0.0	0.0
(C33,C44,C43)	120.0	120.0	120.0	119.9	(C15,C16,C17,C6)	12.9	13.2	13.0	12.8
(C33,C44,C45)	108.2	108.2	108.2	108.2	(C15,C16,C17,C18)	-143.5	-143.7	-143.5	-143.5
(C43,C44,C45)	119.3	119.3	119.3	119.3	(C26,C16,C17,C6)	158.2	158.4	158.3	158.2
(C44,C45,C46)	108.2	108.2	108.2	108.2	(C26,C16,C17,C18)	1.9	1.5	1.8	1.9
(C44,C45,C55)	119.3	119.3	119.3	119.3	(C15,C16,C26,C25)	138.8	138.8	138.8	138.8
(C46,C45,C55)	120.0	120.0	120.0	119.9	(C15,C16,C26,C27)	1.5	1.5	1.6	1.6
(C34,C46,C45)	107.8	107.8	107.8	107.8	(C17,C16,C26,C25)	-2.6	-2.4	-2.6	-2.6
(C34,C46,C47)	120.2	120.2	120.2	120.2	(C17,C16,C26,C27)	-140.0	-139.8	-139.8	-139.9
(C45,C46,C47)	119.8	119.8	119.8	119.9	(C6,C17,C18,C19)	-2.9	-2.9	-2.8	-2.9
(C46,C47,C48)	119.9	119.9	119.9	119.9	(C6,C17,C18,C24)	-150.9	-151.3	-151.1	-150.8
(C46,C47,C53)	120.1	120.2	120.1	120.1	(C16,C17,C18,C19)	149.0	149.7	149.4	149.0
(C48,C47,C53)	107.9	107.9	107.9	107.9	(C16,C17,C18,C24)	1.1	1.3	1.2	1.1
(C35,C48,C47)	120.0	120.0	120.0	120.0	(C17,C18,C19,C7)	-20.5	-21.3	-21.1	-20.3
(C35,C48,C51)	120.2	120.2	120.2	120.1	(C17,C18,C19,C22)	-136.0	-135.7	-135.8	-136.0
(C47,C48,C51)	107.8	107.8	107.8	107.8	(C17,C18,C19,C164)	105.1	103.9	104.7	105.2
(C27,C49,C36)	119.8	119.8	119.8	119.8	(C24,C18,C19,C7)	130.1	129.6	129.8	130.3
(C27,C49,C50)	107.9	107.9	107.9	107.9	(C24,C18,C19,C22)	14.7	15.2	15.1	14.6
(C36,C49,C50)	120.1	120.1	120.1	120.1	(C24,C18,C19,C164)	-104.3	-105.1	-104.4	-104.3
(C25,C50,C49)	107.9	107.9	107.9	107.9	(C17,C18,C24,C23)	142.0	141.9	142.0	142.0
(C25,C50,C51)	119.7	119.7	119.7	119.7	(C17,C18,C24,C25)	-3.3	-3.3	-3.3	-3.4
(C49,C50,C51)	120.1	120.1	120.1	120.1	(C19,C18,C24,C23)	-10.1	-10.5	-10.4	-10.0
(C48,C51,C50)	119.9	119.9	119.9	119.9	(C19,C18,C24,C25)	-155.4	-155.6	-155.6	-155.3

(C48,C51,C52)	108.1	108.1	108.1	108.1	(C7,C19,C22,C21)	20.0	20.8	20.7	19.7
(C50,C51,C52)	120.0	120.0	120.0	120.0	(C7,C19,C22,C23)	-128.8	-128.2	-128.4	-129.0
(C23,C52,C51)	120.3	120.3	120.3	120.2	(C18,C19,C22,C21)	134.4	134.1	134.3	134.5
(C23,C52,C53)	119.2	119.2	119.2	119.2	(C18,C19,C22,C23)	-14.4	-14.9	-14.7	-14.3
(C51,C52,C53)	108.0	108.0	108.0	108.0	(C164,C19,C22,C21)	-106.6	-105.9	-106.3	-106.8
(C47,C53,C52)	108.3	108.3	108.3	108.3	(C164,C19,C22,C23)	104.6	105.1	104.6	104.5
(C47,C53,C54)	120.1	120.1	120.1	120.1	(C8,C20,C21,C22)	-20.3	-21.1	-21.0	-20.1
(C52,C53,C54)	119.0	119.0	118.9	119.0	(C8,C20,C21,C54)	127.7	127.3	127.3	127.9
(C21,C54,C53)	121.0	121.0	121.1	121.0	(C56,C20,C21,C22)	-136.2	-136.1	-136.1	-136.2
(C21,C54,C55)	108.8	108.8	108.8	108.8	(C56,C20,C21,C54)	11.9	12.4	12.2	11.8
(C53,C54,C55)	119.6	119.6	119.6	119.6	(C165,C20,C21,C22)	103.4	102.9	103.1	103.6
(C45,C55,C54)	120.3	120.3	120.3	120.3	(C165,C20,C21,C54)	-108.5	-108.7	-108.7	-108.4
(C45,C55,C56)	120.8	120.8	120.9	120.8	(C8,C20,C56,C55)	-127.9	-127.5	-127.5	-128.1
(C54,C55,C56)	108.7	108.6	108.6	108.7	(C8,C20,C56,C57)	22.8	23.6	23.4	22.5
(C20,C56,C55)	109.4	109.4	109.4	109.3	(C21,C20,C56,C55)	-12.4	-13.0	-12.7	-12.4
(C20,C56,C57)	123.8	124.0	123.9	123.8	(C21,C20,C56,C57)	138.2	138.0	138.1	138.3
(C55,C56,C57)	119.9	119.9	119.9	119.9	(C165,C20,C56,C55)	108.6	108.8	108.8	108.5
(C10,C57,C43)	109.4	109.4	109.4	109.3	(C165,C20,C56,C57)	-100.8	-100.1	-100.3	-100.9
(C10,C57,C56)	123.8	124.0	123.9	123.8	(C20,C21,C22,C19)	-1.6	-1.4	-1.6	-1.6
(C43,C57,C56)	119.9	119.9	119.9	119.9	(C20,C21,C22,C23)	144.4	144.8	144.7	144.3
(C29,C58,C40)	120.3	120.3	120.3	120.2	(C54,C21,C22,C19)	-146.3	-146.6	-146.5	-146.2
(C29,C58,C41)	119.2	119.2	119.2	119.2	(C54,C21,C22,C23)	-0.3	-0.4	-0.3	-0.3
(C40,C58,C41)	108.0	108.0	108.0	108.0	(C20,C21,C54,C53)	-151.8	-152.1	-152.0	-151.8
(C10,C59,C11)	122.2	122.4	122.3	122.2	(C20,C21,C54,C55)	-7.3	-7.6	-7.5	-7.2
(C10,C59,C42)	109.2	109.3	109.3	109.2	(C22,C21,C54,C53)	-3.0	-2.8	-3.0	-3.0
(C11,C59,C42)	120.2	120.2	120.2	120.2	(C22,C21,C54,C55)	141.5	141.6	141.5	141.5
(C30,C60,C34)	119.9	119.9	119.9	119.9	(C19,C22,C23,C24)	9.3	9.6	9.5	9.2
(C30,C60,C35)	108.1	108.1	108.1	108.1	(C19,C22,C23,C52)	153.6	153.8	153.9	153.6
(C34,C60,C35)	119.9	119.9	119.9	119.9	(C21,C22,C23,C24)	-141.0	-140.9	-141.1	-141.0
(C1,C162,C163)	66.7	66.7	66.8	66.7	(C21,C22,C23,C52)	3.3	3.3	3.3	3.4
(C6,C163,C162)	65.8	65.9	65.9	65.8	(C22,C23,C24,C18)	0.4	0.5	0.5	0.5
					(C22,C23,C24,C25)	146.2	146.1	146.2	146.2
					(C52,C23,C24,C18)	-144.6	-144.4	-144.6	-144.6
					(C52,C23,C24,C25)	1.1	1.1	1.1	1.1
					(C22,C23,C52,C51)	-140.5	-140.4	-140.4	-140.4
					(C22,C23,C52,C53)	-3.1	-3.0	-3.1	-3.1
					(C24,C23,C52,C51)	0.1	0.0	0.2	0.2
					(C24,C23,C52,C53)	137.5	137.5	137.5	137.5
					(C18,C24,C25,C26)	2.6	2.3	2.5	2.6
					(C18,C24,C25,C50)	139.8	139.6	139.6	139.7
					(C23,C24,C25,C26)	-138.8	-138.8	-138.8	-138.8
					(C23,C24,C25,C50)	-1.6	-1.5	-1.7	-1.7
					(C24,C25,C26,C16)	0.4	0.5	0.4	0.4
					(C24,C25,C26,C27)	142.2	142.3	142.1	142.1
					(C50,C25,C26,C16)	-141.5	-141.5	-141.5	-141.5
					(C50,C25,C26,C27)	0.3	0.3	0.3	0.2
					(C24,C25,C50,C49)	-141.4	-141.6	-141.4	-141.4
					(C24,C25,C50,C51)	0.8	0.8	0.9	0.9
					(C26,C25,C50,C49)	-0.1	-0.1	-0.1	-0.1
					(C26,C25,C50,C51)	142.2	142.2	142.2	142.2
					(C16,C26,C27,C28)	-1.5	-1.5	-1.5	-1.6
					(C16,C26,C27,C49)	141.0	141.0	141.0	141.0
					(C25,C26,C27,C28)	-142.8	-142.8	-142.8	-142.8
					(C25,C26,C27,C49)	-0.3	-0.3	-0.3	-0.3
					(C26,C27,C28,C14)	0.0	0.0	0.0	0.0
					(C26,C27,C28,C38)	137.9	137.9	137.9	138.0
					(C49,C27,C28,C14)	-137.9	-137.9	-137.9	-138.0
					(C49,C27,C28,C38)	0.0	0.0	0.0	0.0
					(C26,C27,C49,C36)	-142.1	-142.2	-142.2	-142.2

(C26,C27,C49,C50)	0.3	0.3	0.3	0.2
(C28,C27,C49,C36)	0.3	0.2	0.3	0.3
(C28,C27,C49,C50)	142.7	142.7	142.7	142.7
(C14,C28,C38,C37)	142.1	142.2	142.2	142.2
(C14,C28,C38,C39)	-0.3	-0.3	-0.3	-0.2
(C27,C28,C38,C37)	-0.3	-0.2	-0.3	-0.3
(C27,C28,C38,C39)	-142.7	-142.7	-142.7	-142.7
(C11,C29,C58,C40)	140.5	140.4	140.4	140.4
(C11,C29,C58,C41)	3.1	3.0	3.1	3.1
(C12,C29,C58,C40)	-0.1	0.0	-0.2	-0.2
(C12,C29,C58,C41)	-137.5	-137.5	-137.5	-137.5
(C37,C30,C31,C32)	138.0	138.0	138.0	138.0
(C37,C30,C31,C40)	-0.1	-0.1	-0.1	-0.1
(C60,C30,C31,C32)	-0.1	0.0	0.0	-0.1
(C60,C30,C31,C40)	-138.1	-138.1	-138.1	-138.1
(C31,C30,C37,C36)	-142.5	-142.5	-142.5	-142.5
(C31,C30,C37,C38)	0.0	0.0	0.0	0.1
(C60,C30,C37,C36)	0.0	0.0	0.0	0.0
(C60,C30,C37,C38)	142.5	142.5	142.5	142.6
(C31,C30,C60,C34)	0.1	0.1	0.1	0.1
(C31,C30,C60,C35)	142.5	142.5	142.5	142.5
(C37,C30,C60,C34)	-142.4	-142.4	-142.4	-142.4
(C37,C30,C60,C35)	0.0	0.1	0.0	0.1
(C30,C31,C32,C33)	-0.1	-0.1	-0.1	-0.1
(C30,C31,C32,C41)	-142.7	-142.8	-142.8	-142.8
(C40,C31,C32,C33)	142.5	142.5	142.5	142.5
(C40,C31,C32,C41)	-0.1	-0.1	-0.1	-0.1
(C30,C31,C40,C39)	0.1	0.1	0.1	0.1
(C30,C31,C40,C58)	142.7	142.7	142.7	142.7
(C32,C31,C40,C39)	-142.5	-142.5	-142.5	-142.5
(C32,C31,C40,C58)	0.1	0.1	0.1	0.1
(C31,C32,C33,C34)	0.3	0.3	0.3	0.3
(C31,C32,C33,C44)	-137.6	-137.6	-137.6	-137.7
(C41,C32,C33,C34)	138.3	138.4	138.4	138.4
(C41,C32,C33,C44)	0.4	0.5	0.5	0.4
(C31,C32,C41,C42)	141.4	141.5	141.4	141.4
(C31,C32,C41,C58)	0.1	0.1	0.1	0.1
(C33,C32,C41,C42)	-1.1	-1.1	-1.2	-1.1
(C33,C32,C41,C58)	-142.5	-142.5	-142.5	-142.5
(C32,C33,C34,C46)	-142.6	-142.6	-142.6	-142.6
(C32,C33,C34,C60)	-0.2	-0.2	-0.2	-0.2
(C44,C33,C34,C46)	-0.3	-0.2	-0.3	-0.2
(C44,C33,C34,C60)	142.1	142.1	142.1	142.2
(C32,C33,C44,C43)	1.0	1.0	1.1	1.1
(C32,C33,C44,C45)	142.7	142.7	142.7	142.7
(C34,C33,C44,C43)	-141.5	-141.5	-141.4	-141.4
(C34,C33,C44,C45)	0.2	0.2	0.2	0.1
(C33,C34,C46,C45)	0.3	0.2	0.3	0.2
(C33,C34,C46,C47)	142.6	142.6	142.6	142.6
(C60,C34,C46,C45)	-142.1	-142.1	-142.1	-142.2
(C60,C34,C46,C47)	0.2	0.2	0.2	0.2
(C33,C34,C60,C30)	0.0	0.1	0.0	0.1
(C33,C34,C60,C35)	-138.0	-138.0	-138.0	-138.0
(C46,C34,C60,C30)	138.0	138.0	138.0	138.0
(C46,C34,C60,C35)	0.0	-0.1	0.0	-0.1
(C48,C35,C36,C37)	142.5	142.5	142.5	142.5
(C48,C35,C36,C49)	0.0	0.0	0.0	-0.1
(C60,C35,C36,C37)	0.0	0.0	0.0	0.0
(C60,C35,C36,C49)	-142.5	-142.5	-142.5	-142.6

(C36,C35,C48,C47)	-138.0	-138.0	-138.0	-138.0
(C36,C35,C48,C51)	0.1	0.1	0.1	0.1
(C60,C35,C48,C47)	0.1	0.0	0.0	0.1
(C60,C35,C48,C51)	138.1	138.1	138.1	138.1
(C36,C35,C60,C30)	0.0	-0.1	0.0	-0.1
(C36,C35,C60,C34)	142.4	142.4	142.4	142.4
(C48,C35,C60,C30)	-142.5	-142.5	-142.5	-142.5
(C48,C35,C60,C34)	-0.1	-0.1	-0.1	-0.1
(C35,C36,C37,C30)	0.0	0.0	0.0	0.0
(C35,C36,C37,C38)	-142.5	-142.4	-142.5	-142.5
(C49,C36,C37,C30)	142.5	142.4	142.5	142.5
(C49,C36,C37,C38)	0.0	0.0	0.0	0.0
(C35,C36,C49,C27)	137.8	137.8	137.8	137.8
(C35,C36,C49,C50)	0.0	-0.1	0.0	0.0
(C37,C36,C49,C27)	-0.3	-0.2	-0.3	-0.3
(C37,C36,C49,C50)	-138.1	-138.1	-138.2	-138.2
(C30,C37,C38,C28)	-137.8	-137.8	-137.8	-137.8
(C30,C37,C38,C39)	0.0	0.1	0.0	0.0
(C36,C37,C38,C28)	0.3	0.2	0.3	0.3
(C36,C37,C38,C39)	138.1	138.1	138.2	138.2
(C28,C38,C39,C13)	0.1	0.1	0.1	0.1
(C28,C38,C39,C40)	142.2	142.2	142.2	142.2
(C37,C38,C39,C13)	-142.2	-142.2	-142.2	-142.2
(C37,C38,C39,C40)	0.0	-0.1	-0.1	-0.1
(C13,C39,C40,C31)	137.7	137.7	137.7	137.7
(C13,C39,C40,C58)	-0.4	-0.4	-0.4	-0.4
(C38,C39,C40,C31)	0.0	0.0	0.0	0.0
(C38,C39,C40,C58)	-138.2	-138.2	-138.2	-138.2
(C31,C40,C58,C29)	-141.6	-141.7	-141.6	-141.6
(C31,C40,C58,C41)	0.0	-0.1	-0.1	-0.1
(C39,C40,C58,C29)	0.9	0.8	1.0	1.0
(C39,C40,C58,C41)	142.5	142.4	142.5	142.5
(C32,C41,C42,C43)	0.3	0.3	0.4	0.3
(C32,C41,C42,C59)	-140.5	-140.5	-140.5	-140.5
(C58,C41,C42,C43)	137.6	137.7	137.6	137.6
(C58,C41,C42,C59)	-3.2	-3.1	-3.2	-3.2
(C32,C41,C58,C29)	142.0	142.1	142.0	142.0
(C32,C41,C58,C40)	0.0	0.0	0.0	0.0
(C42,C41,C58,C29)	0.2	0.2	0.2	0.2
(C42,C41,C58,C40)	-141.9	-141.9	-141.8	-141.8
(C41,C42,C43,C44)	1.2	1.2	1.2	1.2
(C41,C42,C43,C57)	-144.2	-144.1	-144.3	-144.2
(C59,C42,C43,C44)	146.3	146.2	146.3	146.3
(C59,C42,C43,C57)	0.9	1.0	0.9	0.9
(C41,C42,C59,C10)	151.8	152.1	152.0	151.8
(C41,C42,C59,C11)	3.0	2.8	3.0	3.0
(C43,C42,C59,C10)	7.3	7.6	7.5	7.2
(C43,C42,C59,C11)	-141.5	-141.6	-141.5	-141.5
(C42,C43,C44,C33)	-1.8	-1.8	-1.9	-1.9
(C42,C43,C44,C45)	-139.3	-139.3	-139.3	-139.3
(C57,C43,C44,C33)	139.4	139.3	139.3	139.3
(C57,C43,C44,C45)	1.9	1.8	1.9	1.9
(C42,C43,C57,C10)	-8.7	-9.2	-9.0	-8.7
(C42,C43,C57,C56)	143.3	143.3	143.3	143.3
(C44,C43,C57,C10)	-153.9	-154.3	-154.1	-153.9
(C44,C43,C57,C56)	-1.9	-1.8	-1.9	-1.9
(C33,C44,C45,C46)	0.0	0.0	0.0	0.0
(C33,C44,C45,C55)	-142.0	-142.0	-141.9	-141.9
(C43,C44,C45,C46)	142.0	142.0	141.9	141.9

(C43,C44,C45,C55)	0.0	0.0	0.0	0.0
(C44,C45,C46,C34)	-0.2	-0.2	-0.2	-0.1
(C44,C45,C46,C47)	-142.7	-142.7	-142.7	-142.7
(C55,C45,C46,C34)	141.5	141.5	141.4	141.4
(C55,C45,C46,C47)	-1.0	-1.0	-1.1	-1.1
(C44,C45,C55,C54)	139.3	139.3	139.3	139.3
(C44,C45,C55,C56)	-1.9	-1.8	-1.9	-1.9
(C46,C45,C55,C54)	1.8	1.8	1.9	1.9
(C46,C45,C55,C56)	-139.4	-139.3	-139.3	-139.3
(C34,C46,C47,C48)	-0.3	-0.3	-0.3	-0.3
(C34,C46,C47,C53)	-138.3	-138.4	-138.4	-138.4
(C45,C46,C47,C48)	137.6	137.6	137.6	137.7
(C45,C46,C47,C53)	-0.4	-0.5	-0.5	-0.4
(C46,C47,C48,C35)	0.1	0.1	0.1	0.1
(C46,C47,C48,C51)	-142.5	-142.5	-142.5	-142.5
(C53,C47,C48,C35)	142.7	142.8	142.8	142.8
(C53,C47,C48,C51)	0.1	0.1	0.1	0.1
(C46,C47,C53,C52)	142.5	142.5	142.5	142.5
(C46,C47,C53,C54)	1.1	1.1	1.2	1.1
(C48,C47,C53,C52)	-0.1	-0.1	-0.1	-0.1
(C48,C47,C53,C54)	-141.4	-141.5	-141.4	-141.4
(C35,C48,C51,C50)	-0.1	-0.1	-0.1	-0.1
(C35,C48,C51,C52)	-142.7	-142.7	-142.7	-142.7
(C47,C48,C51,C50)	142.5	142.5	142.5	142.5
(C47,C48,C51,C52)	-0.1	-0.1	-0.1	-0.1
(C27,C49,C50,C25)	-0.1	-0.1	-0.1	-0.1
(C27,C49,C50,C51)	-142.2	-142.2	-142.2	-142.2
(C36,C49,C50,C25)	142.2	142.2	142.2	142.2
(C36,C49,C50,C51)	0.0	0.1	0.1	0.1
(C25,C50,C51,C48)	-137.7	-137.7	-137.7	-137.7
(C25,C50,C51,C52)	0.4	0.4	0.4	0.4
(C49,C50,C51,C48)	0.0	0.0	0.0	0.0
(C49,C50,C51,C52)	138.2	138.2	138.2	138.2
(C48,C51,C52,C23)	141.6	141.7	141.6	141.6
(C48,C51,C52,C53)	0.0	0.1	0.1	0.1
(C50,C51,C52,C23)	-0.9	-0.8	-1.0	-1.0
(C50,C51,C52,C53)	-142.5	-142.4	-142.5	-142.5
(C23,C52,C53,C47)	-142.0	-142.1	-142.0	-142.0
(C23,C52,C53,C54)	-0.2	-0.2	-0.2	-0.2
(C51,C52,C53,C47)	0.0	0.0	0.0	0.0
(C51,C52,C53,C54)	141.9	141.9	141.8	141.8
(C47,C53,C54,C21)	140.5	140.5	140.5	140.5
(C47,C53,C54,C55)	-0.3	-0.3	-0.4	-0.3
(C52,C53,C54,C21)	3.2	3.1	3.2	3.2
(C52,C53,C54,C55)	-137.6	-137.7	-137.6	-137.6
(C21,C54,C55,C45)	-146.3	-146.2	-146.3	-146.3
(C21,C54,C55,C56)	-0.9	-1.0	-0.9	-0.9
(C53,C54,C55,C45)	-1.2	-1.2	-1.2	-1.2
(C53,C54,C55,C56)	144.2	144.1	144.3	144.2
(C45,C55,C56,C20)	153.9	154.3	154.1	153.9
(C45,C55,C56,C57)	1.9	1.8	1.9	1.9

Table S3. The calculated at four different levels and experimental ^{13}C NMR isotropic chemical shifts (in tetrachloromethane with deuteriated chloroform solvent, ppm) of C_{60}Cl_6 molecule

Cal. chemical shifts (ppm) (in tetrachloromethane with deuteriated chloroform)		Exp.*	Cal.chemical shifts (ppm) (in gas)						
Nucleus	B3LYP/6-31G(d,p)		RM062X/6-31G(d,p)	HSEH1PBE/6-31G(d,p),	B3LYP/6-311++G(d,p)	B3LYP/6-31G(d,p)	RM062X/6-31G(d,p)	HSEH1PBE/6-31G(d,p),	B3LYP/6-311++G(d,p)
C2	150.1	164.4	145.7	169.8	152.8	150.0	164.1	145.5	169.7
C7	150.1	164.4	145.7	169.8	151.0	150.0	164.1	145.5	169.7
C11	146.8	160.7	142.6	166.2	148.3	146.8	160.7	142.5	166.2
C22	146.8	160.7	142.6	166.2	148.2	146.8	160.7	142.5	166.2
C36	143.2	156.9	138.6	162.6	147.6	142.8	156.5	138.5	161.9
C37	143.2	156.9	138.6	162.6	147.6	142.8	156.5	138.5	161.9
C30	143.0	156.8	138.5	161.5	147.6	142.5	156.2	137.9	161.5
C35	143.0	156.8	138.5	161.5	147.6	142.5	156.2	137.9	161.5
C5	142.9	156.6	138.4	161.4	147.2	142.3	156.1	137.8	161.1
C17	142.9	156.5	138.4	161.4	147.1	142.3	156.1	137.8	161.1
C60	142.4	156.5	137.9	161.3	146.9	142.2	156.0	137.7	160.7
C32	142.3	156.3	137.9	161.3	146.5	142.2	155.6	137.7	160.7
C47	142.3	156.3	137.7	161.0	146.4	141.8	155.6	137.1	160.3
C21	142.2	156.2	137.7	161.0	146.4	141.6	155.6	137.0	160.3
C59	142.2	156.2	137.7	161.0	146.2	141.6	155.6	137.0	160.3
C33	142.2	156.2	137.7	161.0	144.2	141.6	155.6	137.0	160.3
C46	142.2	156.2	137.7	160.9	144.0	141.6	155.6	137.0	160.2
C27	142.2	155.7	137.5	160.9	143.5	141.5	155.5	136.9	160.2
C28	142.2	155.7	137.5	160.3	143.2	141.5	155.5	136.9	159.6
C40	141.5	155.6	136.8	160.3	142.8	140.8	155.0	136.2	159.6
C51	141.5	155.6	136.8	160.2	142.3	140.8	155.0	136.2	159.5
C38	141.1	155.3	136.1	160.2	142.0	140.5	154.7	135.5	159.5
C49	141.1	155.3	136.1	159.5	141.8	140.5	154.7	135.5	158.9
C34	141.1	155.3	136.1	159.5	141.6	140.4	154.7	135.5	158.9
C31	140.9	155.3	136.1	159.1	140.8	140.3	154.7	135.5	158.4
C48	140.9	155.2	136.1	159.1	140.5	140.3	154.6	135.5	158.4
C39	140.8	155.1	136.0	158.6	140.2	140.2	154.5	135.4	157.9
C50	140.8	155.1	136.0	157.7		140.2	154.5	135.4	157.0
C44	138.1	152.2	133.7	157.4		137.4	151.5	133.1	156.7
C45	138.1	152.2	133.7	157.4		137.4	151.5	133.1	156.7
C14	138.0	152.0	133.4	157.0		137.3	151.3	132.7	156.3
C26	138.0	152.0	133.4	157.0		137.3	151.3	132.7	156.3

C8	137.8	151.7	133.4	156.4		137.1	151.1	132.7	155.8
C9	137.8	151.7	133.4	156.4		137.1	151.1	132.7	155.8
C41	137.3	151.5	132.7	156.3		136.7	151.0	132.3	155.6
C53	137.3	151.5	132.7	156.3		136.7	151.0	132.3	155.6
C43	137.1	151.2	132.6	156.0		136.7	150.6	132.0	155.5
C55	137.1	151.2	132.6	156.0		136.7	150.6	132.0	155.5
C52	136.9	150.6	132.2	155.9		136.3	150.2	131.8	155.3
C58	136.9	150.6	132.2	155.9		136.3	150.2	131.8	155.3
C12	136.6	150.6	132.1	155.3		136.2	150.2	131.5	155.0
C24	136.6	150.6	132.1	155.3		136.2	150.2	131.5	155.0
C15	136.2	150.6	131.9	154.3		135.9	150.2	131.5	153.9
C16	136.2	150.6	131.9	154.3		135.9	150.2	131.5	153.9
C13	136.0	150.2	131.3	154.1		135.4	149.6	130.7	153.6
C25	136.0	150.2	131.3	154.1		135.4	149.6	130.7	153.6
C23	135.5	149.5	131.0	153.9	135.4	135.2	148.9	130.6	153.4
C29	135.5	149.5	131.0	153.9		135.2	148.9	130.6	153.4
C42	134.4	149.3	130.0	153.3		134.5	148.7	130.0	153.4
C54	134.4	149.3	130.0	153.3		134.5	148.7	130.0	153.4
C56	134.4	148.2	129.9	152.7		134.1	148.3	129.5	152.4
C57	134.4	148.2	129.9	152.7		134.1	148.3	129.5	152.4
C4	129.2	143.2	124.7	147.0		129.4	143.3	124.9	147.1
C18	129.2	143.2	124.7	147.0		129.4	143.3	124.9	147.1
C1	76.5	70.8	67.8	89.4	69.4	75.8	70.2	67.2	88.9
C6	74.9	66.9	65.6	87.0	66.5	74.0	66.2	64.9	86.3
C3	63.1	55.2	54.8	75.5	55.4	62.2	54.6	54.1	74.7
C19	63.1	55.2	54.8	75.5		62.2	54.6	54.1	74.7
C10	62.2	54.5	53.9	74.6		61.4	53.9	53.2	73.8
C20	62.2	54.5	53.9	74.6	54.9	61.4	53.9	53.2	73.8

*Taken from ref. [49].



Comparison of Prooxidant Activities of Various Fruit Juices and Herbs via Gold Nanocluster Biosensors and Carbonyl Assay

Esin AKYÜZ^{1,*}

¹ *Istanbul University–Cerrahpasa, Faculty of Engineering, Department of Chemistry, 34320, Istanbul, Türkiye*
kondakci@iuc.edu.tr, ORCID: 0000-0002-3473-8180

Received: 04.06.2021

Accepted: 06.12.2021

Published: 31.12.2021

Abstract

Under oxidative stress conditions, which are defined as the deterioration of antioxidant and prooxidant balance in the organism in favor of prooxidants, ROS species that trigger the formation of various diseases occur. The fact that these harmful species cause oxidative damage to biological macromolecules is expressed as prooxidant activity. In this study, Cu(II)-catalyzed prooxidant activities of pomegranate, apricot, peach, and pear juices and extracts of mint, white tea, and rosehip were measured by using gold nanoclusters synthesized *via* chicken egg white proteins. Fluorometric and spectrophotometric gold nanocluster biosensors and carbonyl assay were used. The fruit juices were used directly by diluting with pure water. Herbal plant samples were extracted in an ultrasonic water bath, filtered through microfilters, and stored in the refrigerator. Total prooxidant activities of fruit juices and herbal plants were calculated in terms of mM epicatechin equivalent, and the results obtained by applying all methods were compared with each other. It has been found that the applied methods can be used to accurately determine the total prooxidant activity of many food products.

Keywords: Prooxidant activity; Gold nanocluster; Protein oxidation; Biosensor; Fruit juice; Herb.



Çeşitli Meyve Suları ve Şifalı Bitkilerin Prooksidan Aktivitelerinin Altın Nanoküme Biyosensörleri ve Karbonil Yöntemi ile Karşılaştırılması

Öz

Organizmadaki antioksidan ve prooksidan dengesinin prooksidanlar lehine bozulması olarak tanımlanan oksidatif stres koşulları altında, çeşitli hastalıkların oluşumunu tetikleyen ROS türleri meydana gelmektedir. Bu zararlı türlerin biyolojik makromoleküllerin oksidatif hasarına sebep olması prooksidan aktivite olarak ifade edilir. Bu çalışmada yumurta akı proteinleri ile sentezlenen altın nanoküme kullanılarak nar, kayısı, şeftali ve armut suları ile nane, beyaz çay ve kuşburnu özütlerinin Cu(II)–katalizli prooksidan aktiviteleri ölçüldü. Florometrik ve spektrofotometrik altın nanoküme biyosensörleri ile karbonil yöntemi kullanıldı. Meyve suları doğrudan saf suyla seyreltilerek kullanıldı. Şifalı bitki örnekleri ise ultrasonik su banyosunda ekstrakte edildikten sonra mikrofiltreden süzülüp buzdolabında saklandı. Meyve suları ve şifalı bitkilerin toplam prooksidan aktiviteleri mM epikateşin eşdeğeri cinsinden hesaplandı ve tüm yöntemlerin uygulanmasıyla elde edilen sonuçlar birbirleri ile karşılaştırıldı. Uygulanan yöntemlerin birçok gıda ürününün toplam prooksidan aktivitesinin hassas bir şekilde tayin edilebilmesi için kullanılabileceği görüldü.

Anahtar Kelimeler: Prooksidan aktivite; Altın nanoküme; Protein oksidasyonu; Biyosensör; Meyve suyu; Şifalı bitki.

1. Introduction

Antioxidants are substances that significantly prevent or delay the negative effects of free radicals and reactive species against biological macromolecules when they are present in foods or the body at low concentrations [1]. Prooxidants are toxic substances that induce oxidative damage in biomacromolecules and various pathological events or diseases. Oxidative stress may form as a result of the imbalance of antioxidants and prooxidants in the presence of transition metal ions present in the organism. These ions induce the redox cycling of natural antioxidants leading to the formation of reactive species such as reactive oxygen, reactive nitrogen, and phenoxy radicals that can damage cellular macromolecules such as DNA, lipids, and proteins in the presence of oxygen resulting appear chronic diseases and cancer [2]. It has recently been acknowledged that almost every disease is owing to some level of oxidative stress.

Proteins the major targets in the organism are attacked by free radicals resulting in some covalent changes directly or indirectly. Direct oxidative attacks involve transition metal ion catalyzed site-specific oxidative damage, where the reduced form of protein-bound metal ion

caused to occur hydroxyl radicals *via* Fenton reaction. Radical-directed oxidation stimulates releasing carbonyl groups in amino acids as markers of oxidative protein damage, especially in lysine, arginine, proline, and tyrosine [3]. Protein oxidation detection is commonly based on the measurement of stable dinitrophenyl (DNP) adduct formed from the reaction between carbonyl groups and 2,4-dinitrophenylhydrazine (DNPH) reagent [4, 5]. Nevertheless, this assay has some critical drawbacks: 1) require the use of a significant amount of decontaminated protein; 2) incapable to distinguish oxidized/nonoxidized proteins in cell or tissue; 3) non-enzymatic glycation of proteins may add carbonyl groups onto amino acid residues; 4) some plant and protein components have considerable absorbance at 370 nm such as hemoglobin and myoglobin; 5) nucleic acids in plant extracts contain carbonyl groups that can cause positive errors [3, 6, 7]. Hence, carbonyl groups must not be accepted quantitatively indicative of protein oxidation.

Pomegranate which has been used in conventional medicine for centuries in ancient cultures is a Middle Eastern fruit that extends throughout the Mediterranean region, eastward to India and China [8]. Major bioactive components of pomegranate are flavan-3-ols, flavonols, hydroxybenzoic acids, hydroxycinnamic acids, anthocyanins, gallotannins, and ellagitannins [9, 10]. Apricot fruit and products are attracted because of their special aroma, nutritive values, and taste. Major compounds of apricot are flavonoids, anthocyanins, and especially phenolic acids [11]. Peach that is native to South Asia and grown worldwide is one of the most widely consumed fruits in some European countries, especially those with the Mediterranean diet. Major compounds of peach are hydroxycinnamic acids, flavanols, anthocyanins, and flavonols [12]. Pear fruits that are widely extended throughout the temperate areas of the world, such as Australia, America, and China are in favor among consumers due to their easy digestibility and desirable taste. Major compounds of pear are phenolic acids and flavonoids [13].

Mint genres which had approximately 25–30 types distributed in Europe, Asia, Africa, North America, and Australia are widely utilized as fresh vegetable, food, herbal tea, medicine, flavoring, and spice for centuries. Major compounds of mint are phenolic acids and flavonoids. Especially the amount of rosmarinic acid in the mint extract is too high in the majority [14, 15]. White tea is originated and mainly produced in the southeastern coastal region of China and is known for its minimum processing and natural characteristics. Major compounds of white tea are phenolic acids, proanthocyanidins, glycosylated, and acetylated derivatives of flavan-3-ols and flavonols especially catechins [16, 17]. Rosehip has healing effects in inflammatory diseases, infections, flu, skincare, anti-ulcer treatments, and chronic pains. It is also used in commercial foodstuffs as a healthy ingredient in tea, marmalades, jellies, jams, soups, probiotic drinks, and

yogurts. Major compounds of rosehip are ascorbate, β -carotene, glutathione, α -tocopherol, anthocyanins, resveratrol, phenolic acids, proanthocyanidins, flavonols, and flavanols [18, 19].

Natural antioxidants are found in abundance in cereals, fruits, herbs, and beverages. Antioxidants found in these sources are vitamins, carotenes, phenolic compounds such as tocopherols, flavonoids, phenolic acids, alkaloids, chlorophyll, and nitrogenous compounds such as protein, amine, polyfunctional organic acids. However, it has been proven that these health-beneficial compounds have prooxidant activity at high metal concentrations, high pH, and in an oxygenated environment [20-22]. Since prooxidant activity is not common unlike antioxidant capacity, its mechanism is not easily understandable and its effects are not well-known, is an area that needs more research and development. In this context, the determination of free radical formation and prooxidant activity is important in terms of applying an antioxidant-rich diet and/or drug use to reduce the risk of developing cancer and chronic diseases. Various prooxidant activity determination methods have been previously developed by our research group such as solution-based, solid protein-based, and nanocluster-based biosensors [23-28]. In this study, among these methods, gold nanocluster biosensors were applied to measure the prooxidant activities of fruit juices and plant extracts.

In the last decades, noble metal nanoclusters (NCs) have received remarkable interest due to their easy synthesis, subnanometer size, photostability, and biocompatibility [29]. These nanomaterials are particularly used for metal detection and imaging in medicine [30-32]. In this study, Cu(II)-induced total prooxidant activities of fruit juices like pomegranate, apricot, peach, and pear juices, and herbal extracts as mint, white tea, and rosehip were measured for the first time with respect to the fluorometric and spectrophotometric chicken egg white protein directed gold nanocluster (CEW-AuNC-FL and CEW-AuNC-UV) biosensors and carbonyl assay. The results obtained from all methods were compared with each other.

2. Materials and Methods

2.1. Reagents and instrumentation

The chemical substances used in this study were supplied from the corresponding sources: tetrachloroauric acid (HAuCl_4), 2,4-dinitrophenylhydrazine (DNPH), and neocuproine (Nc) from Aldrich (Taufkirchen, Germany); ethylenediaminetetraacetic acid (EDTA) disodium salt from Fluka (Buchs, Switzerland); sodium dihydrogen phosphate dihydrate ($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$), sodium hydroxide (NaOH) and ethanol (EtOH) from Sigma-Aldrich (Taufkirchen, Germany); disodium hydrogen phosphate (Na_2HPO_4), hydrochloric acid (HCl) and copper(II) sulphate from Riedel-de Haën (Seelze, Germany).

An Agilent Cary Eclipse fluorescence spectrophotometer (Santa Clara, CA, United States) was used to measure fluorescence intensity values. An Agilent Cary 100 UV–Vis spectrophotometer (Santa Clara, CA, United States) was used to measure absorbance values. A Hanna Edge pH–meter (Woonsocket, RI, United States) was used to adjust pH values of solutions using a combined glass electrode. A Select vortex apparatus was used to stir solutions. A Witeg water bath was used for the synthesis of nanocluster solution. A Bandelin Sonorex ultrasonic water bath (Berlin) was used for the extraction of herbs. A Heidolph MR Hei–Standard magnetic stirrer (Schwabach, Germany) was used to dissolve protein solution. A Shimadzu ATX222 (Kyoto, Japan) analytical balance equipped with UniBloc was used to weigh all chemicals. A Millipore Simpak1 Synergy185 (France) ultra-pure water system was used to obtain pure water.

2.2. Preparation of solutions

Neocuproine (7.5 mM) solution was prepared in EtOH. Phosphate buffer ($\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$, 0.5 M, pH 7.4), 2.0 mM copper, and 0.1 M EDTA solutions were prepared in pure water. DNPH (10 mM) solution was prepared in pure water including 0.2 N HCl solution. Gold nanocluster solutions were synthesized as described in our previous work [26].

Fruit juices and herbal plants were purchased from a local market. Two grams of herb (mint, white tea, and rosehip) were weighed into a beaker, and extraction was started by adding 10 mL pure water. After 15 minutes, the supernatant was decanted, and the extraction process was repeated two times with another 10 and 5 mL pure water, respectively. The extraction process took 45 minutes in total. All extracts were passed through a GF/PET (glass fiber/polyethylene terephthalate) 1.0/0.45- μm microfilter before analysis. Fruit juices were diluted with pure water and filtered through a microfilter.

2.3. CEW–AuNC–FL biosensor

CEW–AuNC–FL was based on the quenching CEW–AuNC fluorescence intensity *via* binding Cu(I) ions to protein thiol groups ($\lambda_{\text{ex}}=360$ nm, $\lambda_{\text{em}}=640$ nm) [26]. CEW–AuNC (1 mL), 0.5 M, 0.5 mL pH 7.4 phosphate buffer, 2 mM, 0.5 mL copper(II), x mL sample solution, and (1 – x) mL of pure water were added to a test tube and incubated for 20 minutes. After adding 0.1 M, 0.5 mL EDTA solution, these mixtures were incubated for 10 minutes. Intensity values were recorded at 640 nm. Blank was prepared by adding all solutions except sample solution.

The total prooxidant activities of fruit juices and herbal extracts were calculated as mM epicatechin (ECAT) using the calibration line formed between intensity difference (ΔI) of blank

and sample, and concentration of epicatechin standard. For the CEW–AuNC–FL biosensor: $\epsilon_{\text{ECAT}} = 2860000 \text{ L mol}^{-1} \text{ cm}^{-1}$ [26].

2.4. CEW–AuNC–UV biosensor

CEW–AuNC–UV method was based on the measuring absorbance value of Cu(I)–Nc chelate formed by the reaction between protein-bound Cu(I) and neocuproine which was thought to primarily bind to the thiol groups of the protein on CEW–AuNC surface [25]. CEW–AuNC (1 mL), 0.2 M, 1 mL pH 7.4 phosphate buffer, 1 mM, 1 mL copper(II), x mL sample solution, and (2 – x) mL of pure water were added to a test tube and incubated for 30 minutes. After adding 0.5 mL, 0.1 M EDTA and 7.5 mM 1 mL Nc solutions, absorbance measurements were performed at 450 nm. Blank was prepared by adding all solutions except sample solution.

The total prooxidant activities of fruit juices and herbal extracts were calculated as mM ECAT by using the calibration line formed between absorbance and concentration of epicatechin standard. For the CEW–AuNC–UV biosensor: $\epsilon_{\text{ECAT}} = 1809 \text{ L mol}^{-1} \text{ cm}^{-1}$ [25].

2.5. Carbonyl assay

Carbonyl assay was based on the measuring absorbance values of the dinitrophenyl (DNP) hydrazone adduct formed from the reaction of the 2,4–DNPH reagent with the carbonyl groups released as a result of protein oxidation [4, 5]. CEW–AuNC (1 mL), 0.5 M, 0.5 mL pH 7.4 phosphate buffer, 2 mM, 0.5 mL copper(II), x mL sample, (1 – x) mL of pure water, and 10 mM, 1 mL DNPH solutions were added to a test tube and incubated for 30 minutes. After this period, absorbance values were recorded at 370 nm. Blank was prepared by adding all solutions except sample solution.

The total prooxidant activities of fruit juices and herbal extracts were calculated as mM ECAT by using the calibration line formed between absorbance and concentration of epicatechin standard. For the carbonyl assay: $\epsilon_{\text{ECAT}} = 6478 \text{ L mol}^{-1} \text{ cm}^{-1}$ [33].

2.6. Statistical analysis

All experiments were performed in triplicate for each sample. Statistical analyses were performed using Excel software (Microsoft Office 2016) for calculating the mean and the standard error of the mean. The significance of differences between the means of TPA values found with the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays was evaluated using the Pearson correlation coefficient.

3. Results and Discussion

In the last decades, nanomaterials have been utilized in various scientific and technological fields. Nanoclusters are particularly preferred in medical applications due to their very small size and not to accumulate in the body unlike nanoparticles [34]. In this study, NCs were used as a prooxidant biosensor for measuring copper-catalyzed prooxidant activities of some fruit juices and herbal extracts.

Antioxidant compounds are widely used as food ingredients or food additives [20]. Since these compounds are known to exhibit prooxidant behavior under certain conditions, the amount of dose used in the food industry becomes important. Thus, measuring prooxidant activity of phenolics known to be health-beneficial compounds has gained importance as it may guide diets carefully and in certain concentrations and prevent getting sick.

3.1. Fluorescence response to fruit juices and herbs

The specific fluorescence response of CEW–AuNC–FL at 640 nm was quenched via bonding cuprous ions to protein thiols on the NC surface. To examine the effect of fruit juices and herbs on the fluorescence response of CEW–AuNC–FL, the responses of pomegranate, apricot, peach, and pear juices and mint, white tea, and rosehip extracts in phosphate buffer medium (pH 7.4) without other solutions, i.e. nanocluster, metal ion were measured at 640 nm.

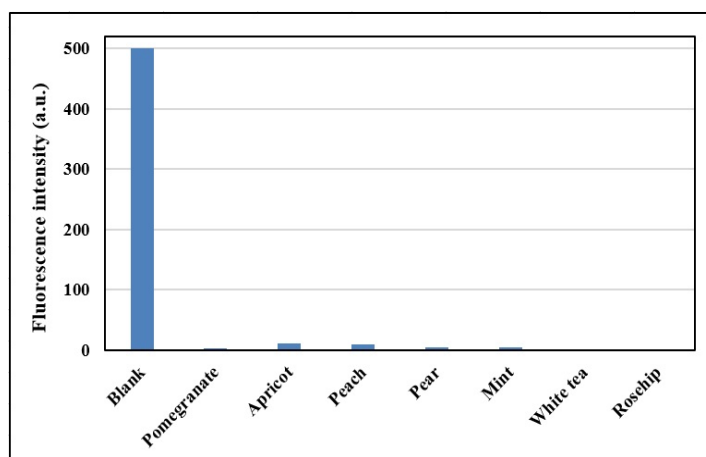


Figure 1: Fluorescence intensities of fruit juices and herbs in phosphate buffer medium (pH 7.4) at 640 nm without other solutions, i.e. nanocluster, metal ion (n=3)

The blank shown in Figure 1 was a nanocluster solution including phosphate buffer and Cu(II) solutions except for the sample. As can be seen in Figure 1, it was found that fruit juices and plant extracts did not affect the measured results as they did not have their fluorescence response.

3.2. Total prooxidant activities of pomegranate, apricot, peach, and pear juices

Total prooxidant activities (TPAs) of pomegranate, apricot, peach, and pear juices were investigated with respect to the fluorometric (CEW–AuNC–FL) and spectrophotometric (CEW–AuNC–UV) gold nanocluster biosensors and carbonyl assay. TPAs were calculated and expressed as mM ECAT equivalent by using the molar absorptivities of epicatechin given in the “Materials and Methods” section. All experiments were performed in triplicate for each sample. The results obtained can be seen in Figure 2.

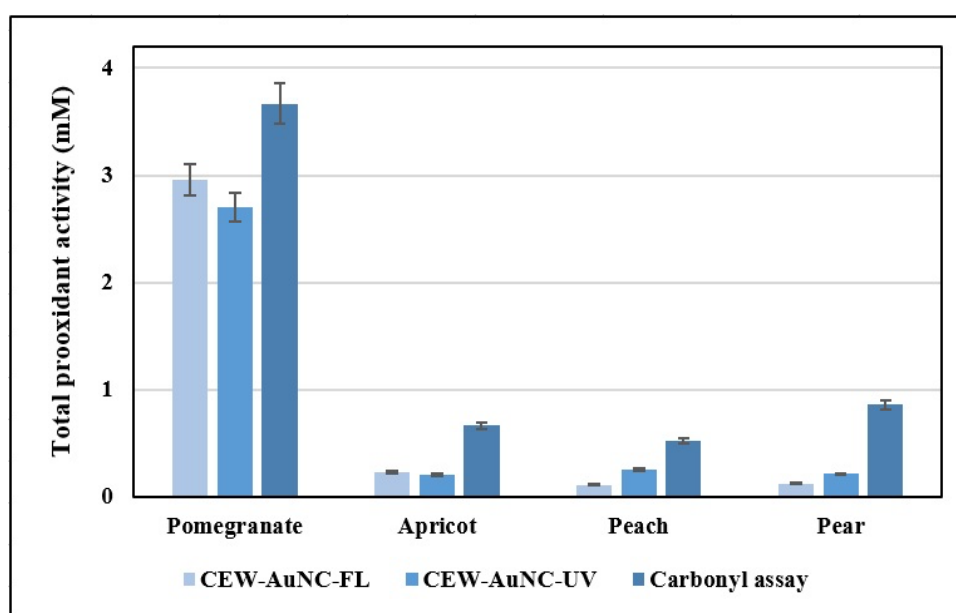


Figure 2: Total prooxidant activities of pomegranate, apricot, peach, and pear juices were calculated and expressed as mM ECAT equivalent in the bar diagram concerning the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays (n=3)

The TPA results obtained with AuNC biosensors were compatible unlike carbonyl assay results which were relatively higher. It is thought that the reason for the higher prooxidant activity results obtained via carbonyl assay is the presence of organic acids such as citric and malic acids, which are abundant in the content of fruits [9]. Since pomegranate fruit contains a high amount of both flavonoids and phenolic acids [10], prooxidant activity was found to be quite high compared to the other fruit juices, i.e. for pomegranate 2.96, 2.7, 3.67 mM; for apricot 0.23, 0.2, 0.66 mM; for peach 0.11, 0.25, 0.52 mM, and for pear 0.12, 0.21, 0.86 mM with respect to the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays respectively. The main reason why pear and apricot juice was weak prooxidant was that hydroxycinnamics were in majority compared to the flavanols in pear and apricot [35]. In the study of Slezak et al. (2017), they added high concentrations of pomegranate peels to the cell cultures and reported that pomegranate increased intracellular level

of ROS (reactive oxygen species) as behaving prooxidant activity [36]. Girard-Lalancette et al. (2009) reported that peach was slightly prooxidant at $16 \mu\text{g mL}^{-1}$ concentration on inhibiting DCFH oxidation [37]. On the other hand, no study could be found in the literature for the determination of the prooxidant activity of apricot and pear juices. In this context, prooxidant activities of related fruit juices were determined for the first time with the present study.

3.3. Total prooxidant activities of mint, white tea, and rosehip

TPAs of mint, white tea, and rosehip extracts were investigated with respect to the fluorometric (CEW–AuNC–FL) and spectrophotometric (CEW–AuNC–UV) gold nanocluster biosensors and carbonyl assay. The results were calculated as mM epicatechin equivalent using the molar absorptivities of ECAT. All experiments were performed in triplicate for each sample. The results obtained can be seen in Figure 3.

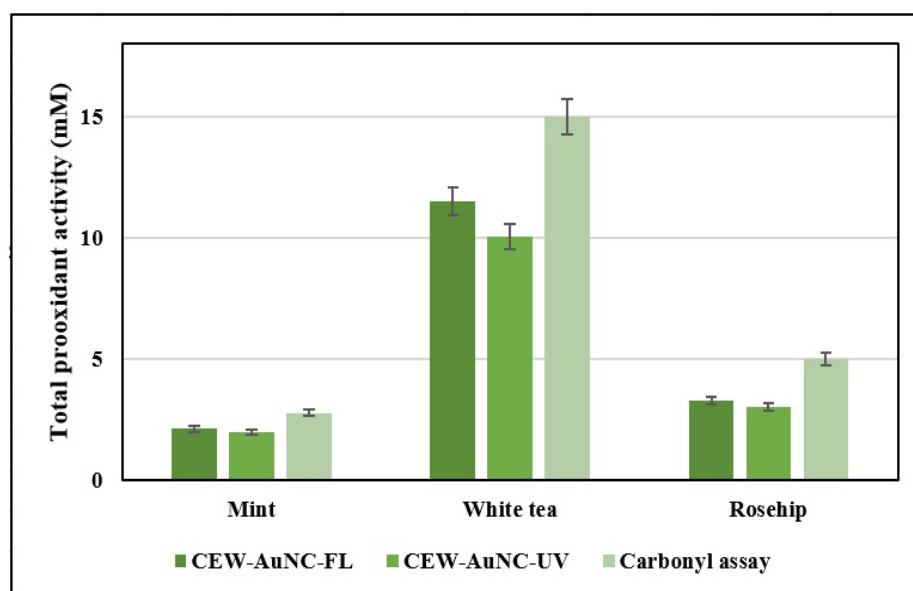


Figure 3: Total prooxidant activities of mint, white tea, and rosehip extracts were calculated and expressed as mM ECAT equivalent in the bar diagram with respect to the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays (n=3)

The TPA order of the herbal extracts was white tea > rosehip > mint using CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays. This order was the same with respect to the findings of all three methods. The presence of organic acids included the carbonyl moieties in herbs cause to increase TPAs according to the carbonyl assay [19]. Therefore, carbonyl assay results were higher than nanocluster biosensors such as for mint 2.13, 2.02, 2.79 mM; for white tea 11.5, 10.06, 15.0 mM, and for rosehip 3.3, 3.05, 5.02 mM with respect to the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays respectively. Considering that the prooxidant activities of catechin

species are generally higher than phenolic acids, it is reasonable that the prooxidant activity of white tea rich in catechins is higher than other plants. In our previous studies, in which we used solid biosensors, the prooxidant activity of mint was found to be considerably lower than that of green tea. In these studies, the total prooxidant activity of mint was calculated as 0.50 mM ECAT equivalent according to the Cu(II)-catalyzed protein-based solid biosensor method, and 3.30 mM ECAT according to the Fe(III)-catalyzed protein-based solid biosensor method [24, 27]. Although there are many studies in the literature for the measurement of the antioxidant capacity of white tea, there is no study for determining prooxidant activity. Moldovan et al. (2016) reported that the prooxidant activity of rosehip was 6 mg mL⁻¹ in tea extract [38]. Likewise, since it is known that white tea also contains catechin compounds, such as green tea, this knowledge supports the results obtained. In the study of Wang et al. (2000), it was reported that the content of flavan-3-ols found in unfermented white tea were in the following order (-)-EGCG > (-)-EGC > (-)-EC > (-)-ECG > (+)-C [39]. Also in our previous study, it was explained that especially EGCG compound had higher prooxidant activity than other catechin species [25]. Due to containing a high amount of EGCG and this compound is a stronger prooxidant than other catechins, it can be explained that white tea has much higher prooxidant activity than other herbs.

3.4. Determining significance level via Pearson correlation

The mean of total prooxidant activity values of pomegranate, apricot, peach and pear juices, and mint, white tea, and rosehip extracts measured by the CEW–AuNC–FL, CEW–AuNC–UV, and carbonyl assays were utilized to calculate significance level *via* Pearson's correlation. The Pearson correlation coefficients showed that applied methods were strongly positively correlated with each other at 95% confidence level with the values of R were 0.9998 for CEW–AuNC–FL/CEW–AuNC–UV, 0.998 for CEW–AuNC–FL/carbonyl, and 0.9979 for CEW–AuNC–UV/carbonyl (*P*-value<.00001).

4. Conclusion

In this study, total prooxidant activities of pomegranate, apricot, peach, and pear juices and mint, white tea, and rosehip herbal extracts were determined successfully with respect to the fluorometric and spectrophotometric gold nanocluster (CEW–AuNC–FL and CEW–AuNC–UV) biosensors and carbonyl assay. It was proved that the samples studied in the CEW–AuNC–FL method did not have their own fluorescence responses at the wavelength measured. Thus, it was understood that the determined prooxidant activity was only due to the fluorescence quenching of NCs with reducing the copper ion by the phenolics in the samples. The results obtained by the fluorometric and spectrophotometric CEW–AuNC biosensors were more compatible where the

carbonyl assay results were higher than the others. Since the studied samples contain high amounts of organic acid, it is thought that higher prooxidant activity was measured with positive error compared to other methods. Also, there was no significant difference at 95% confidence level *via* Pearson's correlation. The calculations showed that applied methods were strongly positively correlated with each other with the values of R were 0.9998 for CEW–AuNC–FL/CEW–AuNC–UV, 0.998 for CEW–AuNC–FL/carbonyl, and 0.9979 for CEW–AuNC–UV/carbonyl (P -value<.00001). The prooxidant activity of pomegranate juice was approximately 10 times higher than other fruit juices, and the prooxidant activity of white tea was approximately 5 times higher than other plant extracts. When the results of juices and herbal teas are compared, it is seen that the prooxidant activity of herbal teas is much higher. Considering these results, the consumption of herbal teas should be taken into consideration in terms of healthy life and nutrition.

References

- [1] Halliwell, B., Aeschbach, R., Loliger, J., Aruoma, O.I., *The characterization of antioxidants*, Food and Chemical Toxicology, 33, 601–617, 1995.
- [2] Banerjee, A., Kunwar, A., Mishra, B., Priyadarsini, K.I., *Concentration dependent antioxidant/pro-oxidant activity of curcumin: Studies from AAPH induced hemolysis of RBCs*, Chemico-Biological Interactions, 174, 134–139, 2008.
- [3] Shacter, E., *Protein Oxidative Damage*, Methods in Enzymology, 319, 428–436, 2000.
- [4] Levine, R.L., Williams, J.A., Stadtman, E.P., Shacter, E., *Carbonyl assays for determination of oxidatively modified proteins*, Methods in Enzymology, 233, 346–357, 1994.
- [5] Reznick, A.Z., Packer, L., *Oxidative damage to proteins: Spectrophotometric method for carbonyl assay*, Methods in Enzymology, 233, 357–363, 1994.
- [6] Manning, K., *Isolation of nucleic acids from plants by differential solvent precipitation*, Analytical Biochemistry, 195, 45–50, 1991.
- [7] Dalle-Donne, I., Rossi, R., Giustarini, D., Milzani, A., Colombo, R., *Protein carbonyl groups as biomarkers of oxidative stress*, Clinica Chimica Acta, 329, 23–38, 2003.
- [8] Robert, P., Gorena, T., Romero, N., Sepulveda, E., Chavez, J., Saenz, C., *Encapsulation of polyphenols and anthocyanins from pomegranate (Punica granatum) by spray drying*, International Journal of Food Science and Technology, 45, 1386–1394, 2010.
- [9] Mena, P., Calani, L., Dall'Asta, C., Galaverna, G., García-Viguera, C., Bruni, R., Crozier, A., Del Rio, D., *Rapid and comprehensive evaluation of (poly)phenolic compounds in pomegranate (Punica granatum L.) juice by UHPLC-MSⁿ*, Molecules, 17, 14821–14840, 2012.
- [10] Hmid, I., Elothmani, D., Hanine, H., Oukabli, A., Mehinagic, E., *Comparative study of phenolic compounds and their antioxidant attributes of eighteen pomegranate (Punica granatum L.) cultivars grown in Morocco*, Arabian Journal of Chemistry, 10, 2675–2684, 2017.
- [11] Dragovic-Uzelac, V., Pospišil, J., Levaj, B., Delonga, K., *The study of phenolic profiles of raw apricots and apples and their purees by HPLC for the evaluation of apricot nectars and jams authenticity*, Food Chemistry, 91(2), 373–383, 2005.

- [12] Mokrani, A., Krisa, S., Cluzet, S., Da Costa, G., Temsamani, H., Renouf, E., Mérillon, J.-M., Madani, K., Mesnil, M., Monvoisin, A., Richard, T., *Phenolic contents and bioactive potential of peach fruit extracts*, *Food Chemistry*, 202, 212-220, 2016.
- [13] Wang, Z., Barrow, C.J., Dunshea, F.R., Suleria, H.A.R., *A comparative investigation on phenolic composition, characterization and antioxidant potentials of five different Australian grown pear varieties*, *Antioxidants*, 10(2), 151, 2021.
- [14] Tang, K.S., Konczak, I., Zhao, J., *Identification and quantification of phenolics in Australian native mint (*Mentha australis* R. Br.)*, *Food Chemistry*, 192, 698-705, 2016.
- [15] Bahadori, M.B., Zengin, G., Bahadori, S., Dinparast, L., Movahhedini, N., *Phenolic composition and functional properties of wild mint (*Mentha longifolia* var. *calliantha* (Stapf) Briq.)*, *International Journal of Food Properties*, 21(1), 183-193, 2018.
- [16] Zielinski, A.A.F., Haminiuk, C.W.I., Beta, T., *Multi-response optimization of phenolic antioxidants from white tea (*Camellia sinensis* L. Kuntze) and their identification by LC–DAD–Q–TOF–MS/MS*, *LWT-Food Science and Technology*, 65, 897-907, 2016.
- [17] Tan, J., Engelhardt, U.H., Lin, Z., Kaiser, N., Maiwald, B., *Flavonoids, phenolic acids, alkaloids and theanine in different types of authentic Chinese white tea samples*, *Journal of Food Composition and Analysis*, 57, 8-15, 2017.
- [18] Stănilă, A., Diaconeasa, Z., Roman, I., Sima, N., Măniuțiu, D., Roman, A., Sima R., *Extraction and characterization of phenolic compounds from rose hip (*Rosa canina* L.) using liquid chromatography coupled with electrospray ionization-mass spectrometry*, *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 43(2), 349-354, 2015.
- [19] Demir, N., Yildiz, O., Alpaslan, M., Hayaloglu, A.A., *Evaluation of volatiles, phenolic compounds and antioxidant activities of rose hip (*Rosa* L.) fruits in Turkey*, *LWT-Food Science and Technology*, 57(1), 126-133, 2014.
- [20] Rietjens, I.M.C.M., Boersma, M.G., Haan, L., Spenkeink, B., Awad, H.M., Cnubben, N.H.P., Zanden, J.J., Woude, H., Alnk, G.M., Koeman, J.H., *The pro-oxidant chemistry of the natural antioxidants vitamin C, vitamin E, carotenoids and flavonoids*, *Environmental Toxicology and Pharmacology*, 11, 321–333, 2002.
- [21] Childs, A., Jacobs, C., Kaminski, T., Halliwell, B., Leeuwenburgh, C., *Supplementation with vitamin C and N-acetyl-cysteine increases oxidative stress in humans after an acute muscle injury induced by eccentric exercise*, *Free Radical Biology & Medicine*, 31(6), 745–753, 2001.
- [22] Hanif, S., Shamim, U., Ulah, M.F., Azmi, A.S., Bhat, S.H., Hadi, S.M., *The anthocyanidin delphinidin mobilizes endogenous copper ions from human lymphocytes leading to oxidative degradation of cellular DNA*, *Toxicology*, 249, 19–25, 2008.
- [23] Kondakçı, E., Özyürek, M., Güçlü, K., Apak, R., *Novel pro-oxidant activity assay for polyphenols, vitamins C and E using a modified CUPRAC method*, *Talanta*, 115, 583–589, 2013.
- [24] Akyüz, E., Sözgen Başkan, K., Tütem, E., Apak, R., *Novel protein-based solid-biosensor for determining pro-oxidant activity of phenolic compounds*, *Journal of Agricultural and Food Chemistry*, 65(28), 5821–5830, 2017.
- [25] Akyüz, E., Şen, F.B., Bener, M., Sözgen Başkan, K., Tütem, E., Apak, R., *Protein-protected gold nanocluster-based biosensor for determining the prooxidant activity of natural antioxidant compounds*, *ACS Omega*, 4(1), 2455–2462, 2019.
- [26] Akyüz, E., Şen, F.B., Bener, M., Sözgen Başkan, K., Apak, R., *A novel gold nanocluster-based fluorometric biosensor for measuring prooxidant activity with a large Stokes shift*, *Talanta*, 208, 1204252, 2020.

- [27] Akyüz, E., Sözgen Başkan, K., Tütem, E., Apak, R., *Novel iron(III)-induced prooxidant activity measurement using a solid protein sensor in comparison with a copper(II)-induced assay*, *Analytical Letters*, 53(9), 1489–1503, 2020.
- [28] Akyüz, E., *One-pot green synthesized protein-based silver nanocluster as prooxidant biosensor*, *Turkish Journal of Chemistry*, 45, 1422-1431, 2021.
- [29] Li, J., Zhu, J.-J., Xu, K., *Fluorescent metal nanoclusters: From synthesis to applications*, *Trends in Analytical Chemistry*, 58, 90–98, 2014.
- [30] Li, X.-J., Ling, J., Han, C.-L., Chen, L.-Q., Cao, Q.-E., Ding, Z.-T., *Chicken egg white-stabilized Au nanoclusters for selective and sensitive detection of Hg(II)*, *Analytical Sciences*, 33, 671–675, 2017.
- [31] Zhao, Q., Chen, S., Zhang, L., Huang, H., Zeng, Y., Liu, F., *Multiplex sensor for detection of different metal ions based on on-off of fluorescent gold nanoclusters*, *Analytica Chimica Acta*, 852, 236–243, 2014.
- [32] Kong, Y., Chen, J., Gao, F., Brydson, R., Johnson, B., Heath, G., Zhang, Y., Wu, L., Zhou, D., *Near-infrared fluorescent ribonuclease-a-encapsulated gold nanoclusters: Preparation, characterization, cancer targeting and imaging*, *Nanoscale*, 5, 1009–1017, 2013.
- [33] Akyüz, E., “*Determining total prooxidant activity of green tea and black tea using gold nanocluster biosensor*” in *Proceedings of the 5th International Academic Studies Conference*, pp. 194-196, 2021.
- [34] Zhang, X.-D., Wu, D., Shen, X., Liu, P.-X., Fan, F.-Y., Fan, S.-J., *In vivo renal clearance, biodistribution, toxicity of gold nanoclusters*, *Biomaterials*, 33, 4628–4638, 2012.
- [35] Sánchez, A.C.G., Gil-Izquierdo, A., Gill, M.I., *Comparative study of six pear cultivars in terms of their phenolic and vitamin C contents and antioxidant capacity*, *J Sci Food Agric*, 83, 995–1003, 2003.
- [36] Slezak, A., Moreira, H., Szyjka, A., Oszmianski, J., Gasiorowski, K., *Conditions of prooxidant activity of cistus and pomegranate polyphenols in V79 cell cultures*, *Acta Poloniae Pharmaceutica*, 74(2), 670-678, 2017.
- [37] Girard-Lalancette, K., Pichette, A., Legault, J., *Sensitive cell-based assay using DCFH oxidation for the determination of pro- and antioxidant properties of compounds and mixtures: Analysis of fruit and vegetable juices*, *Food Chemistry*, 115, 720–726, 2009.
- [38] Moldovan, B., Hosu, A., David, L., Cimpoiu, C., *Total phenolics, total anthocyanins, antioxidant and pro-oxidant activity of some red fruits teas*, *Acta Chimica Slovenica*, 63, 213–219, 2016.
- [39] Wang, H., Provan, G.J., Helliwell, K., *Tea flavonoids: their functions, utilisation and analysis*, *Trends in Food Science and Technology*, 11, 152-160, 2000.



Density Functional Theory Studies of Some Barbiturates on Lipophilicity

Sümeyya SERİN^{1,*}, Ali BAYRI²

¹*Inonu University, Scientific and Technological Research Center, 44280, Malatya, Türkiye
sumeyya.alatas@inonu.edu.tr, ORCID: 0000-0002-4637-1734*

²*Inonu University, Faculty of Science and Art, Department of Physics, 44280, Malatya, Türkiye
ali.bayri@inonu.edu.tr, ORCID: 0000-0002-8197-1604*

Received: 13.07.2021

Accepted: 10.12.2021

Published: 31.12.2021

Abstract

This paper deals with the evaluation of lipophilicity expressed by $\log P_{ow}$ parameter of ten barbiturate derivatives generally used as sedative-hypnotics based on Density Functional Theory (DFT) calculations. All geometry optimizations and frequency calculations have been carried out by using DFT/B3LYP/ 6-311++G (d,p) basis set in gas phase and also in water and n-octanol phases. Gibbs free energies of solvation for studied barbiturates were calculated to predict $\log P_{ow}$. The correlation between the calculated $\log P_{ow}$ values and available data in literature has been examined. Root mean square error (RMSE), mean square error (MSE), mean absolute deviation (MAD) and mean absolute percentage error (MAPE) statistics were utilized in measuring predictive accuracy (forecast performance) of DFT method used in this study. Accordingly, the reasonable results have been obtained in estimating the partition coefficient of the mentioned ten barbiturate derivatives by DFT/B3LYP/6-311++G (d,p) method. The lipophilicity tendency of the studied barbiturates was interpreted with the help of the calculated quantum chemical descriptors such as HOMO energy (EHOMO), LUMO energy (ELUMO), molecular volume (Vm), electrophilicity index (ω). ELUMO, Vm, and ω descriptors gave reasonable results rather than EHOMO. Also, the 3D molecular lipophilicity potential (MLP) maps that display the accumulative lipophilic contributions of each atom in studied barbiturates were visualized.



Keywords: DFT; Barbiturate; Solvation free energy; Error analysis.

Bazı Barbitüratların Lipofilikliği Üzerine Yoğunluk Fonksiyonel Teori Çalışmaları

Öz

Bu makale, Yoğunluk Fonksiyonel Teori (YFT) hesaplamalarına dayalı olarak, genellikle sedatif-hipnotik olarak kullanılan on barbitürat türevinin $\log P_{ow}$ parametresi ile ifade edilen lipofilikliğin değerlendirilmesini ele almaktadır. Tüm geometri optimizasyonları ve frekans hesaplamaları, gaz fazında ve ayrıca su ve n-oktanol fazlarında DFT/B3LYP/6-311++G (d,p) temel seti kullanılarak yapılmıştır. $\log P_{ow}$ değerlerini tahmin etmek için, çalışılan barbitüratların Gibbs serbest solvasyon enerjileri hesaplanmıştır. Hesaplanan $\log P_{ow}$ değerleri ile literatürdeki mevcut veriler arasındaki korelasyon incelenmiştir. Bu çalışmada kullanılan YFT yönteminin tahmin doğruluğunun (tahmin performansı) ölçülmesinde ortalama karekök hata (RMSE), ortalama kare hata (MSE), ortalama mutlak sapma (MAD) ve ortalama mutlak yüzde hata (MAPE) istatistiklerinden yararlanılmıştır. Buna göre, bahsedilen on barbitürat türevinin dağılım katsayısının DFT/B3LYP/6-311++G (d,p) yöntemi ile tahmin edilmesinde makul sonuçlar elde edilmiştir. İncelenen barbitüratların lipofilisite eğilimi, HOMO enerjisi (EHOMO), LUMO enerjisi (ELUMO), moleküler hacim (V_m), elektrofilik indeks (ω) gibi hesaplanan kuantum kimyasal tanımlayıcılar yardımıyla yorumlanmıştır. ELUMO, V_m ve ω tanımlayıcıları EHOMO değerine kıyasla daha makul sonuçlar vermiştir. Ayrıca, çalışılan barbitüratlarda her bir atomun birikimli lipofilik katkılarını gösteren 3 boyutlu moleküler lipofiliklik potansiyeli (MLP) haritaları görselleştirilmiştir.

Anahtar Kelimeler: YFT; Barbiturat; Solvasyon serbest enerjisi; Hata analizi.

1. Introduction

Epilepsy is one of the most common and severe neurological disorders observed as seizures that occur with some symptoms as a result of sudden, abnormal and hyper synchronized discharges of a group of neurons in the central nervous system [1]. A wide variety of treatments to prevent seizure activity options are available. The purpose of the treatment with antiepileptic drugs (AEDs) is to provide the best possible quality of life by not only elimination of seizures or reducing the number of seizures but also avoidance of drug interactions and adverse effects [2]. AEDs are chosen first of all according to clinical success, then tolerability, drug interaction, and ease of use [3-5]. By discovery of phenobarbital, a barbituric acid derivative, many new substances have begun to be used as anticonvulsants in pharmacotherapy [6-7]. Although

barbituric acid derivatives were initially considered as AEDs due to their anticonvulsive effects, they could be used in very small doses and in mild cases because of their sedative-hypnotic effects. Although barbituric acid itself is not pharmacologically active, 5,5-disubstituted derivatives have been observed to have hypnotic effect. The duration and depth of efficacy of barbiturates varies according to characteristics of substituents. Consequently, the purpose of use in treatment also changes. For instance, long-acting derivatives are used for antiepileptic and hypnotic purposes, while short-acting derivatives are used as injection anesthetics. Due to these features, barbiturates have been included in many Structure-Activity Relationships (SAR), Quantitative Structure-Activity Relationships (QSAR) and also Quantitative Structure-Pharmacokinetic Relationships (QSPkR) studies [8-15].

Computer-aided methods have become a fundamental research tool for scientists from a wide range of fields including biology, physics, chemistry and pharmacology and play a central role in combining theoretical and experimental results. QSAR studies can contribute to the designing new drugs by determining the important interactions that can have an effect on bioactivity, and predicting major parameters such as absorption, hydrophilicity, lipophilicity and toxicity. Among these parameters, the lipophilicity and hydrophilicity have been explored in detail both experimentally and theoretically [16-24]. The capability of a drug to dissolve in a lipid phase when an aqueous phase also exists frequently referred to as lipophilicity. The lipophilicity can be defined numerically by partition coefficient of a molecule in *n*-octanol-water system. The partition coefficient, *P* is dimensionless, and its logarithm ($\log P$) is often used as the measure of lipophilicity [25]. In medicinal chemistry, $\log P$ is an extremely major physicochemical parameter and has private benefit in pharmacology and toxicology [26].

Sedative-hypnotics, despite their different chemical structure, show certain common physicochemical and structural features. All of them contain polar (hydrophilic) groups as well as nonpolar (lipophilic) groups in their structures. Such compounds have dominant lipophilic character. All barbiturates mentioned in this study have two Hydrogen Bonding Donor (HD) and three Hydrogen Bonding Acceptor (HA) sites that influence the pharmacological activity. For barbiturates, maximum effect and pharmacological optimization is achieved by having *n*-octanol-water partition coefficients around $P = 100$ ($\log P = 2$) [17]. This property of the compounds is an important criterion in crossing the blood-brain barrier and in their reabsorption.

The main purpose of this study is to explore the efficacy of substitutions at 5 position of barbiturate ring on its lipid solubility based on DFT methods. Thus, DFT calculations for non-ionic forms of ten barbiturate derivatives generally used as sedative-hypnotics have been

performed. The chemical structures and IUPAC names of studied barbiturates are shown in Fig. 1. Correlations between $\log P_{ow}$ and computed descriptors have been presented.

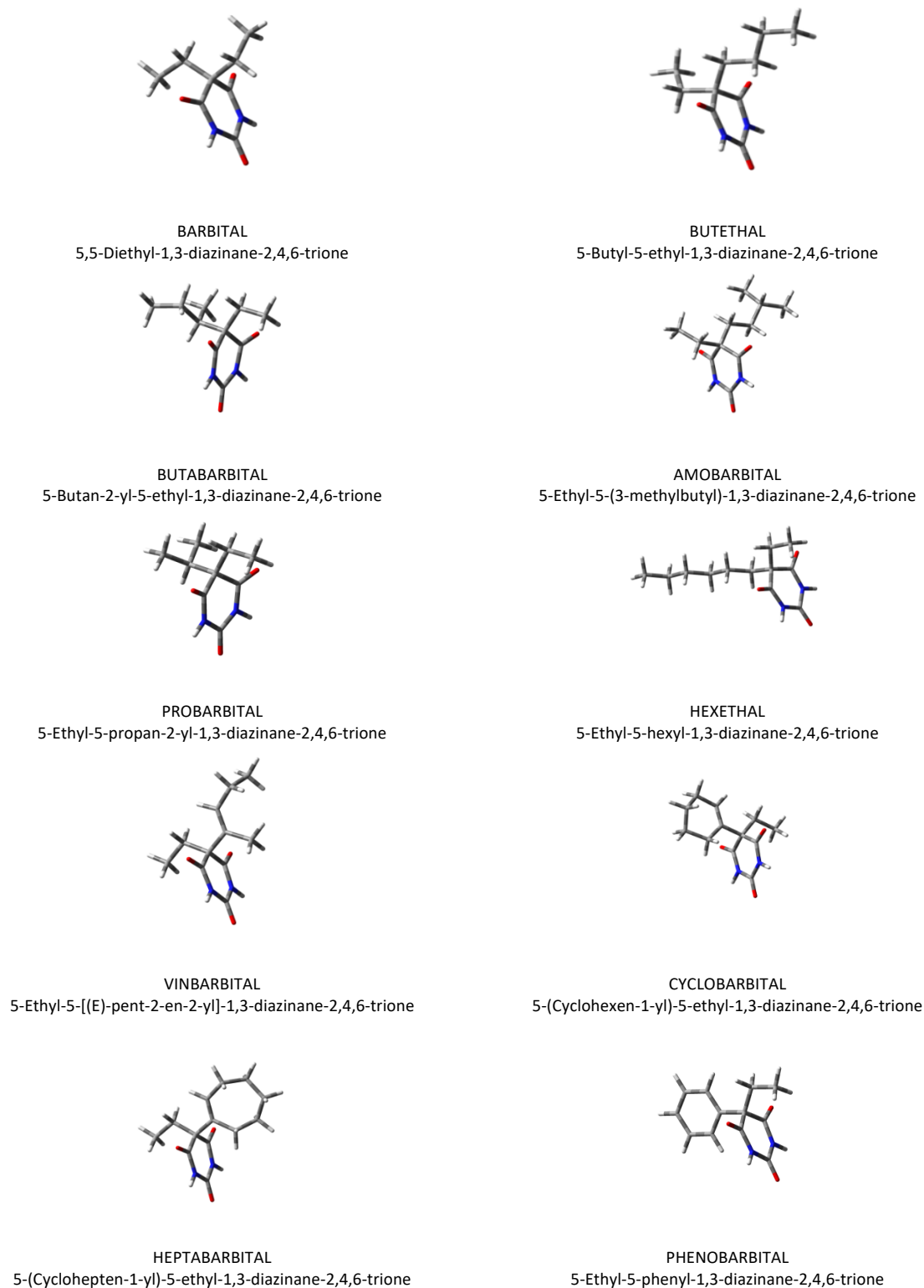


Figure 1: The chemical structures and IUPAC names of studied barbiturates

2. Materials and Methods

DFT calculations of the barbiturate derivatives were performed by using Gaussian 09 software package [27]. Also, 3D frontier molecular orbital diagrams were visualized by means of GaussView 5 molecular visualization software [28]. Geometry optimizations of all molecules were completed in vacuum and also in water ($\epsilon = 78.36$) and n-octanol ($\epsilon = 9.863$) media by using Becke, three-parameter, Lee-Yang-Parr (B3LYP) method and 6-311++G (d, p) basis set [29, 30]. The effect of the medium dielectric constant was examined by means of self-consistent reaction field (SCRF) theory calculations using the **I**ntegral **E**quation **F**ormalism **P**olarizable **C**ontinuum **M**odel (IEFPCM) for water and n-octanol phases [31-34]. The Gibbs free solvation energies of the studied molecules were calculated with the data obtained from IEFPCM calculation outputs.

As known, in accordance with Koopmans theorem [35], the ionization energy (I) and electron affinity (A) can be described via HOMO and LUMO orbital energies [36, 37] as follows (Eqn.(1) and Eqn. (2)) :

$$I = -E_{HOMO} \quad (1)$$

$$A = -E_{LUMO} \quad (2)$$

In addition, this study includes the calculation of theoretical physicochemical parameters like energy gap (ΔE), chemical softness (S), chemical hardness (η), electronegativity (χ), elec33trophilicity index (ω) and chemical potential (μ). Computational chemists make extensive use of the quantum chemical descriptors proposed by Parr and co-workers [38-42] to predict the chemical behavior of a particular molecule. The relevant formulas are given below (Eqn. (3)-(7)):

$$\mu = \frac{E_{HOMO} + E_{LUMO}}{2} \quad (3)$$

$$\chi = \frac{I+A}{2} \quad (4)$$

$$\eta = \frac{I-A}{2} \quad (5)$$

$$S = \frac{1}{2\eta} \quad (6)$$

$$\omega = \frac{\mu^2}{2\eta} \quad (7)$$

3. Results and Discussion

3.1. Partition coefficient calculations of barbiturates

Thermochemistry is as relevant to solution chemistry as it is for molecules and reactions in the gas phase. An often-desired quantity is the free energy, which can be used to compute the solvation energy of a molecule: the energy going from the gas phase to solution. To get an idea about the relative solubility of a solute in different environment, free energy of solvation can be computed for the same solute with different solvents [43]. In this study, Gibbs free energies of solvation for barbiturates were calculated in water and n-octanol phases by using IEFPCM solvent model mentioned in Materials and Methods section in order to predict logP. The theoretical logP partition coefficient can be estimated according to formula given below (Eqn. (8)) [44].

$$\text{Log}P_{ow} = \frac{\Delta G_{\text{water}} - \Delta G_{\text{n-octanol}}}{2.303RT} \quad (8)$$

In Formula, R and T are the gas constant and temperature, respectively. ΔG_{water} and $\Delta G_{\text{n-octanol}}$ are the free energy differences of compounds in solvent and in gas phase. Results of calculations for solvation free energy and theoretical $\text{Log}P_{ow}$ values are shown in Table 1.

Table 1: Solvation free energy and $\text{log}P_{ow}$ values of studied barbiturates

	ΔG_{solv} (kcal/mol)			$\text{Log}P_{ow(\text{lit})}$				
	Water ($\epsilon=78.39$)	Octanol ($\epsilon=9.863$)	$\text{Log}P_{ow}$ (calc.)	Ref. [10]	Ref. [21]	Ref. [8]	Ref. [24]	
Barbital	9.35	8.09	0.92	0.65	0.65	0.68	0.65	0.72
Butethal	9.28	7.94	0.98	1.65	1.89	1.65	1.70	1.78
Butabarbital	8.88	7.61	0.93	1.45	1.69	1.56	1.69	1.66
Amobarbital	9.29	7.97	0.97	1.95	2.07	2.07	2.09	2.19
Probarbital	9.68	8.27	1.04	0.95	-	-	0.95	1.13
Hexethal	9.95	6.73	2.36	2.65	-	-	3.08	2.84
Vinbarbital	10.02	8.44	1.15	-	1.65	-	1.95	1.95
Cyclobarbital	9.73	8.43	0.96	1.20	-	-	1.20	1.24
Heptabarbital	9.88	8.50	1.01	-	-	-	2.03	1.77
Phenobarbital	10.68	9.23	1.06	1.42	1.42	1.42	1.41	1.39

When the values given in the literature are compared with the $\log P_{ow}$ values calculated at the 6-311++G (d,p) level of theory, the graph in Fig. 2 appears. Figure 2 shows the relationship between the calculated $\log P_{ow}$ values and the values given in different sources. It is understood from these results that the DFT/B3LYP/6-311 ++ G (d, p) theoretical computational method has the power to predict the partition coefficients of the mentioned barbiturates with a ratio of over 50%. Additionally, method performance analysis has been conducted to investigate the compatibility of the DFT/B3LYP/6-311++G (d, p) method with different literatures. Root mean square errors (RMSE), mean square errors (MSE), mean absolute deviations (MAD), and mean absolute percentage errors (MAPE) were calculated [45]. Error analysis results of DFT method used for $\log P_{calc.}$ are shown in Table 2. The closer RMSE, MSE, MAD results are to zero, the better the predictive power of the method. Also, the smaller MAPE means the better forecast. Therefore, it can be concluded that the DFT/B3LYP/6-311++G (d, p) method used in this study provides reasonable results in estimating the partition coefficient of the mentioned ten barbiturate derivatives.

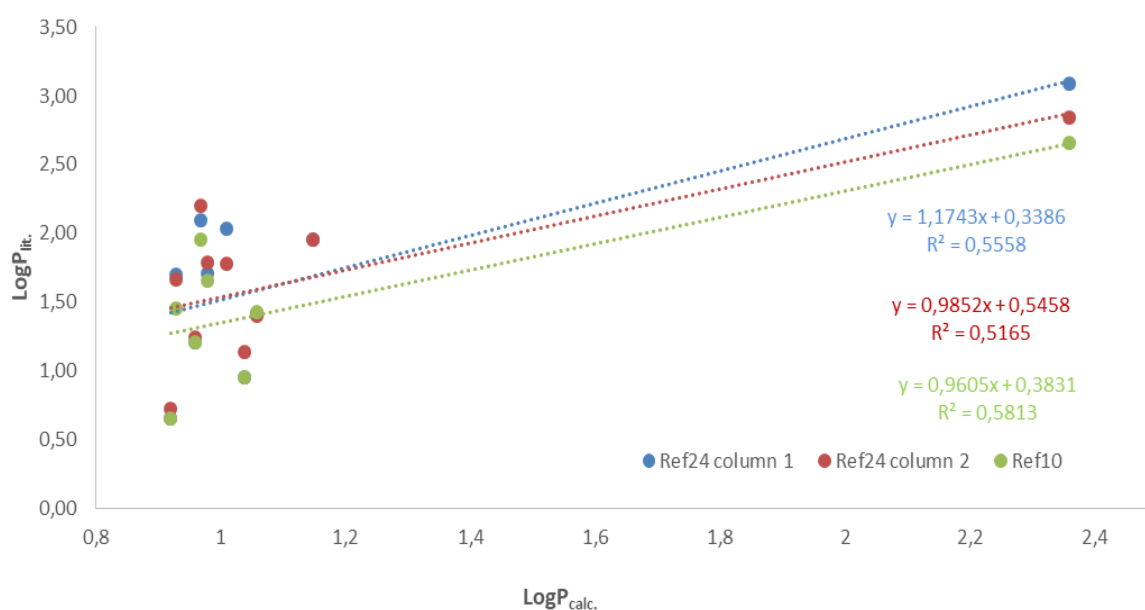


Figure 2: Linear correlations of $\log P_{calc.}$ and $\log P_{lit.}$ values

Table 2: Error analysis of DFT method used for prediction of $\log P_{calc.}$

	RMSE	MSE	MAD	MAPE
Ref. [10]	0.50	0.25	0.43	29.2
Ref. [24] Column 1	0.69	0.48	0.61	35.1
Ref. [24] Column 2	0.66	0.43	0.57	32.8

3.2. HOMO-LUMO analysis of barbiturates

As it is already well known, HOMO and LUMO energy levels are very important in molecular reactivity. EHOMO and ELUMO can describe the hydrogen bond basicity and hydrogen bond acidity of a molecule respectively [46]. Therefore, in this study, frontier molecular orbital energy levels and energy gaps (ΔE) were calculated for barbiturates to investigate the chemical reactivity behaviors. Some physicochemical properties such as Chemical Hardness (η), Softness (S), Electronegativity (χ), Chemical Potential (μ) and Electrophilicity index (ω) were also calculated with the same level of theory. We also carried out calculations in two different solvent media by using IEFPCM solvent model which is the most widely used one to evaluate the solvent effect [34]. The calculated results are listed in Table 3 and Table 4. Figure 3 represents the ΔE values of barbiturates according to changing dielectric media. No sharp increases or decreases in energy gap values were observed on going from gas phase to solvent phase. However, a decrease was observed in ΔE values of barbiturates containing cyclic and unsaturated substituents compared to others. The high energy gap value indicates good stability and low reactivity. The 3D diagrams for HOMO-LUMO energy levels of studied barbiturates are given in Fig. 4. According to Figure 4, while the LUMO locations do not change, it is observed that the HOMO locations change in barbiturates containing cyclic and unsaturated substituents.

Table 3: EHOMO, ELUMO and Energy gap ($\Delta E = ELUMO - EHOMO$) results for studied barbiturates at different dielectric media (in eV)

	Gas			Water ($\epsilon=78.39$)			Octanol ($\epsilon=9.863$)		
	EHOMO	ELUMO	ΔE	EHOMO	ELUMO	ΔE	EHOMO	ELUMO	ΔE
Barbital	-7.735	-1.729	6.005	-7.740	-1.698	6.041	-7.739	-1.706	6.033
Butethal	-7.719	-1.707	6.011	-7.735	-1.690	6.045	-7.734	-1.696	6.037
Butabarbital	-7.683	-1.775	5.907	-7.697	-1.763	5.933	-7.696	-1.769	5.926
Amobarbital	-7.716	-1.703	6.013	-7.734	-1.687	6.046	-7.731	-1.692	6.039
Probarbital	-7.704	-1.707	5.997	-7.728	-1.699	6.029	-7.727	-1.704	6.022
Hexethal	-7.701	-1.792	5.912	-7.713	-1.764	5.948	-7.711	-1.771	5.940
Vinbarbital	-7.072	-1.681	5.391	-7.026	-1.757	5.268	-7.014	-1.746	5.267
Cyclobarbital	-7.008	-1.667	5.340	-6.962	-1.754	5.207	-6.948	-1.742	5.206
Heptabarbital	-6.932	-1.669	5.263	-6.892	-1.750	5.141	-6.880	-1.738	5.141
Phenobarbital	-7.280	-1.759	5.521	-7.271	-1.820	5.451	-7.255	-1.810	5.445

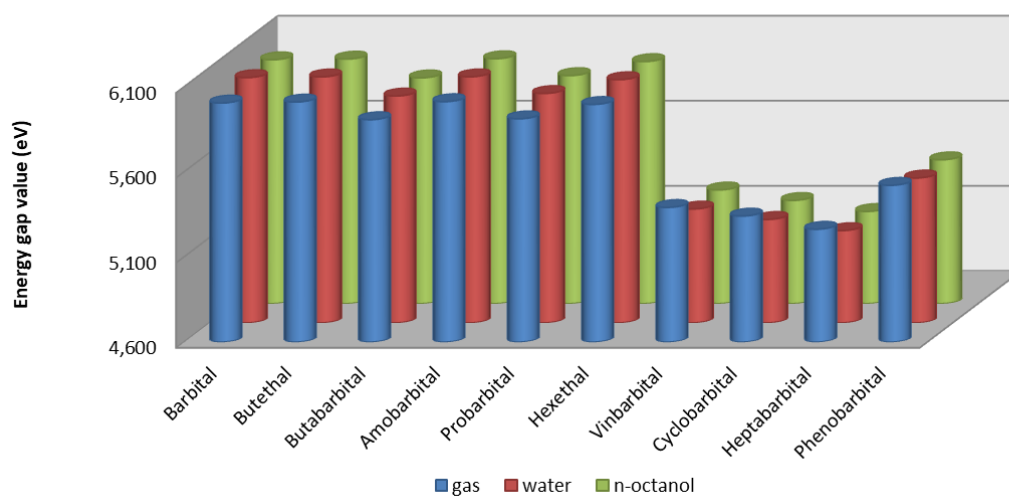


Figure 3: ΔE values of barbiturates according to changing dielectric media

Table 4: Calculated quantum chemical descriptors of studied barbiturates

	Compound	Chemical Hardness (η) (eV)	Softness (S) (eV^{-1})	Chemical Potential (μ) (eV)	Electronegativity (χ) (eV)	Electrophilicity index (ω) (eV)
gas	Barbital	3.00265	0.16652	-4.73247	4.73247	3.72942
	Butethal	3.00564	0.16635	-4.71342	4.71342	3.69577
	Butabarbitital	2.95380	0.16927	-4.72934	4.72934	3.78608
	Amobarbitital	3.00673	0.16629	-4.71016	4.71016	3.68932
	Probarbitital	2.99870	0.16674	-4.70622	4.70622	3.69302
	Hexethal	2.95639	0.16913	-4.74853	4.74853	3.81352
	Vinbarbitital	2.69570	0.18548	-4.37709	4.37709	3.55361
	Cyclobarbitital	2.67025	0.18725	-4.33804	4.33804	3.52375
	Heptabarbitital	2.63162	0.19000	-4.30077	4.30077	3.51430
	Phenobarbitital	2.76074	0.18111	-4.51995	4.51995	3.70008
water	Barbital	3.02088	0.16551	-4.71968	4.71968	3.68690
	Butethal	3.02292	0.16540	-4.71301	4.71301	3.67401
	Butabarbitital	2.96672	0.16854	-4.73056	4.73056	3.77154
	Amobarbitital	3.02319	0.16539	-4.71111	4.71111	3.67072
	Probarbitital	3.01475	0.16585	-4.71410	4.71410	3.68567
	Hexethal	2.97421	0.16811	-4.73914	4.73914	3.77570
	Vinbarbitital	2.63447	0.18979	-4.39233	4.39233	3.66156
	Cyclobarbitital	2.60373	0.19203	-4.35832	4.35832	3.64764
	Heptabarbitital	2.57094	0.19448	-4.32172	4.32172	3.63238
	Phenobarbitital	2.72563	0.18344	-4.54580	4.54580	3.79074
n-octanol	Barbital	3.01666	0.16575	-4.72281	4.72281	3.69696
	Butethal	3.01870	0.16563	-4.71533	4.71533	3.68277
	Butabarbitital	2.96346	0.16872	-4.73274	4.73274	3.77917
	Amobarbitital	3.01979	0.16557	-4.71179	4.71179	3.67591
	Probarbitital	3.01135	0.16604	-4.71587	4.71587	3.69260
	Hexethal	2.97013	0.16834	-4.74186	4.74186	3.78523
	Vinbarbitital	2.63379	0.18984	-4.38049	4.38049	3.64279
	Cyclobarbitital	2.60304	0.19208	-4.34539	4.34539	3.62699
	Heptabarbitital	2.57066	0.19450	-4.30947	4.30947	3.61221
	Phenobarbitital	2.72264	0.18365	-4.53301	4.53301	3.77357

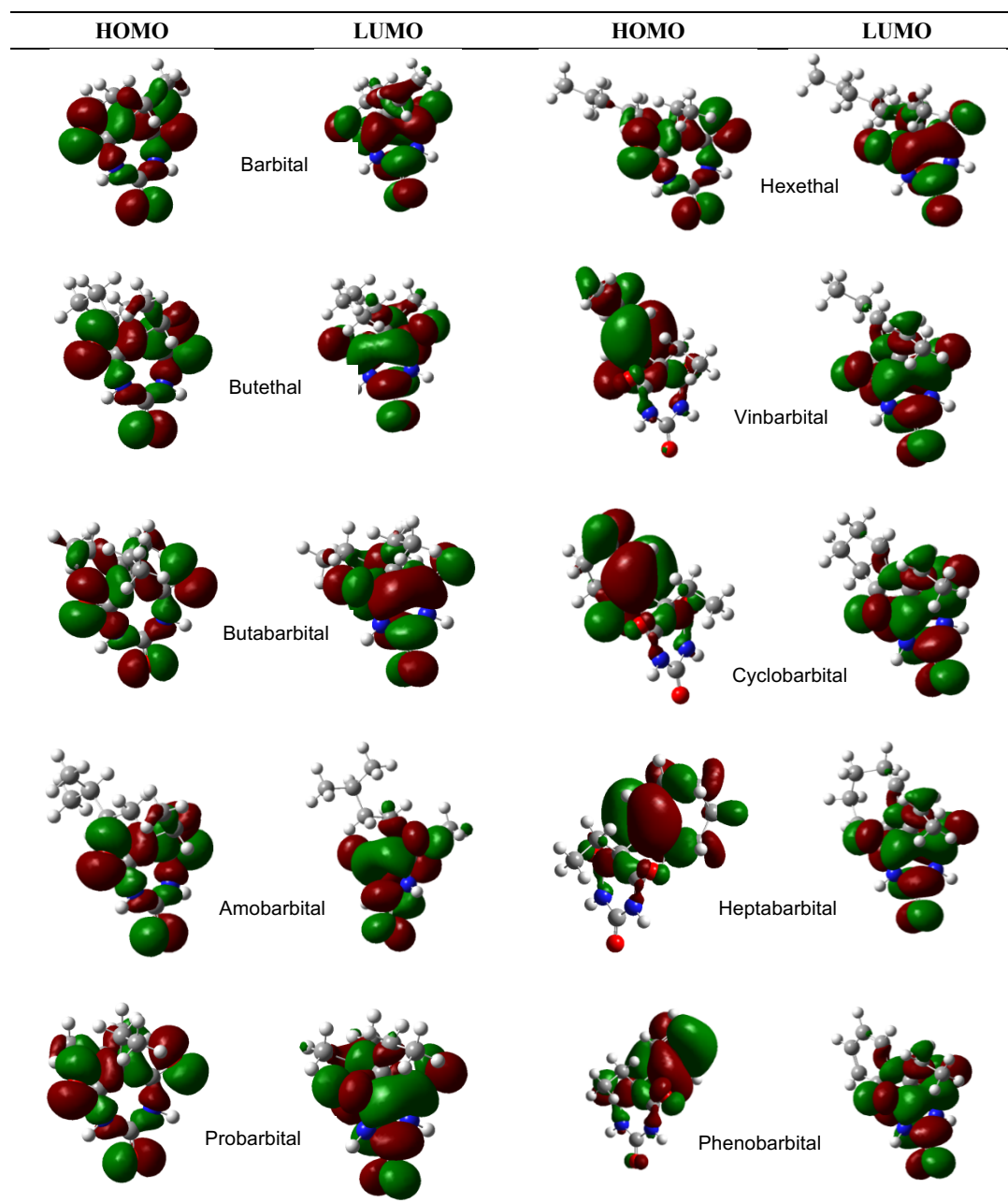


Figure 4: Frontier molecular orbital diagrams of studied barbiturates

It is pretty much practical to use computational methods including solvent models to determine the hydrophilic and lipophilic properties of drug candidates and to get an insight about correlations between computed descriptors and partition coefficients. Some studies reported the correlations between $\log P_{ow}$ and theoretically calculated descriptors such as EHOMO, ELUMO, molecular polarizability (α), molecular volume (V_m), Electrophilicity index (ω) in order to analyze the relationship between physicochemical properties and molecular structure [47-49]. It is mentioned in these publications that the value of $\log P$ is inversely proportional to EHOMO and

ELUMO because of their roles in formation of hydrogen bond and directly proportional to molecular volume and electrophilicity index. In the light of this information, if we evaluate the relationship between $\log P_{ow}$ values and computed descriptors, the ranking should be as follows:

	Straight-chain barbiturates	Branched-chain barbiturates	Cyclic and unsaturated side chain barbiturates
Considering EHOMO	Barb>Bute>Hex	Amo>Pro>Buta	Phen>Vin>Cyclo>Hepta
Considering ELUMO	Hex>Barb>Bute	Buta>Pro>Amo	Phen>Vin>Cyclo>Hepta
Considering V_m	Hex>Bute>Barb	Amo>Buta>Pro	Hepta>Cyclo>Vin>Phen
Considering ω	Hex>Barb>Bute	Buta>Pro>Amo	Phen>Vin>Cyclo>Hepta

Barb: Barbital, Bute: Butethal, Hex: Hexethal, Amo: Amobarbital, Pro: Probarbital, Buta: Butabarbital, Phen: Phenobarbital, Vin: Vinbarbital, Cyclo: Cyclobarbital, Hepta: Heptabarbital

When looking at this ranking, it can be seen that ELUMO, V_m , and ω descriptors give reasonable results rather than EHOMO. The robustness, reliability and validity of the property predictive power of a QSPR model vary depending on the compatibility of computed molecular descriptors.

Many sources of molecular descriptors can be found through quantum chemical calculations. Meanwhile, other lipophilicity descriptors Molecular Lipophilicity Potential (MLP) and Polar Surface Area (PSA) were calculated by using molinspiration cheminformatics software. The maps of MLP were visualized in Molinspiration Galaxy 3D Structure Generator v2018.01 beta [50, 51].

The map of Molecular Lipophilicity Potential defines qualitatively the 3D distribution of lipophilicity of a molecule on molecular surface. Polar surface area is characterized as the surface area (\AA^2) of O- and N-centered polar fragments and hydrogens bonded to them in a molecule and it is mightily related to hydrogen bonding capacity and polarity. It has been stated that for a molecule to penetrate the brain, the polar surface area must be around 90\AA^2 at most [52]. The barbiturates mentioned in this study are structurally similar, they all contain the same number of Hydrogen Binding Donor and Hydrogen Binding Acceptor sites, because of this the calculated PSA values (75.27\AA^2) are the same for studied barbiturates. Therefore, the lipophilic character of the substituents attached at the position 5 has a major influence on the $\log P_{ow}$ value.

Figure 5 represents maps of MLP and calculated molecular volume values for ten barbiturate derivatives. When Fig. 5 is examined, it is clearly seen which surfaces of the

barbiturate derivatives are lipophilic and which are hydrophilic. The most lipophilic surfaces are coded by violet and blue, the intermediate lipophilic surfaces are coded by green, and finally the hydrophilic surfaces are coded by orange and red. This situation demonstrated the effect of the structural properties of the substituents on lipid solubility of barbiturates.

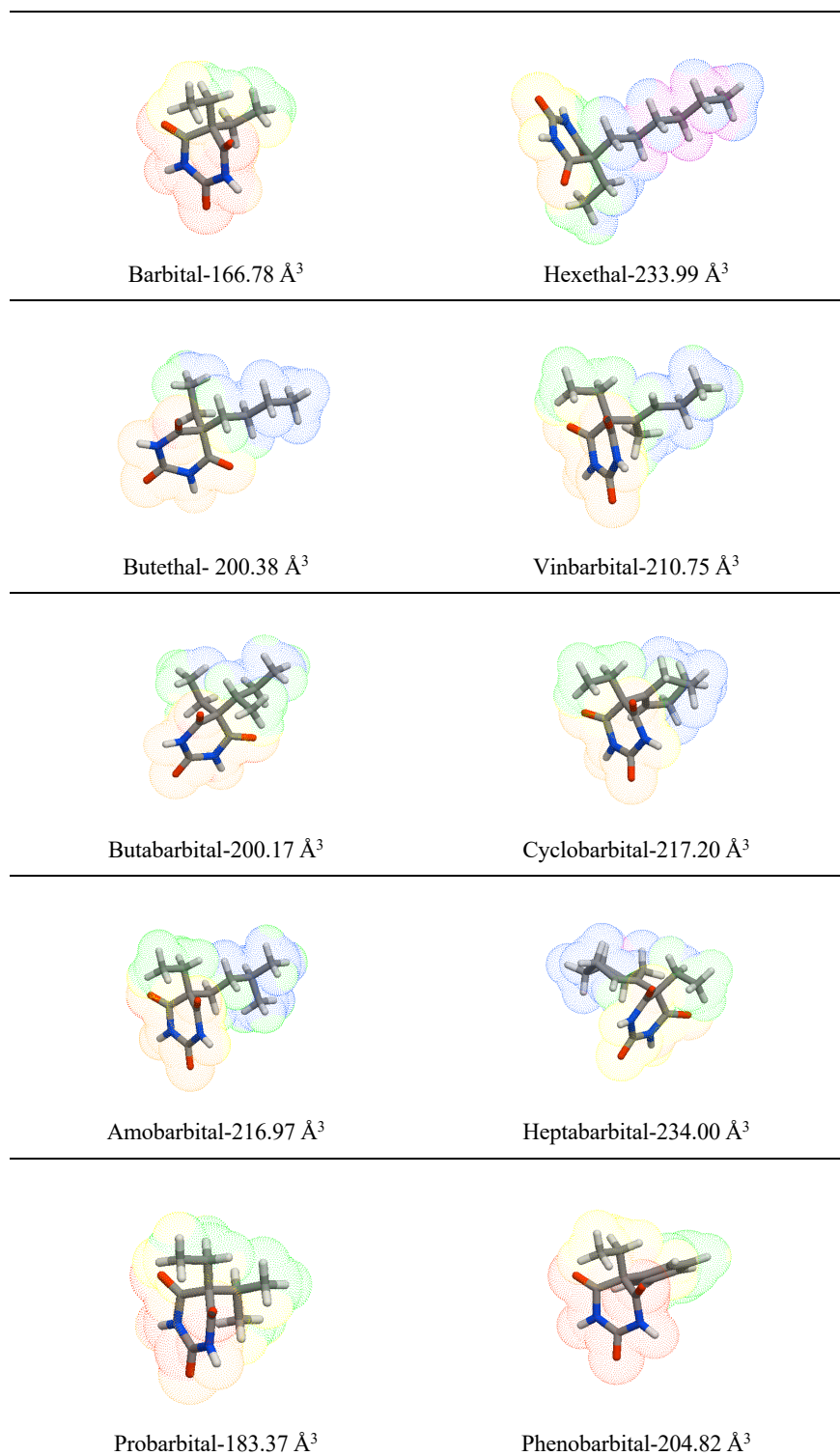


Figure 5: Maps of MLP and calculated molecular volume values for ten barbiturate derivatives

4. Conclusion

Partition coefficient estimations of ten barbituric acid derivatives containing straight, branched, cyclic or unsaturated side chain at C-5 position have been performed by using DFT/B3LYP method and 6-311++G(d,p) basis set. The correlation between the calculated values and the values given in the literature was examined. According to the scatter diagram obtained as a result of the linear correlation study, it was determined that the theoretical method used could predict the partition coefficients of barbiturate derivatives with a ratio of over 50%. Method performance analysis has been conducted by calculating RMSE, MSE, MAD, and MAPE to investigate the compatibility of this method in estimation of $LogP_{ow}$. The error analysis results show that DFT/B3LYP/6-311++G (d, p) method has a predictive capability for mentioned ten barbiturate derivatives. Additionally, HOMO-LUMO evaluation and calculations of quantum chemical descriptors such as Chemical Hardness (η), Softness (s), Electronegativity (χ), Chemical Potential (μ) and Electrophilicity index (ω) have been carried out with the same level of theory for not only gas phase but also water and n-octanol phases. It has been observed that there is no notable change in the energy gap values, which help to characterize chemical reactivity and kinetic stability, when passing from the gas phase to the solvent phase. According to 3D HOMO-LUMO diagrams of studied barbiturates, while the LUMO locations do not change, the HOMO locations change in barbiturates containing cyclic and unsaturated substituents.

Besides, the relationships between $\log P_{ow}$ and computed descriptors EHOMO, ELUMO, V_m and ω have been presented. ELUMO, V_m , and ω descriptors gave reasonable results rather than EHOMO. Last, it is clear that the maps of Molecular Lipophilicity Potential (MLP) visualized in Molinspiration Galaxy 3D Structure Generator v2018.01 beta are consistent with the results. According to the MLP map, it was seen that the most lipophilic regions coded with blue and violet colors were more intense in the Hexethal derivative with the highest partition coefficient. Hence, the usage of computational methods may offer an influential strategy in order to derive novel descriptors that may assist to determine lipophilicity in drug design studies.

References

- [1] Ngugi, A.K., Bottomley, C., Kleinschmidt, I., Sander, J.W., Newton, C.R., *Estimation of the burden of active and life-time epilepsy: a meta-analytic approach*, *Epilepsia*, 51(5), 883–890, 2010.
- [2] Adams, R.D., Victor, M., Ropper, A.H., *Epilepsy and other seizure disorder*. Adams RD (Ed). Principles of Neurology. Seventh ed., New York; McGraw-Hill 331-365, 2001.
- [3] Brodie, M.J., Kwan, P., *The star systems: overview and use in determining antiepileptic drug choice*, *CNS Drugs*, 15(1), 1–12, 2001.

- [4] Sander, J.W., *The use of antiepileptic drugs-principles and practice*, *Epilepsia*, 45, 28-34, 2004.
- [5] Yamatogi, Y., *Principles of antiepileptic drug treatment of epilepsy*. *Psychiatry and Clinical Neurosciences*, 58(3), 3-6, 2004.
- [6] Brenner, G.M., Stevens, C.W., *Pharmacology* (4th ed.). Philadelphia, PA: Elsevier/Saunders, 204p, 2013.
- [7] Engel, J., *Epilepsy: A comprehensive textbook* (2nd ed.). Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins., 1431p, 2008.
- [8] Cuenca-Benito, M., Sagrado, S., Villanueva-Camanas, R.M., Medina- Hernandez, M. J., *Quantitative retention-structure and retention-activity relationships of barbiturates by micellar liquid chromatography*, *Journal of Chromatography A*, 814, 121-132, 1998.
- [9] Henry, D., Block, J.H., Andersen, J.L., and Carlson, G.R., *Use of high-pressure liquid chromatography for quantitative structure-activity relationship studies of sulfonamides and barbiturates*. *Journal of Medicinal Chemistry*, 19(5), 619-626, 1976.
- [10] Hansch, C., Andersen, S.M., *Structure-activity relationship in barbiturates and its similarity to that in other narcotics*, *Journal of Medicinal Chemistry*, 10(5), 745-753, 1967.
- [11] Mayer, J.M., van de Waterbeemd, H., *Development of quantitative structure-pharmacokinetic relationships*. *Environmental Health Perspectives*, 61, 295-306, 1985.
- [12] Gupta, S.P., *QSAR studies on drug acting at the central nervous system*, *Chemical Reviews*, 89, 1765-1800, 1989.
- [13] Carter, M.D., Stephenson, V.C., Weaver, D.F., *Are anticonvulsants 'two thirds' of local anesthetics? A quantum pharmacology study*, *Journal of Molecular Structure: THEOCHEM*, 638, 57-62, 2003.
- [14] Tasso, S.M., Bruno-Blanch, L.E., Moon, S.C., Estiu, G.L., *Pharmacophore searching and QSAR analysis in the design of anticonvulsant drugs*. *Journal of Molecular Structure: THEOCHEM*, 504, 229-240, 2000.
- [15] Serdaroğlu, G., Ortiz, J.V., *Ab initio calculations on some antiepileptic drugs such as phenytoin, phenobarbital, ethosuximide and carbamazepine*, *Structural Chemistry*, 28, 957-964, 2017.
- [16] Fong, C.W., *Statins in therapy: understanding their hydrophilicity, lipophilicity, binding to 3-hydroxy-3-methylglutaryl-CoA reductase, ability to cross the blood brain barrier and metabolic stability based on electrostatic molecular orbital studies*. *European Journal of Medicinal Chemistry*, 85, 661-674, 2014.
- [17] Hansch, C., Björkroth, J. P., and Leo, A., *Hydrophobicity and central nervous system agents: on the principle of minimal hydrophobicity in drug design*. *Journal of Pharmaceutical Sciences*, 76(9), 663-687, 1987.
- [18] Michalik, M., Lukes, V., *The validation of quantum chemical lipophilicity prediction of alcohols*, *Acta Chimica Slovaca*, 9(2), 89-94, 2016.
- [19] Iwase, K., Komatsu, K., Hirono, S., Nakagawa, S., Moriguchi, I., *Estimation of hydrophobicity based on the solvent-accessible surface area of molecules*. *Chemical and Pharmaceutical Bulletin*, 33(5), 2114-2121, 1985.
- [20] Du, Q., Arteca, G.A., Mezey, P.G., *Heuristic lipophilicity potential for computer-aided rational drug design*, *Journal of Computer-Aided Molecular Design*, 11, 503-515, 1997.
- [21] Hansch, C. and Dunn, W.J., *Linear relationships between lipophilic character and biological activity of drugs*, *Journal of Pharmaceutical Sciences*, 61(1), 1-19, 1972.

- [22] Hansch, C., Fujita, T., *p- σ - π Analysis. A method for the correlation of biological activity and chemical structure*, Journal of the American Chemical Society, 85, 1616-1626, 1964.
- [23] Fujita, T., Hansch, C., Iwasa, J., *The correlation of biological activity of plant growth regulators and chloromycetin derivatives with hammett constants and partition coefficient*, Journal of the American Chemical Society, 85, 2817-2824, 1963.
- [24] Pinal, R., Yalkowsky, S. H., *Solubility and partitioning VII: solubility of barbiturates in water*, Journal of Pharmaceutical Sciences, 76(1), 75-85, 1987.
- [25] Nadendla, R.R., *Principles of organic medicinal chemistry*. New Age International (P) Ltd., New Delhi, Publishers, chapter 3, 16-17p, 2005.
- [26] Sekulić, T.D., Smoliński, A., Mandić, A., Lazić, A., *Chromatographic and in silico assessment of logP measures for new spirohydantoin derivatives with anticancer activity*, Journal of Chemometrics, 32, 1-13, 2018.
- [27] Frisch, M.J., Trucks, G.W., Schlegel, H. B., Scuseria, G. E., Robb, M. A. et al. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford, CT 2009.
- [28] Dennington, R., Keith, T., Millam, J., 2009, Gauss View, Version 5., Semichem Inc., Shawnee Mission, KS.
- [29] Becke, A.D., *A new mixing of Hartree–Fock and local density functional theories*, The Journal of Chemical Physics, 98, 1372–1377, 1993.
- [30] Lee, C., Yang, W., Parr, R.G., *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Physical Review B, 37, 785–789, 1988.
- [31] Tomasi, J., Mennucci, B., Cance`s, E., *The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level*, Journal of Molecular Structure (Theochem), 464, 211–226, 1999.
- [32] Mennucci, B., Cance`s, E., and Tomasi, J., *Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications*, The Journal of Physical Chemistry B, 101, 10506-10517, 1997.
- [33] Cance`s, E., Mennucci, B., Tomasi, J., *A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics*, The Journal of Physical Chemistry, 107, 3032-3041, 1997.
- [34] Foresman, J.B., Keith, T.A., Wiberg, K.B., Snoonian, J., Frisch, M.J., *Solvent effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation on ab initio reaction field calculations*, The Journal of Physical Chemistry, 100, 16098–16104, 1996.
- [35] Koopmans, T. *Über die zuordnung von wellenfunktionen und eigenwertenzu den einzelnen elektronen eines atoms*, Physica., 1, 104-113, 1934.
- [36] Fukui, K., *The Role of Frontier Orbitals in chemical reactions*, Science, 218: 747–754, 1982.
- [37] Janak, J.F., *Proof that $\partial E/\partial n_i = \epsilon_i$ in density-functional theory*, Physical Review B, 18 (12), 7165-7168, 1978.
- [38] Parr, R.G., and Pearson, R.G., *Absolute hardness: companion parameter to absolute electronegativity*, Journal of the American Chemical Society, 105, 7512-7516, 1983.
- [39] Pearson, R.G., *Absolute electronegativity and hardness correlated with the molecular orbital theory*, Proceedings of the National Academy of Sciences, USA.83, 8440-8441, 1986.

[40] Parr, R.G., Szentpaly, L.V., Liu, S., *Electrophilicity Index*, Journal of the American Chemical Society, 121, 1922-1924, 1999.

[41] Perdew, J.P., Levy, M., *Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities*, Physical Review Letters, 51 (20), 1884-1887, 1983.

[42] Perdew, J.P., Parr, R.G., Levy, M., Balduz, J.L., *Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy*, Physical Review Letters, 49 (23), 1691- 1694, 1982.

[43] Foresman, J.B., Frisch, A.E. *Exploring chemistry with electronic structure methods*, third edition, Gaussian, Inc. Wallingford, CT USA, 2015.

[44] Garrido, N.M., Queimada, A. J., Jorge, M., Macedo, E. A., and Economou, I. G., *1-Octanol/water partition coefficients of n-alkanes from molecular simulations of absolute solvation free energies*, Journal of Chemical Theory and Computation, 5, 2436-2446, 2009.

[45] Pernot, P., Civalleri, B., Presti, D., Savin, A., *Prediction uncertainty of density functional approximations for properties of crystals with cubic symmetry*, The Journal of Physical Chemistry A, 119, 5288–5304, 2015.

[46] Lewis, D.F.V., *The calculation of molar polarizabilities by the CNDO/2 method: correlation with the hydrophobic parameter, logP*, Journal of Computational Chemistry, 10, 145-151, 1989.

[47] Gao, S., Cao, C., *A new approach on estimation of solubility and n-octanol/ water partition coefficient for organohalogen compounds*, International Journal of Molecular Sciences, 9, 962-977, 2008.

[48] Zhou, W., Zhai, Z., Wang, Z., Wang, L., *Estimation of n-octanol/water partition coefficients (K_{ow}) of all PCB congeners by density functional theory*, Journal of Molecular Structure: THEOCHEM, 755, 137–145, 2005.

[49] Padmanabhan, J., Parthasarathi, R., Subramanian, V., Chattaraj, P. K., *QSPR models for polychlorinated biphenyls:n-octanol/water partition coefficient*, Bioorganic & Medicinal Chemistry, 14, 1021–1028, 2006.

[50] <http://www.molinspiration.com>

[51] Gaillard, P., Carrupt, P.A., Testa, B., Boudon, A., *Molecular lipophilicity potential, a tool in 3D QSAR: method and applications*, Journal of Computer-Aided Molecular Design, 8, 83–96, 1994.

[52] Pajouhesh, H., Lenz, G.R., *Medicinal chemical properties of successful central nervous system drugs*, NeuroRx, 2, 541–553, 2005.



Evaluation of Radiation Attenuation Properties of Some Cancer Drugs

İlyas ÇAĞLAR^{1,*}, Gülçin BİLGİCİ CENGİZ²

¹*Kafkas University, Kazım Karabekir Vocational School of Technical Sciences, Department of Electricity and Energy, 36100 Kars, Türkiye*

ilyas.caglar@kafkas.edu.tr, ORCID: 0000-0002-6958-8469

²*Kafkas University, Faculty of Arts and Sciences, Department of Physics, 36100 Kars, Türkiye*

gulcincengiz@kafkas.edu.tr, ORCID: 0000-0002-6164-3232

Received: 06.08.2021

Accepted: 10.12.2021

Published: 31.12.2021

Abstract

The present study was conducted to estimate the radiation attenuation parameters of six different antineoplastic drugs used in the cure of cancer diseases. The effective atomic number and electron density of anastrozole, epirubicin, gemcitabine, ifosfamide, methotrexate and paclitaxel were computed theoretically in the energy region of 1 keV to 100 GeV. The energy absorption buildup factors (EABF) and exposure buildup factors (EBF) for these chemotherapy drugs were also examined by applying Geometric Progression (GP) fitting method. The variation of EABF and EBF values with photon energy and penetration depth were presented graphically and discussed. The results obtained from this study pointed out that buildup factors rely on the chemical combination of the drugs, incident photon energy and penetration thickness. It was observed that ifosfamide had a significantly better radiation absorption effect compared to other drugs. The data obtained from this study are expected to be useful in the fields of radiation biology, radiation dosimetry and radiotherapy.

Keywords: Cancer drugs; Radiation; Effective atomic number; Effective electron density; Buildup factors.

Bazı Kanser İlaçlarının Radyasyon Soğurma Özelliklerinin Değerlendirilmesi



Öz

Bu çalışma, kanser hastalıklarının tedavisinde kullanılan altı farklı antineoplastik ilacın radyasyon zayıflama parametrelerini değerlendirmek için gerçekleştirilmiştir. Anastrozol, epirubisin, gempitabin, ifosfamid, metotreksat ve paklitakselin etkin atom numarası ve etkin elektron yoğunluğu 1 keV ile 100 GeV enerji bölgesinde teorik olarak hesaplandı. Bu kemoterapi ilaçları için enerji soğurma yığılma faktörleri (EABF) ve maruz kalma yığılma faktörleri (EBF) de GP fit yöntemi uygulanarak incelenmiştir. EABF ve EBF değerlerinin foton enerjisi ve nüfuz etme derinliği ile değişimi grafiksel olarak sunulmuş ve tartışılmıştır. Bu çalışmadan elde edilen sonuçlar, yığılma faktörlerinin ilaçların kimyasal kombinasyonuna, gelen foton enerjisine ve nüfuz etme kalınlığına bağlı olduğuna işaret etti. İfosfamidin diğer ilaçlara göre önemli ölçüde daha iyi radyasyon soğurma etkisine sahip olduğu gözlemlendi. Bu çalışmadan elde edilen verilerin radyasyon biyolojisi, radyasyon dozimetrisi ve radyoterapi alanlarında faydalı olması beklenmektedir.

Anahtar Kelimeler: Kanser ilaçları; Radyasyon; Etkin atom numarası; Etkin elektron yoğunluğu; Yığılma faktörleri.

1. Introduction

Ionizing radiation (X, gamma rays, etc.) performs a major act in the diagnosis and cure of illness in medical applications such as diagnostic radiology, radiotherapy and like, as it has the ability to acquire images and destroy cells or tumours [1]. Radiotherapy technologies, which are being updated day by day with the use of high energy radiation in cancer treatments, are now being implemented as a priority option for many types of cancer. Today, 60-70% of cancer cases receive radiotherapy at least once in the process after the disease is diagnosed [2]. In some cases, simultaneously chemotherapy and radiotherapy are effectively applied in the treatment of some types of cancer, and this treatment is called chemoradiotherapy. In this way, by increasing the sensitivity of the cells to radiation with chemotherapy, radiation is provided to be more effective on the cells. When ionizing radiation passes through biological tissues, it can cause chemical alterations in tissues, leading to cell damage or cell and tissue death [3]. Therefore, evaluation of the interaction parameters of X or gamma rays with chemotherapy drugs may be beneficial for the estimation of absorbed radiation doses and radiation dose limits in chemoradiotherapy treatments.

The prior knowledge of radiation attenuation parameters such as mass attenuation coefficients (μ_m), effective atomic number (Z_{eff}), effective electron density (N_{e1}) and buildup factors are critical in various practices such as medical physics, radiation physics, radiation

dosimetry, radiotherapy, computerized tomography, and radiation biology. The most basic parameter among these parameters is the mass attenuation coefficient, which measures the probability of photon interaction (absorption or scattering) with the drug sample and this parameter and other parameters can be calculated using μ_m [4]. Buildup factor, an important term in radiation dosimetry besides that shield design can be categorized as energy absorption buildup factor (EABF) and exposure buildup factor (EABF). It depends on the atomic number of the absorber medium [5]. EABF is defined as the amount of absorbed or deposited energy in the interacting material and the detector response function is as the absorption in the interacting medium. EBF is defined as the amount of exposure and the detector response function is as the absorption in air [2, 6]. Various methods have been developed in the literature to work out the buildup factors that take into account various parameters such as photon energy, absorbing medium properties and distance [7-10]. The GP fitting method [7, 11] is the most widely used method to calculate buildup factors of various materials. Using the GP fitting method, many researchers have studied buildup factors of various materials such as concretes [12], alloys [13, 14], glasses [15, 16], polymers [17], building materials [18] human organs and tissues [19], teeth [20], bioactive compounds [21], amino acids [22], enzyme inhibitors [23], thermoluminescent dosimetric (TLD) materials [24], solvents [6] and like. In these studies, it was emphasized that GP method is a suitable method for calculating photon buildup factors for various materials.

Antineoplastic drugs used in the cure of cancer illness are drugs conceived to disturb or avoid cellular proliferation by inhibiting deoxyribonucleic acid (DNA) synthesis. Antineoplastic drugs are generally classified as alkylating agents, antimetabolites, antitumor antibiotics, hormone and hormone antagonists, alkaloids and other antineoplastic drugs [25]. Anastrozole ($C_{17}H_{19}N_5$), epirubicin ($C_{27}H_{26}NO_{11}$), gemcitabine ($C_9H_{11}F_2N_3O_4$), ifosfamide ($C_7H_{15}Cl_2N_2O_2P$), methotrexate ($C_{20}H_{22}N_8O_5$) and paclitaxel ($C_{47}H_{51}NO_{14}$) are some of the commonly used antineoplastic drugs in chemotherapy. Anastrozole is an aromatase inhibitor utilized in the handling of second-level breast cancer and avoids the production of the hormone oestrogen, which triggers the formation of breast cancer. Epirubicin is an anthracycline antitumor antibiotic and is used alone or in combination with other chemotherapy drugs to cure certain diseases such as breast and ovarian cancer. Ifosfamide, an alkylating oxazophosphorine agent, is one of the chemotherapy drugs that avoid the reproduction of cancer cell DNA by averting its reproduction. Today, it is used in the treatment of lymphoma, soft tissue sarcoma and advanced breast, testicle, ovarian, stomach and lung cancers. Methotrexate is a type of antimetabolite agent that inhibits DNA replication or causes apoptosis by synthesizing incorrect codes and is extensively used in the cure of various types of cancer diseases such as head and neck cancers, ovarian, bladder, cervix, stomach, large intestine, testicle, breast, bone cancer, choriocarcinoma cancers and etc.

Similar to methotrexate, Gemcitabine is a type of antimetabolite agent. It is a chemotherapy drug used in the treatment of pancreas, lung, bladder and breast cancers as well as other tumours such as ovarian cancer, mesothelioma and head and neck cancers. Paclitaxel is an antitumor agent active against a wide variety of cancers that are generally considered to be resistant to conventional chemotherapy. It is effective in the treatment of metastatic breast or ovarian cancer [25-31].

Radio protective effects of various drugs have been estimated by many researchers. Oto et al. computed gamma ray interaction parameters (i.e. μ_m , Z_{eff} , N_{el} , EABF and EBF) of different drugs used in cholinergic medications using WinXCOM computer program [1]. Sayyed et al. calculated Z_{eff} , N_{el} , EABF and EBF for Nonsteroidal anti-inflammatory drugs (NSAIDs) [32]. Kavaz et al. computed photon buildup factors of some chemotherapy drugs by using the GP fitting method in the energy region 0.015–15 MeV up to penetration depths of 40 mean free paths (mfp) [33]. Akman and Kaçal calculated some essential radiation attenuation parameters such as μ_m , Z_{eff} and N_{el} of some drugs used in Chemotherapy with the help of the WinXCOM program [34]. Ekinici et al. investigated the EABF and EBF of some anti-inflammatory drugs by using the GP fitting method [35]. Yorgun and Kavaz determined μ_m , Z_{eff} , N_{el} of some cancer drugs at 13.81, 17.7, 26.34 and 59.54 keV photon energies. They also computed EABF and EBF of these cancer drugs in the energy region 0.015–15 MeV up to penetration depths of 40 mfp [2].

The aim of this research work is to estimate radiation interaction parameters of six different antineoplastic chemotherapy drugs commonly used for cancer treatment. There is almost no study in the literature on the radiation absorption parameters of these drugs. Therefore, it may be useful to investigate the radiation interaction parameters of these drugs for chemoradiation dose limits and dose calculations. For this purpose, Z_{eff} and N_{el} values of some chemotherapy drugs such as anastrozole, epirubicin, gemcitabine, ifosfamide, methotrexate and paclitaxel were computed in the energy range of 1 keV–100 GeV. Additionally, the energy absorption and exposure buildup factors of these antineoplastic chemotherapy drugs were calculated by means of GP fitting method for the energy range 0.015-15 MeV up to the penetration depth of 40 mfp.

2. Materials and Methods

2.1. Computation of Z_{eff} and N_{el}

The mass attenuation coefficients for investigated drug samples were computed by using mixture rule given in the following equation [36, 37]:

$$\mu_m = \frac{\mu}{\rho} = \sum_i w_i \left(\frac{\mu}{\rho}\right)_i \quad (1)$$

where ρ is the density, w_i and $\left(\frac{\mu}{\rho}\right)_i$ are the weight fraction and mass attenuation coefficient for individual element in drugs, respectively. The theoretical μ_m values for the investigated drugs were determined using WinXCOM software package [38]. For the drug samples, the effective atomic number were computed with the help of the following formula [1, 33]:

$$Z_{eff} = \frac{\sum_i f_i A_i \left(\frac{\mu}{\rho}\right)_i}{\sum_j f_j \frac{A_j}{Z_j} \left(\frac{\mu}{\rho}\right)_j} \quad (2)$$

where f_i , A_i , Z_i and $\left(\frac{\mu}{\rho}\right)_i$ are the molar fraction, atomic weight, atomic number and mass attenuation coefficient of relative element in the drug samples, respectively. In addition, the effective electron density, a parameter closely related to the effective atomic number, can be calculated by following equation [33, 35]:

$$N_{el} = N_A \frac{n Z_{eff}}{\sum_i n_i A_i} = N_A \frac{Z_{eff}}{\langle A \rangle} \text{ (electrons/g)} \quad (3)$$

where N_A represents the Avogadro constant and $\langle A \rangle$ indicates average atomic mass of the material.

2.2. Computation of EABF and EBF

The energy absorption and exposure buildup of chemotherapy drugs under the study were determined in three steps. In the first step, Compton partial mass attenuation coefficient $((\mu/\rho)_{Compton})$ and the total mass attenuation coefficient $(\mu/\rho)_{Total}$ values were determined for different elements ($Z = 4-30$) and also for investigated drug samples using WinXCOM program. Then the interpolation formula given in Eqn. (4) was employed to compute the equivalent atomic number (Z_{eq}) of chosen drug by matching the ratio $R ((\mu_m)_{Compton}/(\mu_m)_{Total})$ at a particular photon energy with the convenient ratio of the pure element at the same energy [7,39];

$$Z_{eq} = \frac{Z_1(\log R_2 - \log R) + Z_2(\log R - \log R_1)}{\log R_2 - \log R_1} \quad (4)$$

where R signifies the ratio for chosen drug samples at particular energy which lies between R_1 and R_2 . Z_1 and Z_2 denote atomic numbers of the elements corresponding to the ratios R_1 and R_2 , respectively. More details for the calculation procedure can be found in reference [38]. In the

second step, in order to evaluate buildup factors the GP fitting coefficients for elements were acquired from the ANSI/ANS-6.4.3 database [40]. This database provides the GP fitting parameters for 23 elements ($Z=4-92$), water, air and concrete in the energy region of 0.015-15 MeV up to 40 mfp [39]. The obtained Z_{eq} values were used to determine geometric progression (GP) fitting coefficients ($a, b, c, d,$ and X_k) for the drug samples using the following relation [33, 39, 41];

$$P = \frac{P_1(\log Z_2 - \log Z_{eq}) + P_2(\log Z_{eq} - \log Z_1)}{\log Z_2 - \log Z_1} \tag{5}$$

where P denotes GP fitting parameters of studied drug samples. P_1 and P_2 are the values of GP fitting coefficients corresponding to the Z_1 and Z_2 atomic numbers at a specific energy, respectively. In the last step, the computed GP fitting coefficients were used to generate the energy absorption and exposure build-up for selected antineoplastic drugs at some standard photon energies in the energy range 0.015-15 MeV up to 40 mfp penetration depths. This calculation was accomplished by using the following GP fitting formulas [7, 33, 39, 41];

$$B(E, X) = 1 + \frac{b - 1}{K - 1} (K^X - 1) \quad \text{for } K \neq 1 \tag{6}$$

$$B(E, X) = 1 + (b - 1) \quad \text{for } K = 1 \tag{7}$$

$$K(E, X) = cx^a + d \frac{\tanh\left(\frac{x}{X_k} - 2\right) - \tanh(-2)}{1 - \tanh(-2)} \quad \text{for } x \leq 40 \text{ mfp} \tag{8}$$

where E, x and K (E, X) are the photon energy, penetration depth in mfp and dose multiplicative factor, respectively. a, b, c, d and X_k are the GP fitting parameters and b is the buildup factor at one mfp.

3. Results and Discussion

The chemical formula and elemental composition of the studied antineoplastic chemotherapy drugs are presented in Table 1. The μ_m values of anastrozol, epirubicin, gemcitabine, ifosfamide, methotrexate and paclitaxel drugs were computed using WinXCOM software package [37]. From the computed μ_m values Z_{eff} and N_{el} values of investigated drugs were worked out with the help of the Eqn. (2) and (3) in the energy range from 1 keV to 100 GeV.

Table 1: Chemical formula and elemental composition of investigated drugs

Drug	Chemical Formula	Weight fraction of elements (%)					
		H	C	N	O	F	P

Anastrozole	C ₁₇ H ₁₉ N ₅	0.065	0.696	0.239	-	-	-	-
Epirubicin	C ₂₇ H ₂₆ NO ₁₁	0.048	0.600	0.026	0.326	-	-	-
Gemcitabine	C ₉ H ₁₁ F ₂ N ₃ O ₄	0.042	0.411	0.160	0.243	0.144	-	-
Ifosfamide	C ₇ H ₁₅ Cl ₂ N ₂ O ₂ P	0.058	0.322	0.107	0.123	-	0.119	0.272
Methotrexate	C ₂₀ H ₂₂ N ₈ O ₅	0.049	0.529	0.247	0.176	-	-	-
Paclitaxel	C ₄₇ H ₅₁ NO ₁₄	0.060	0.661	0.016	0.262	-	-	-

Figure 1 and 2 show the variation of the computed Z_{eff} and N_{el} values with photon energy for the investigated antineoplastic chemotherapy drugs. Z_{eff} and N_{el} values of ifosfamide have a peak at 1.892 keV photon energy, which correspond to the K absorption edge of chlorine. The computed Z_{eff} values ranged from 3.68-6.30 for Anastrozole, 4.35-6.98 for epirubicin, 4.69-7.47 for gemcitabine, 4.70–12.35 for ifosfamide, 4.32-6.81 for methotrexate and 4.00-6.83 for paclitaxel, respectively. The N_{el} values also ranged from 3.10×10^{23} – 4.26×10^{23} electrons/g for Anastrozole, 3.31×10^{23} – 4.11×10^{23} electrons/g for epirubicin, 3.30×10^{23} – 4.05×10^{23} electrons/g for gemcitabine, 3.15×10^{23} – 8.26×10^{23} electrons/g for ifosfamide, 3.30×10^{23} – 4.09×10^{23} electrons/g for methotrexate and 3.35×10^{23} – 4.28×10^{23} electrons/g for paclitaxel, respectively. The highest Z_{eff} and N_{el} values of the studied drugs were observed in the low energy region ($E < 0.1$ MeV) where photoelectric absorption which is cross section proportional to the Z^{4-5} and $E^{-3.5}$ is the dominant interaction mechanism [23].

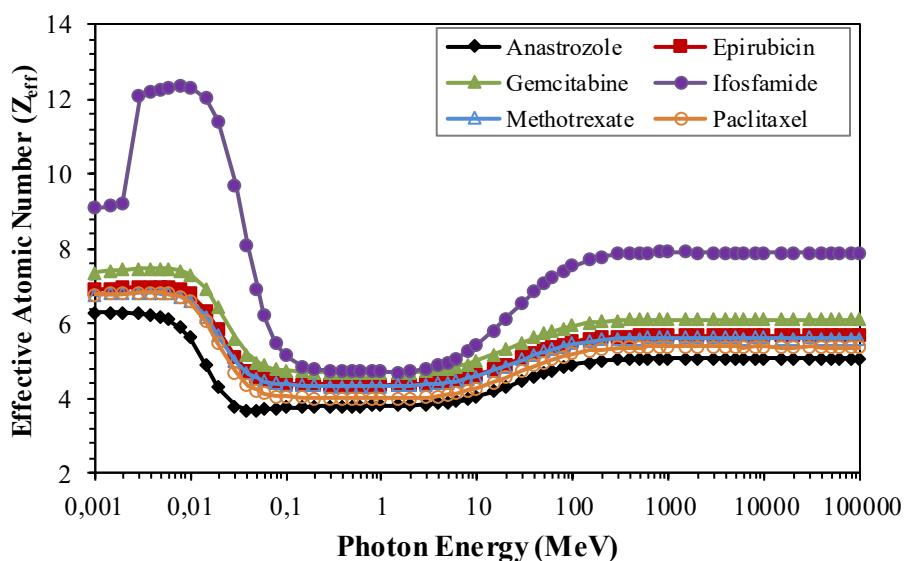


Figure 1: Z_{eff} values of investigated chemotherapy drugs versus photon energy

As seen from Figs. 1 and 2, the lowest Z_{eff} and N_{el} values of the studied drugs were determined in the intermediate energies ($0.1 \text{ MeV} < E < 10 \text{ MeV}$). It was seen that in this energy region where Compton scattering is the most important interaction process, Z_{eff} and N_{el} values are almost independent to photon energy. This may be due to the weak dependence of the Compton scattering cross section to atomic number and photon energy (i.e. proportional to Z and E^{-1}) [37].

At energies greater than 10 MeV, Z_{eff} and N_{el} values increase slowly with increasing photon energy and are almost constant at further energies. This change can be clarified by actually that pair production is the dominant interaction process at high energies. Because the pair productions cross section is directly proportional to E and Z^2 [13]. This observed trend in Z_{eff} and N_{el} values is consistent with the results of the study reported by Oto et al. [1] who investigated radiation interaction parameters of some cholinergic drugs. As shown in Figs. 1 and 2, ifosfamide has considerably higher Z_{eff} and N_{el} values than other studied drugs in the low and high energy regions. In the intermediate energy region, the Z_{eff} values of ifosfamide are slightly larger than those of the other drugs, while the N_{el} values are approximately the same. The reason for this apparent difference in the Z_{eff} and N_{el} values of ifosfamide may be that ifosfamide contains phosphate ($Z=15$, weight fraction= 0.119) and chlorine ($Z=17$, weight fraction =0.272), unlike other drugs.

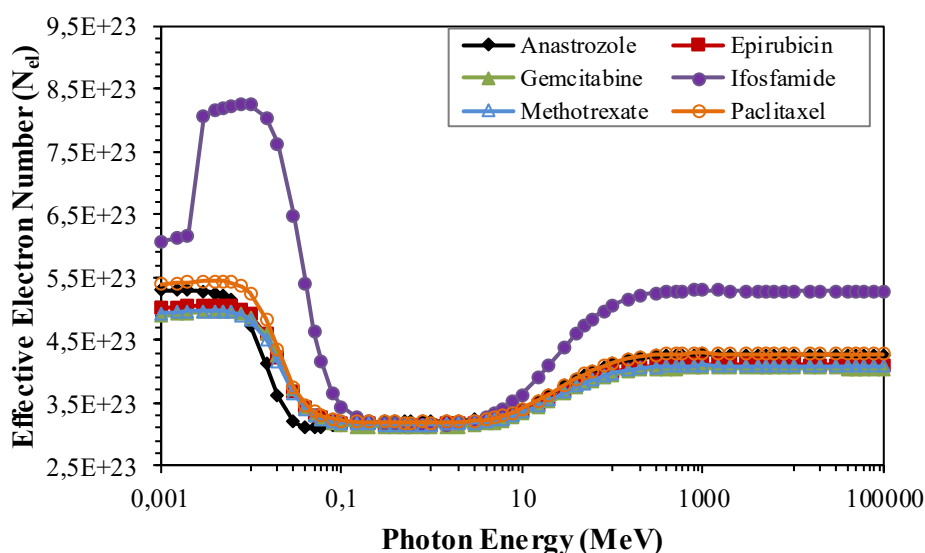


Figure 2: N_{el} values of investigated chemotherapy drugs versus photon energy

The Z_{eq} values calculated using the interpolation formula given in Eqn. (4) in the 0.015-15 MeV energy range for the examined drugs are given in Table 2. It is obviously sighted from Table 2 that ifosfamide has the highest Z_{eq} values among the examined antineoplastic drugs, while anastrozole has the lowest Z_{eq} values. The high Z_{eq} values of ifosfamide are due to the presence of phosphate ($Z=15$, weight fraction= 0.119) and chlorine ($Z=17$, weight fraction=0.272) in the chemical structure of ifosfamide, unlike other drugs. Similarly, the reason why anastrozole has low Z_{eq} values is due to the lack of oxygen, fluorine, phosphate and chlorine in the chemical structure of anastrozole. The GP fitting coefficients of the EABF and EBF for investigated radioprotectors are listed in Table 3-8 at the energy region of 0.015-15 MeV.

Table 2: Z_{eq} values of investigated drugs at the energy region of 0.015-15 MeV

Energy (MeV)	Anastrozole	Epirubicin	Gemcitabine	Ifosfamide	Methotrexate	Paclitaxel
0.015	6.009	6.639	7.094	11.85	6.514	6.446
0.02	6.014	6.636	7.099	11.97	6.509	6.443
0.03	6.008	6.618	7.097	12.09	6.489	6.422
0.04	5.989	6.598	7.086	12.13	6.468	6.399
0.05	5.970	6.578	7.072	12.14	6.451	6.378
0.06	5.951	6.562	7.059	12.13	6.436	6.360
0.08	5.922	6.535	7.036	12.08	6.413	6.331
0.1	5.901	6.515	7.017	12.00	6.396	6.309
0.15	5.867	6.483	6.980	11.79	6.368	6.274
0.2	5.850	6.463	6.954	11.62	6.352	6.253
0.3	5.831	6.444	6.929	11.39	6.335	6.232
0.4	5.823	6.435	6.917	11.26	6.327	6.222
0.5	5.819	6.429	6.910	11.18	6.323	6.217
0.6	5.816	6.427	6.906	11.14	6.321	6.214
0.8	5.813	6.424	6.903	11.09	6.319	6.211
1	5.813	6.423	6.902	11.08	6.318	6.210
1.5	5.673	6.262	6.701	10.28	6.177	6.032
2	5.630	6.206	6.637	9.77	6.126	5.968
3	5.620	6.193	6.623	9.60	6.114	5.953
4	5.617	6.190	6.620	9.56	6.111	5.949
5	5.617	6.189	6.620	9.53	6.110	5.948
6	5.617	6.188	6.618	9.53	6.110	5.947
8	5.616	6.186	6.617	9.51	6.108	5.944
10	5.617	6.184	6.617	9.51	6.106	5.942
15	5.618	6.184	6.614	9.50	6.107	5.942

Table 3: GP fitting parameters for anastrozole in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.397	0.526	0.153	14.421	-0.077	1.385	0.538	0.146	14.310	-0.072
0.02	1.895	0.737	0.077	16.436	-0.037	1.872	0.730	0.080	16.546	-0.040
0.03	3,716	1.150	-0.026	12.601	0.008	3.503	1.150	-0.026	12.846	0.008
0.04	5.073	1.740	-0.128	14.102	0.056	5.276	1.750	-0.130	13.917	0.064
0.05	5.598	2.078	-0.165	14.488	0.070	6.749	2.123	-0.171	14.344	0.074
0.06	5.420	2.304	-0.187	14.668	0.079	7.270	2.408	-0.200	14.522	0.089
0.08	4.841	2.510	-0.206	14.823	0.083	6.888	2.725	-0.230	14.456	0.101
0.1	4.280	2.569	-0.210	14.926	0.083	6.107	2.836	-0.239	14.368	0.104
0.15	3.538	2.484	-0.203	15.230	0.076	4.573	2.892	-0.247	14.292	0.107
0.2	3.184	2.360	-0.194	15.211	0.073	3.775	2.771	-0.239	14.949	0.108
0.3	2.785	2.142	-0.176	14.990	0.070	3.148	2.494	-0.219	14.343	0.099
0.4	2.613	1.942	-0.155	14.758	0.063	3.027	2.268	-0.201	13.324	0.085
0.5	2.452	1.811	-0.140	15.794	0.062	2.661	2.072	-0.180	13.936	0.084
0.6	2.392	1.673	-0.121	14.937	0.047	2.544	1.904	-0.159	13.615	0.068
0.8	2.196	1.567	-0.110	14.100	0.047	2.320	1.726	-0.139	13.727	0.068
1	2.087	1.463	-0.094	14.158	0.041	2.203	1.567	-0.114	13.737	0.057
1.5	1.939	1.277	-0.061	14.308	0.027	2.031	1.330	-0.073	13.718	0.037
2	1.840	1.173	-0.039	14.390	0.016	1.918	1.199	-0.046	14.147	0.023
3	1.715	1.051	-0.012	13.942	0.004	1.764	1.062	-0.015	12.253	0.008
4	1.627	0.989	0.003	13.594	-0.003	1.664	0.983	0.005	22.560	-0.007
5	1.567	0.944	0.015	14.609	-0.008	1.584	0.937	0.017	14.663	-0.011
6	1.521	0.901	0.029	12.647	-0.017	1.531	0.907	0.026	14.563	-0.016

8	1.438	0.874	0.037	11.762	-0.018	1.443	0.870	0.037	16.022	-0.030
10	1.382	0.858	0.040	14.389	-0.022	1.377	0.854	0.042	12.741	-0.020
15	1.287	0.837	0.047	15.238	-0.030	1.280	0.837	0.047	14.811	-0.028

Table 4: GP fitting parameters for epirubicin in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.292	0.499	0.160	14.518	-0.078	1.286	0.497	0.162	14.284	-0.081
0.02	1.677	0.634	0.114	15.440	-0.056	1.659	0.637	0.112	15.544	-0.055
0.03	3.081	0.956	0.023	14.782	-0.021	2.945	0.955	0.024	14.657	-0.022
0.04	4.400	1.450	-0.080	13.915	0.032	4.450	1.458	-0.082	13.707	0.034
0.05	5.176	1.806	-0.132	14.170	0.056	5.711	1.829	-0.136	14.048	0.059
0.06	5.261	2.073	-0.166	14.143	0.073	6.222	2.124	-0.172	14.036	0.078
0.08	4.890	2.342	-0.194	14.045	0.083	6.040	2.465	-0.209	13.808	0.094
0.1	4.495	2.388	-0.195	14.666	0.080	5.526	2.561	-0.213	14.395	0.093
0.15	3.636	2.395	-0.197	14.684	0.078	4.175	2.659	-0.227	14.125	0.100
0.2	3.266	2.283	-0.188	14.784	0.076	3.579	2.558	-0.221	14.231	0.098
0.3	2.817	2.095	-0.172	14.613	0.068	3.061	2.287	-0.197	14.248	0.086
0.4	2.625	1.913	-0.152	14.554	0.062	2.778	2.097	-0.179	13.733	0.075
0.5	2.459	1.794	-0.138	15.185	0.059	2.604	1.932	-0.160	14.163	0.071
0.6	2.383	1.671	-0.121	14.650	0.048	2.482	1.800	-0.143	13.827	0.059
0.8	2.200	1.555	-0.107	14.139	0.045	2.281	1.642	-0.124	13.902	0.057
1	2.096	1.447	-0.090	14.430	0.038	2.161	1.513	-0.104	13.864	0.049
1.5	1.938	1.276	-0.060	14.315	0.026	1.999	1.300	-0.066	14.001	0.031
2	1.841	1.169	-0.038	14.403	0.015	1.889	1.188	-0.043	13.981	0.020
3	1.714	1.051	-0.011	14.104	0.003	1.745	1.059	-0.014	12.385	0.006
4	1.627	0.988	0.004	13.123	-0.003	1.649	0.987	0.004	23.623	-0.007
5	1.565	0.944	0.015	14.740	-0.008	1.572	0.939	0.017	14.315	-0.011
6	1.514	0.907	0.028	13.308	-0.018	1.523	0.907	0.027	13.980	-0.016
8	1.430	0.881	0.034	12.096	-0.017	1.437	0.872	0.037	16.010	-0.031
10	1.375	0.861	0.040	14.322	-0.022	1.371	0.859	0.041	12.715	-0.021
15	1.281	0.838	0.047	15.732	-0.033	1.275	0.841	0.046	15.225	-0.030

Table 5: GP fitting parameters for gemcitabine in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.228	0.479	0.166	14.429	-0.081	1.226	0.471	0.173	14.292	-0.087
0.02	1.538	0.570	0.137	14.855	-0.068	1.523	0.577	0.133	14.963	-0.065
0.03	2.657	0.824	0.059	15.588	-0.038	2.563	0.829	0.056	15.778	-0.040
0.04	3.914	1.246	-0.045	13.748	0.014	3.871	1.253	-0.047	13.568	0.016
0.05	4.838	1.606	-0.107	13.918	0.045	4.964	1.614	-0.109	13.826	0.047
0.06	5.119	1.896	-0.148	13.764	0.069	5.454	1.913	-0.151	13.684	0.070
0.08	4.914	2.213	-0.185	13.441	0.105	5.407	2.268	-0.192	13.316	0.110
0.1	4.660	2.245	-0.183	14.466	0.089	5.195	2.326	-0.191	14.380	0.082
0.15	3.745	2.302	-0.189	14.420	0.077	3.977	2.461	-0.210	14.111	0.095
0.2	3.357	2.202	-0.180	14.799	0.076	3.410	2.420	-0.209	13.395	0.090
0.3	2.838	2.061	-0.168	14.259	0.067	2.974	2.160	-0.182	14.083	0.077
0.4	2.621	1.900	-0.151	14.259	0.060	2.711	1.993	-0.165	14.005	0.070
0.5	2.461	1.786	-0.138	14.283	0.056	2.548	1.850	-0.148	14.135	0.063
0.6	2.366	1.679	-0.124	14.305	0.050	2.418	1.747	-0.135	14.065	0.058
0.8	2.202	1.546	-0.105	14.190	0.044	2.247	1.590	-0.114	14.009	0.050
1	2.106	1.432	-0.087	14.672	0.036	2.129	1.480	-0.097	13.953	0.044
1.5	1.934	1.276	-0.060	14.355	0.026	1.980	1.282	-0.061	14.493	0.027

2	1.838	1.173	-0.039	14.151	0.017	1.872	1.182	-0.041	13.964	0.019
3	1.711	1.054	-0.012	13.224	0.004	1.730	1.060	-0.014	13.243	0.005
4	1.628	0.984	0.006	13.777	-0.005	1.639	0.988	0.004	19.265	-0.006
5	1.566	0.937	0.018	14.068	-0.012	1.567	0.940	0.018	13.913	-0.012
6	1.504	0.922	0.022	15.397	-0.017	1.520	0.904	0.029	13.151	-0.017
8	1.430	0.874	0.037	12.066	-0.021	1.429	0.880	0.035	13.645	-0.023
10	1.369	0.866	0.039	14.326	-0.022	1.365	0.866	0.039	13.530	-0.022
15	1.276	0.839	0.048	15.335	-0.034	1.273	0.841	0.047	15.125	-0.032

Table 6: GP fitting parameters for ifosfamide in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.039	0.402	0.209	13.063	-0.127	1.039	0.398	0.213	13.098	-0.131
0.02	1.087	0.428	0.183	14.507	-0.094	1.086	0.438	0.179	14.318	-0.092
0.03	1.288	0.442	0.192	14.242	-0.102	1.282	0.447	0.190	14.444	-0.102
0.04	1.625	0.548	0.146	15.245	-0.077	1.598	0.550	0.146	15.117	-0.078
0.05	2.175	0.615	0.133	13.753	-0.072	2.050	0.638	0.121	14.471	-0.063
0.06	2.732	0.767	0.081	13.357	-0.056	2.429	0.781	0.075	14.637	-0.057
0.08	3.748	1.048	0.002	14.331	-0.016	2.926	1.030	0.007	13.644	-0.021
0.1	4.267	1.285	-0.049	12.623	0.008	3.101	1.229	-0.035	12.292	-0.004
0.15	4.088	1.619	-0.107	13.330	0.039	3.064	1.478	-0.079	15.006	0.016
0.2	3.636	1.721	-0.122	13.723	0.044	2.897	1.570	-0.094	15.490	0.023
0.3	3.037	1.743	-0.126	13.991	0.046	2.634	1.605	-0.101	15.270	0.026
0.4	2.729	1.688	-0.120	14.211	0.042	2.469	1.575	-0.099	15.188	0.027
0.5	2.541	1.628	-0.113	14.286	0.040	2.343	1.540	-0.096	15.076	0.029
0.6	2.404	1.565	-0.103	14.563	0.036	2.249	1.494	-0.089	15.400	0.026
0.8	2.228	1.467	-0.090	14.856	0.032	2.115	1.419	-0.080	15.223	0.025
1	2.112	1.386	-0.077	14.872	0.028	2.021	1.356	-0.070	15.699	0.023
1.5	1.937	1.255	-0.054	14.286	0.021	1.890	1.241	-0.051	15.049	0.018
2	1.843	1.159	-0.035	14.701	0.013	1.811	1.158	-0.035	14.792	0.013
3	1.708	1.053	-0.010	12.225	0.000	1.696	1.054	-0.011	11.500	0.001
4	1.617	0.984	0.008	13.076	-0.010	1.614	0.991	0.005	16.231	-0.008
5	1.548	0.944	0.019	12.944	-0.014	1.544	0.952	0.002	14.850	-0.016
6	1.491	0.919	0.026	15.568	-0.027	1.499	0.915	0.006	13.305	-0.024
8	1.398	0.900	0.032	12.303	-0.020	1.411	0.897	0.003	12.969	-0.023
10	1.339	0.882	0.038	13.916	-0.028	1.352	0.874	0.005	13.557	-0.030
15	1.241	0.872	0.043	14.734	-0.034	1.264	0.832	0.058	14.664	-0.048

Table 7: GP fitting parameters for methotrexate in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.312	0.504	0.158	14.500	-0.078	1.305	0.505	0.159	14.289	-0.079
0.02	1.720	0.654	0.107	15.635	-0.052	1.700	0.655	0.106	15.740	-0.052
0.03	3.210	0.996	0.013	14.338	-0.015	3.059	0.995	0.014	14.288	-0.016
0.04	4.540	1.509	-0.090	13.954	0.037	4.619	1.518	-0.092	13.749	0.039
0.05	5.265	1.861	-0.139	14.236	0.059	5.923	1.888	-0.143	14.110	0.062
0.06	5.297	2.120	-0.170	14.255	0.074	6.436	2.180	-0.178	14.143	0.080
0.08	4.884	2.375	-0.196	14.214	0.083	6.207	2.515	-0.213	13.948	0.095
0.1	4.452	2.424	-0.198	14.714	0.081	5.608	2.620	-0.219	14.396	0.095
0.15	3.610	2.417	-0.199	14.748	0.078	4.222	2.707	-0.232	14.128	0.102
0.2	3.244	2.302	-0.190	14.781	0.075	3.619	2.591	-0.224	14.430	0.100
0.3	2.813	2.103	-0.173	14.695	0.068	3.082	2.317	-0.200	14.287	0.088
0.4	2.626	1.916	-0.152	14.623	0.062	2.794	2.121	-0.182	13.670	0.077

0.5	2.458	1.796	-0.138	15.394	0.060	2.617	1.951	-0.162	14.170	0.073
0.6	2.387	1.669	-0.121	14.730	0.047	2.497	1.813	-0.145	13.772	0.060
0.8	2.199	1.557	-0.108	14.127	0.045	2.288	1.654	-0.126	13.877	0.059
1	2.094	1.450	-0.091	14.374	0.039	2.169	1.521	-0.105	13.844	0.050
1.5	1.939	1.275	-0.060	14.307	0.027	2.003	1.303	-0.067	13.901	0.032
2	1.842	1.168	-0.037	14.452	0.015	1.893	1.190	-0.043	13.985	0.020
3	1.714	1.051	-0.011	14.272	0.003	1.748	1.058	-0.014	12.222	0.006
4	1.626	0.989	0.003	12.999	-0.003	1.650	0.986	0.004	24.450	-0.008
5	1.564	0.945	0.015	14.868	-0.008	1.573	0.939	0.017	14.391	-0.011
6	1.516	0.904	0.029	12.913	-0.018	1.524	0.907	0.026	14.137	-0.015
8	1.430	0.882	0.034	12.102	-0.017	1.438	0.870	0.037	16.455	-0.033
10	1.376	0.860	0.040	14.321	-0.022	1.372	0.857	0.041	12.562	-0.021
15	1.282	0.838	0.047	15.805	-0.033	1.275	0.841	0.046	15.244	-0.030

Table 8: GP fitting parameters for paclitaxel in the energy range 0.015-15 MeV

Energy (MeV)	EABF					EBF				
	b	c	a	Xk	d	b	c	a	Xk	d
0.015	1.323	0.507	0.158	14.490	-0.078	1.315	0.509	0.158	14.291	-0.078
0.02	1.742	0.665	0.103	15.739	-0.050	1.723	0.665	0.103	15.844	-0.050
0.03	3.279	1.016	0.008	14.103	-0.012	3.119	1.016	0.008	14.093	-0.013
0.04	4.616	1.542	-0.095	13.975	0.039	4.711	1.551	-0.097	13.772	0.041
0.05	5.317	1.893	-0.143	14.274	0.061	6.046	1.923	-0.147	14.146	0.064
0.06	5.318	2.148	-0.173	14.324	0.075	6.568	2.215	-0.181	14.209	0.082
0.08	4.880	2.396	-0.198	14.330	0.083	6.321	2.549	-0.216	14.043	0.096
0.1	4.421	2.450	-0.201	14.750	0.081	5.668	2.663	-0.223	14.397	0.097
0.15	3.588	2.435	-0.200	14.801	0.079	4.262	2.746	-0.235	14.131	0.103
0.2	3.224	2.319	-0.191	14.778	0.075	3.655	2.620	-0.226	14.608	0.101
0.3	2.808	2.111	-0.173	14.775	0.069	3.101	2.346	-0.203	14.324	0.090
0.4	2.627	1.919	-0.153	14.691	0.062	2.809	2.145	-0.185	13.607	0.078
0.5	2.457	1.797	-0.138	15.606	0.060	2.630	1.971	-0.165	14.176	0.075
0.6	2.391	1.667	-0.120	14.812	0.047	2.512	1.825	-0.147	13.716	0.060
0.8	2.199	1.559	-0.108	14.115	0.045	2.296	1.666	-0.128	13.852	0.061
1	2.092	1.454	-0.091	14.316	0.039	2.177	1.528	-0.107	13.822	0.051
1.5	1.941	1.275	-0.060	14.293	0.027	2.010	1.310	-0.069	13.729	0.033
2	1.843	1.167	-0.037	14.518	0.014	1.900	1.193	-0.044	14.003	0.021
3	1.715	1.050	-0.011	14.451	0.003	1.753	1.058	-0.014	12.013	0.006
4	1.626	0.990	0.003	12.920	-0.002	1.654	0.986	0.004	25.241	-0.008
5	1.564	0.947	0.014	14.992	-0.007	1.576	0.939	0.017	14.521	-0.011
6	1.519	0.900	0.030	12.390	-0.018	1.526	0.908	0.026	14.387	-0.015
8	1.431	0.883	0.034	12.061	-0.016	1.440	0.868	0.038	16.930	-0.034
10	1.379	0.859	0.040	14.330	-0.022	1.374	0.855	0.042	12.407	-0.021
15	1.284	0.838	0.047	15.811	-0.032	1.277	0.840	0.046	15.202	-0.029

The variation of EABF and EBF values with incident photon energy for anastrozole, epirubicin, gemcitabine, ifosfamide, methotrexate and paclitaxel at some chosen penetration depth were plotted in Figs. 3 and 4, respectively. It was monitored that the EABF and EBF values of the studied drugs at 1, 5, 10 and 40 mfp, increased with increasing energy, reached the maximum value in the intermediate energies and then decreased again in the further energies. As can be seen from Figs. 3 and 4, the EABF and

EBF values of ifosfamide are smaller than EABF and EBF values of other drugs. The maximum EABF and EBF values were seen at the 0.3 MeV photon energy for ifosfamide and 0.1 MeV photon energy for anastrozole, epirubicin, gemcitabine, methotrexate and paclitaxel. This trend observed in EABF and EBF values can be explained on the basis that while Compton scattering is the dominant interaction mechanism in the medium energy region, photoelectric absorption and pair production are the dominant interaction mechanisms in the low and high energy region, respectively. On the other hand, this trend is also in line with the observation of Sayyed et al. [32] who estimated photon buildup factors of some anti-inflammatory drugs. It was also seen that the values of EABF and EBF increased with increasing depth of penetration and became very high for the greatest at penetration depth 40 mfp. This increase is a result of multiple scattering events for large penetration depths [17].

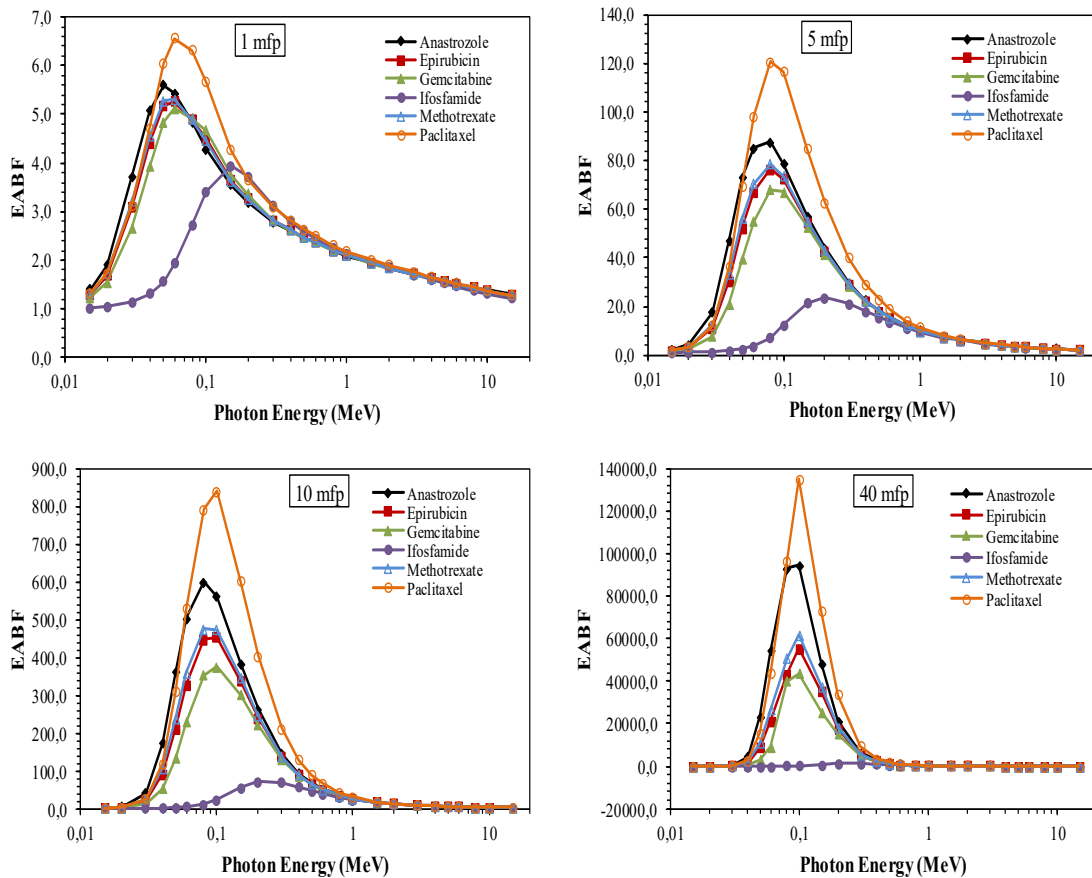


Figure 3: The EABF values of drugs in the energy range of 0.015-15 MeV at 1, 5, 10 and 40 mfp

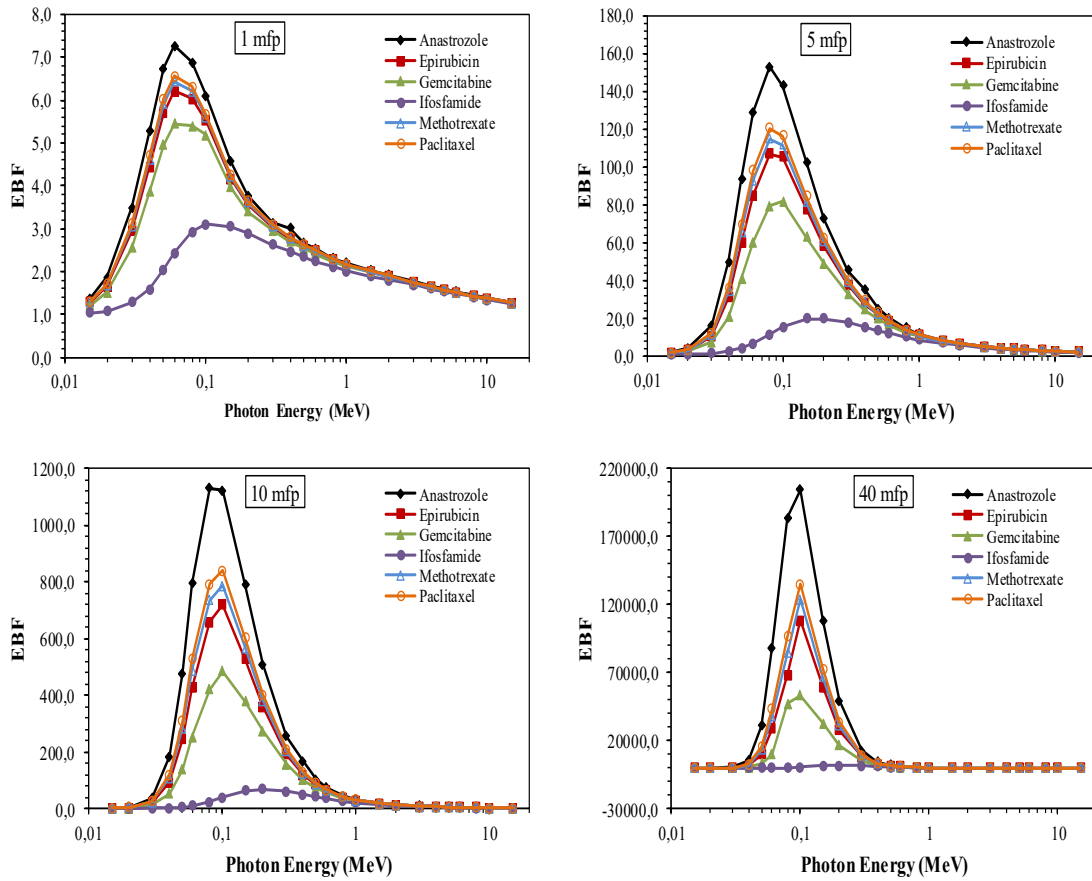


Figure 4: The EBF values of drugs in the energy range of 0.015-15 MeV at 1, 5, 10 and 40 mfp

The EABF and EBF values of the investigated anti neoplastic chemotherapy drugs against the mfp for 0.015, 0.15, 1.5 and 15 MeV photon energies are given in Figs. 5 and 6, respectively. From Figs. 5 and 6, it was observed that the EABF and EBF values of the studied chemotherapy drugs increased with the increase in mfp values. It is clear from these figures that EABF and EBF values of the investigated drugs at 0.15 and 1.5 MeV energies are higher than the others. It was seen that ifosfamide, which has the highest Z_{eq} value, has the smallest EABF and EBF values at low energies (0.015 and 0.15 MeV). On the other hand, Anastrozole has the highest EABF and EBF values at 0.015 and 0.15 MeV photon energies owing to its low Z_{eq} value. Also, the EABF and EBF values of ifosfamide remains nearly constant at 0.015 MeV energy, while the EABF and EBF values of other drugs increase with increasing mpf values. At 0.15 MeV, EABF and EBF values of the studied chemotherapy drugs increased with the rising up in mfp values. The maximum values were observed for anastrozole and paclitaxel and minimum value was observed for ifosfamide in this energy. It was seen that the EABF and EBF values of drug samples

decreased with the increasing Z_{eq} values at lower than the 0.15 MeV, because the cross section of photoelectric absorption, which is the effective interaction process at low energies, is strongly dependent on Z_{eq}^{4-5} . Contrary to others, the EABF and EBF values for 1.5 MeV photon energy are almost independent of the chemical composition (i.e. Z_{eq}) of the drugs as reported by Kavaz et al. [34]. This result can be clarified by the predominance of Compton scattering at 1.5 MeV energy. As can be seen from Figs. 5 and 6, the values of EABF and EBF of the drugs increase with increase in Z_{eq} at 15 MeV between 10 and 40 mfp and ifosfamide showed higher values than other drugs due to the dominance of pair production in this region. This result is agreeing with the findings reported by Kavaz et al. [42] who evaluated EABF and EBF factors of some radio protective agents.

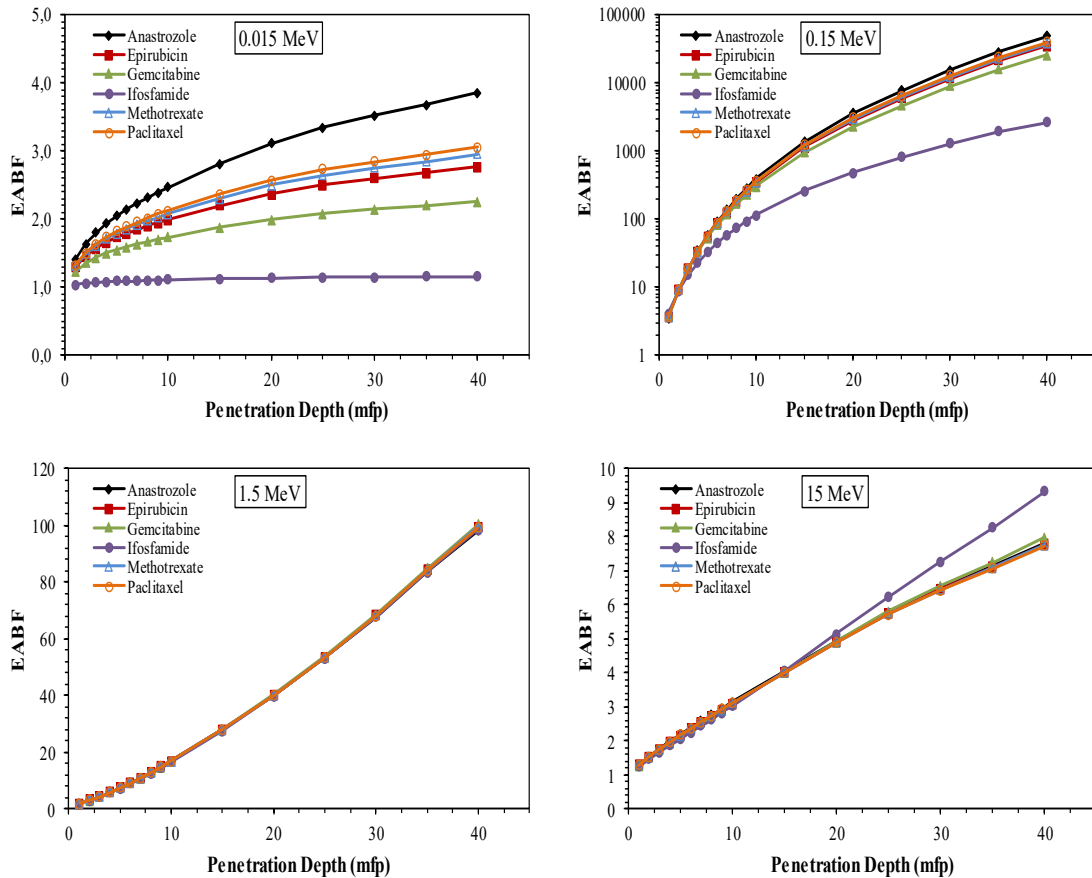


Figure 5: The energy absorption buildup factor for the drugs up to 40 mfp at 0.015, 0.15, 1.5 and 15 MeV

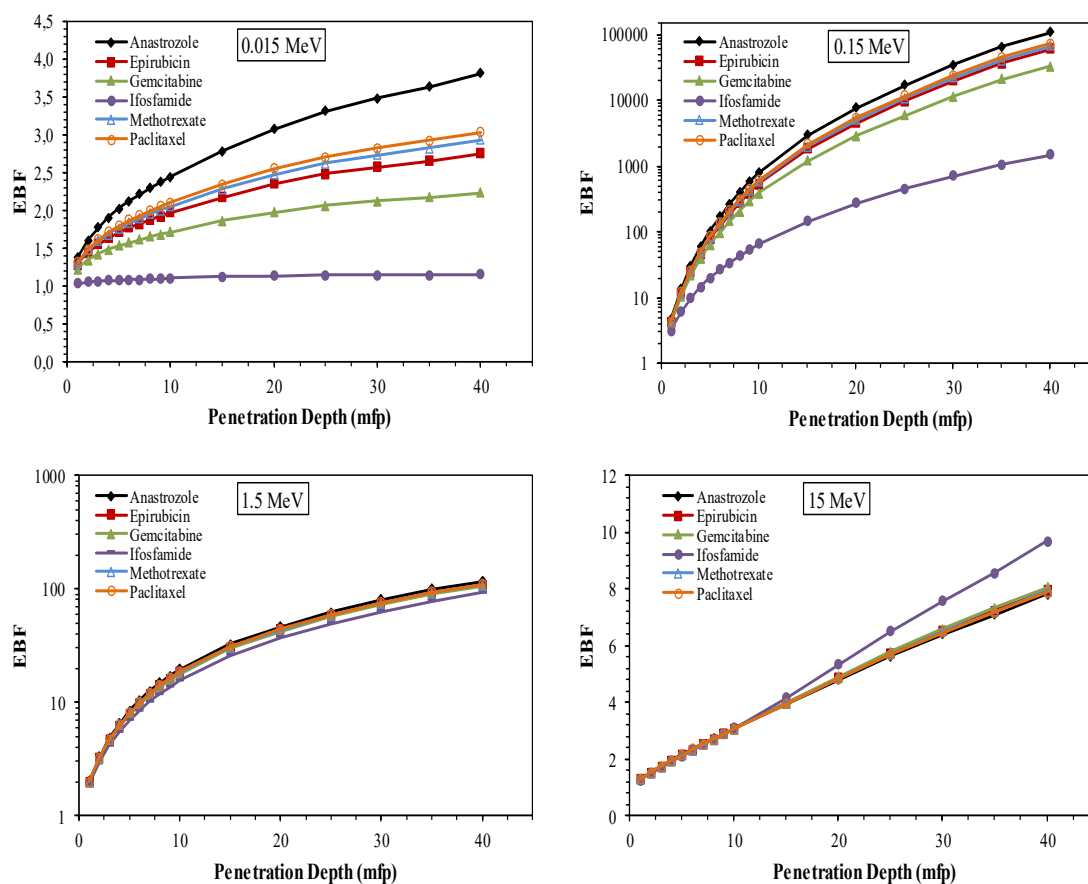


Figure 6: The energy exposure buildup factor for the drugs up to 40 mfp at 0,015, 0.15, 1.5 and 15 MeV

4. Conclusion

The present study was carried out to obtain information on photon interaction parameters of different antineoplastic chemotherapy drugs. The energy absorption buildup factors and exposure buildup factors of anastrozole ($C_{17}H_{19}N_5$), epirubicin ($C_{27}H_{26}NO_{11}$), gemcitabine ($C_9H_{11}F_2N_3O_4$), ifosfamide ($C_7H_{15}Cl_{12}N_2O_2P$), methotrexate ($C_{20}H_{22}N_8O_5$) and paclitaxel ($C_{47}H_{51}NO_{14}$) chemotherapy drugs were computed using GP fitting method. Also, Z_{eff} and N_{el} values were determined with the help of the WinXCOM program for photon energies from 1 keV to 100 GeV. The results obtained in this study showed that Z_{eff} and N_{el} values are dependent on the photon energy. The highest Z_{eff} values were found for ifosfamide and the lowest Z_{eff} values for anastrozole. This study also showed that buildup factors vary depending on the chemical composition (i.e. Z_{eq}) of the drugs, photon energy and mean free path. Among the studied samples, ifosfamide has the largest buildup factor values at 15 MeV photon energy, while it has the smallest values at 0.015, 0.15 and 1.5 MeV photon energies. It was concluded that ifosfamide has

better photon absorption properties since it contains phosphorus and chlorine in its chemical structure, unlike other drugs. It is predictable that the results of this work will be beneficial in areas such as radiation dosimetry and chemoradiotherapy.

References

- [1] Oto, B., Oto, G., Madak, Z., Kavaz, E., *The interaction of gamma radiation with drugs used in cholinergic medications*, *Journal of Radiation Biology*, 96(2), 236-244, 2020.
- [2] Yorgun, N.Y., Kavaz, E., *Gamma photon protection properties of some cancer drugs for medical applications*. *Results in Physics*, 13 (102150), 1-6, 2019.
- [3] Veranda, E., Tavares, D., *Radioprotection mechanisms and radioprotective agents, including honeybee venom*, *Journal of Venomous Animals and Toxins*, 4(1), 5-20, 1988.
- [4] Sayyed, M.I., Kaky, K.M., Şakar, E., Akbaba, U., Taki, M.M., Agar, O., *Gamma radiation shielding investigations for selected germanate glasses*, *Journal of Non-Crystalline Solids*, 512, 33–40, 2019.
- [5] Manohara, S.R., Hanagotimath. S.M., Gerward, L. *Energy absorption buildup factors of human organs and tissues at energies and penetration depths relevant for radiotherapy and diagnostics*, *Journal of Applied Clinical Medical Physics*, 12, 3557–3566, 2011.
- [6] Singh, S.P., Singh, T., Kaur, P., *Variation of energy absorption buildup factors with incident photon energy and penetration depth for some commonly used solvents*, *Annals of Nuclear Energy*, 35, 1093–1097, 2008.
- [7] Harima, Y., Sakamoto, Y., Tanaka, S., Kawai, M. *Validity of the geometric progression formula in approximating gamma ray buildup factors*, *Nuclear Science and Engineering*, 94, 24-35, 1986.
- [8] Shimizu, A., *Calculations of gamma ray buildup factors up to depths of 100 mfp by the method of invariant embedding, (I) analysis of accuracy and comparison with other data*, *Journal of Nuclear Science and Technology*, 39, 477-486, 2002.
- [9] Suteau, C., Chiron, M., *An iterative method for calculating gamma ray buildup factors in multi-layer shields*, *Radiation Protection Dosimetry*, 116, 489-492, 2005.
- [10] Sardari, D., Saudi, S., Tajik, M., *Evaluation of gamma ray buildup factor data in water with MCNP4C code*, *Annals of Nuclear Energy*, 38(23), 628-631, 2011.
- [11] Harima Y., *An approximation of gamma ray buildup factors by modified geometrical progression*, *Nuclear Science and Engineering*, 83, 299-309, 1983.
- [12] Singh, V.P., Badiger, N.M., *Investigation on radiation shielding parameters of ordinary, heavy and super heavy concretes*, *Nuclear Technology and Radiation Protection* 29, 149–156, 2014.
- [13] Şakar, E., Büyükyıldız, M., Alım, B., Şakar, B.C., Kurudirek, M., *Leaded brass alloys for gamma-ray shielding applications*, *Radiation Physics and Chemistry*, 159, 64–69, 2019.
- [14] Şakar, E., *Determination of photon-shielding features and build-up factors of nickel–silver alloys*, *Radiation Physics and Chemistry*, 172, 108778, 2021.

- [15] Kaur, P., Singh, D., Singh, T., *Gamma ray shielding and sensing application of some rare earth doped lead-alumino-phosphate glasses*, *Radiation Physics and Chemistry*, 144, 336–343, 2018.
- [16] Rammah, Y.S., Özpolat, Ö.F., Alım, B., Şakar, E., El-Mallawany, R., El-Agawany F.I., *Assessment of gamma-ray attenuation features for La⁺³ co-doped zinc borotellurite glasses*, *Radiation Physics and Chemistry*, 176, 109069, 2020.
- [17] Sayyed, M., AlZaatreh, M., Matori, K., Sidek, H., Zaid, M., *Comprehensive study on estimation of gamma ray exposure buildup factors for smart polymers as a potent application in nuclear industries*, *Results in Physics*, 9, 585–592, 2018.
- [18] Sharaf, J.M., Saleh, H., *Gamma-ray energy buildup factor calculations and shielding effects of some Jordanian building structures*, *Radiation Physics and Chemistry*, 110, 87–95, 2015.
- [19] Manjunatha, H.C., Rudraswamy, B., *Computation of exposure build-up factors in teeth*, *Radiation Physics and Chemistry*, 80 (1), 14–21, 2011.
- [20] Kurudirek, M., Topcuoglu, S., *Investigation of human teeth with respect to the photon interaction, energy absorption and buildup factor*, *Nuclear Instruments and Methods in Physics Research B*, 269, 1071–1081, 2011.
- [21] Yilmaz, D., Gedik, Z., Tugrak, M., Gul, H.I., *Energy absorption buildup factors of some potential bioactive compounds in the energy region 0.015-15 MeV*, *Spectroscopy Letters*, 50(6), 301-306, 2017.
- [22] Bursalıoglu, E., Balkan, B., Kavanoz, H.B., Okutan, M., Icelli, O., Yalcın, Z., *Energy absorption and exposure buildup factors of essential amino acids*, *Biomed Research International*, 359754, 2014.
- [23] Turhan, M.F., Durak, R., Kaçal, M.R., *Determination of Gamma Ray Buildup Factors of Some Enzyme Inhibitors*, *International Journal of Scientific and Engineering Research*, 10 (9), 8-13, 2019.
- [24] Manohara, S.R., Hanagodimath, S.M., Gerward, L., *Energy absorption buildup factors for thermoluminescent dosimetric materials and their tissue equivalent*, *Radiation Physics and Chemistry*, 79, 575–582, 2010.
- [25] Kummerer, K., Haiss, A., Schuster, A., Hein, A., Ebert, I., *Antineoplastic compounds in the environment-substances of special concern*, *Environmental Science and Pollution Research*, 23(15), 14791-14804, 2016.
- [26] Yeo, W., Johnson, P.J., *Radiation-recall skin disorders associated with the use of antineoplastic drugs, Pathogenesis, prevalence, and management*, *American Journal of Clinical Dermatology*, 1, 113–116, 2000.
- [27] Wiseman, L.R., Adkins J.C., *Anastrozole: a review of its use in the management of postmenopausal women with advanced breast cancer*, *Drugs Aging*, 13, 321-32, 1998.
- [28] Coukell, A.J., Faulds, D., *Epirubicin: An updated review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in the management of breast cancer*, *Drugs*, 53, 453-482, 1997.
- [29] Mini, E., Nobili, S., Caciagli, B., Landini, I., Mazzei, T., *Cellular pharmacology of gemcitabine*, *Annals of Oncology*, 17, 7–12, 2006.

[30] Dechant, K.L., Brogden, R.N., Pilkington, T., Faulds, D., *Ifosfamide/mesna: A review of its antineoplastic activity, pharmacokinetic properties and therapeutic efficacy in cancer*, *Drugs* 42, 428–467, 1991.

[31] Rowinsky, E.K., Donehower, R.C., *Paclitaxel (taxol)*, *The New England Journal of Medicine*, 332, 1004 -1014, 1995.

[32] Sayyed M.I., Issa S.A., Auda S.H., *Assessment of radio-protective properties of some anti-inflammatory drugs*, *Progress in Nuclear Energy*, 100, 297–308, 2017.

[33] Kavaz, E., Ahmadishadbad, N., Özdemir, Y., *Photon buildup factors of some chemotherapy drugs*, *Biomed Pharmacother*, 69, 34–41, 2015.

[34] Akman, F., Kaçal, M.R., *Investigation of radiation attenuation parameters of some drugs used in Chemotherapy in Wide Energy Region*, *Journal of Radiology and Oncology*, 2, 047-052, 2018.

[35] Ekinci, N., Kavaz, E., Ozdemir, Y., *A study of the energy absorption and exposure buildup factors of some anti-inflammatory drugs*, *Applied Radiation and Isotopes*, 90, 265–273, 2014.

[36] Jackson, D.F., Hawkes, D.J., *X-ray attenuation coefficients of elements and mixtures*, *Physics Reports*, 70, 169–233, 1981.

[37] Alim, B., Şakar, E., Baltakesmez, A., Han, İ., Sayyed, M.I., Demir, L., *Experimental investigation of radiation shielding performances of some important AISI-coded stainless steels: Part I*, *Radiation Physics and Chemistry*, 160, 108455, 2020.

[38] Gerward, L., Guilbert, N.K., Jensen, B., Levring, H., *WinXCom—a program for calculating X-ray attenuation coefficients*, *Radiation Physics and Chemistry*, 71(3–4), 653–654, 2004.

[39] Şakar, E., Özpolat, O.F., Alim, B., Sayyed, M.I., Kurudirek, M., *Phy-X/PSD: Development of a user friendly online software for calculation of parameters relevant to radiation shielding and dosimetry*, *Radiation Physics and Chemistry*, 108496, 1-12, 2020.

[40] ANSI/ANS-6.4.3, *Gamma ray attenuation coefficient and buildup factors for engineering materials*, Illinois:American Nuclear Society, La Grange Park, 1991.

[41] Raut, S.D., Awasarmol, V.V., Shaikh, S.F., Ghule, B.G., Ekar, S.U., Mane, R.S., Pawar, P.P., *Study of gamma ray energy absorption and exposure buildup factors for ferrites by geometric progression fitting method*, *Radiation Effects and Defects in Solids*, 173 (3-4), 429-438, 2018.

[42] Kavaz, E., Perisanoglu, U., Ekinci, N., Ozdemir, Y., *Determination of energy absorption and exposure buildup factors by using GP fitting approximation for radioprotective agents*, *International Journal of Radiation Biology*, 92(7); 380–387, 2016.