

Evaluation of Irisin Levels in Cancer Anorexia Cachexia Syndrome and the Relationship between Nutrition Education and Quality of Life

Diler Us Altay ^{1,a,*}, Duygu Mataraci Değirmenci ^{1,b}, Salih Can Çelik ^{2,c}, Abdullah Üner ^{3,d}, Tefvik Noyan ^{3,e}, Çağrı Akalin ^{4,f}

¹Department of Nutrition and Dietetics, Faculty of Health Science, Ordu University, Ordu, Türkiye

²Department of General Surgery, Ordu State Hospital, Ordu, Türkiye

³Department of Medical Biochemistry, Faculty of Medicine, Ordu University, Ordu, Türkiye

⁴Department of General Surgery, Faculty of Medicine, Ordu University, Ordu, Türkiye

*Corresponding author

Research Article

History

Received: 03/05/2024

Accepted: 14/11/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This study examines the effects of nutritional education given to individuals diagnosed with cancer anorexia cachexia syndrome (CACS) on serum irisin levels, cachectic factors, proinflammatory cytokines, quality of life scale results, and anthropometric and routine biochemical parameters. Forty-one patients diagnosed with CACS were randomly divided into two groups, experimental and control. Nutrition education was given for 12 weeks to the patients in the experimental group (n=23), while those in the control group (n=18) were not subjected to any intervention. All participants' serum irisin, proteolysis-inducing factor, zinc- α -2 glycoprotein, interleukin-6, tumor necrosis factor- α , routine biochemical parameters, and body weight were measured at the beginning and end of the study. The patients were also evaluated in terms of 24-hour recall food intake, body mass index, and quality of life scale values. No significant differences were observed at baseline between the experimental and control groups in terms of quality of life scale values, cachectic factors, inflammatory cytokines, or irisin levels. However, at the end of the study, hemoglobin levels were higher in the experimental group than in the control group. Energy and nutrient intakes were similar between the groups initially, but were higher in the experimental group at the end of the study. Nutrition education did not significantly alter the quality of life scale, BMI, or biochemical parameters. However, education yielded the expected increase in nutrient intake in the experimental group.

Keywords: CACS, QOL, Cachectic factors, Irisin, Inflammatory cytokines.

^a surelid@hotmail.com

^c salihcancelik@gmail.com

^e tevfiknoyan@hotmail.com

^b <https://orcid.org/0000-0002-0465-8403>

^d <https://orcid.org/0000-0002-6537-2206>

^f <https://orcid.org/0000-0002-7733-0177>

^b duygu.mataraci@hotmail.com

^d abdullahuner66.an@gmail.com

^f dr.cagriakalin@gmail.com

^b <https://orcid.org/0000-0003-2136-1545>

^d <https://orcid.org/0000-0001-7060-2216>

^f <https://orcid.org/0000-0003-3370-9879>

Introduction

Cancer is the uncontrolled proliferation of a group of cells in the body. Side-effects related to cancer and treatments thereof are among the issues that particularly challenge patients and physicians [1]. The second main cause of cancer-related deaths after sepsis (multiple organ failure due to infection), cachexia progresses with weight loss and cannot be reversed using normal nutritional support [2]. Cancer cachexia is characterized by the co-occurrence of decreased energy intake and increased energy expenditure, leading to a negative energy balance. The primary contributors to reduced energy intake are loss of appetite (anorexia), dysphagia, pain, fatigue, and depression or anxiety. Cachectic cancer patients have a low tolerance for chemo- and radiation therapies and a low quality of life [3,4]. Recent studies have shown that cancer cachexia is associated with anorexia (loss of appetite), a condition known as cancer anorexia cachexia syndrome (CACS). Patients with CACS experience loss of appetite, along with involuntary weight loss, which reduces their quality of life. CACS is frequently seen in individuals with advanced disease, the occurrence of the syndrome depending on the type of cancer. Weight loss is observed before starting treatment in 50-85% of

gastrointestinal, pancreatic, lung, and colorectal cancers. CACS is implicated in 20% of cancer-related deaths. However, cachexia is often considered unimportant and is frequently left untreated [5].

The etiology of CACS is multifactorial, and the biochemical mechanisms responsible have not yet been fully elucidated. In this process, proteolysis-inducing factor (PIF) and zinc- α -2 glycoprotein (ZAG), recognized cachectic factors, are released from tumor cells and cause muscle and fat atrophy. Tumor necrosis factor- α (TNF- α), interleukin-1 (IL-1), IL-6, epidermal growth factor (EGF), transforming growth factor (TGF), and platelet-derived growth factor (PDGF) are proinflammatory cytokines associated with the development of cancer cachexia, although the search for biomarkers for the early diagnosis of cachexia continues. Irisin is a newly discovered adipocytokine known to exhibit anti-inflammatory, anti-diabetic, anti-apoptotic, and anti-obesity effects that cause weight loss [6,7]. Irisin levels have been measured in various types of cancer. However, no study to date has investigated whether irisin represents a cachectic factor in patients with CACS. In addition to traditional mortality and morbidity parameters, quality of life is an important factor

that should be taken into consideration when determining individuals' health levels. The Turkish-language version of the 36-item Short Form Survey (SF-36) is a tool that can be used in cancer research in Türkiye [8]. SF-36 was used to evaluate quality of life in CACS patients in the present study before and after nutrition education. This study set out to find answers to the following questions;

- a) Is there any relationship between cachectic factors, pro-inflammatory cytokines, and irisin in serum in patients diagnosed with CACS?
- b) can irisin represent a novel cachectic factor for patients diagnosed with CACS? And
- c) can the provision of nutritional education for these patients provide quality survival in line with data obtained from the quality-of-life assessment scale?

Within the scope of this project, irisin, cachectic factors and cytokines were measured at the protein level in serum samples taken from 41 patients diagnosed with CACS using the enzyme linked immunosorbent assay (ELISA) method. Body weight, the amount of food consumed in the previous 24 hours, BMI (kg/m^2), routine biochemical tests, and quality of life values were evaluated at baseline and the 12th week for all patients, with and without nutrition education.

Materials and Methods

This prospective study was planned between December 2021 and December 2022. The sample collection time was extended by one year due to the COVID-19 pandemic.

Study Population

The study protocol was approved by the Ordu University (ODU) Clinical Research Ethics Committee (Decision No: 2021/184). This investigation involved 41 patients diagnosed with CACS. The patients were recruited from the Ordu University Training and Research Hospital General Surgery and Ordu State Hospital gastroenterology clinics.

Clinical findings of CACS

- > 6% weight loss in the previous six months, skeletal muscle loss (biceps and quadriceps),
- anorexia, weakness, fatigue, decreased movements and quality of life, decreases daily performance scores,
- increased acute phase response C-reactive protein (CRP), fibrinogen, and α -1 antitrypsin)
- decreases in serum proteins (albumin and total protein)
- anemia (a decrease in hemoglobin level)s, increased resting energy use, decreased lean body mass and body mass.

CACS should be considered and treated promptly in the presence of these clinical findings in cancer patients. More importantly, malnutrition should be detected long before the patient reaches this condition, and screening should even be performed to identify patients at nutritional risk. Nutritional screening can be performed

using the Nutrition Risk Screening (NRS-2002) scoring system [9]. Homogeneity of the study population is important, and every effort was made in this study to ensure that there were no significant differences between the groups in terms of age, gender, or BMI. Inclusion criteria were age over 18 years, similar age groups, and clinical findings of CACS being confirmed by a clinician. Exclusion criteria were age under 18, a history of any surgical procedure in the previous weeks, receipt of radiotherapy, chemotherapy, or drug therapy capable of significantly modulating metabolism and weight, diagnoses of different cancers, and having been previously followed-up in the outpatient clinic.

Patients who met the inclusion and exclusion criteria were followed-up in the general and gastroenterological surgery outpatient clinic. The research was explained, and those individuals willing to take part signed informed consent forms and received nutrition education from a nutrition and dietetics specialist for 12 weeks (three months). No intervention was performed on the patients in the control group during the study period.

Data Collection Tools Used in Assessing Nutritional Status

National risk screening (NRS-2002)

The NRS screening instrument was used to assess the risk of malnutrition in the cancer patients enrolled in the trial. This evaluates the likelihood of developing undernutrition as well as the presence of undernutrition. The main screening, which inquires into nutritional status disorders (BMI and previous unplanned weight loss), and grades the severity of disease with its metabolic effects, is carried out if any of the four prescreening questions are answered affirmatively. A further point is added for patients aged 70 or older. A total score of 3 or higher indicates that the patient is at nutritional risk.

24-Hour food consumption record

The 24-hour dietary recall approach was used to determine food intake. A qualified dietitian interviewed the patients by telephone, asking them to recall their entire food and fluid intake on the previous day. The investigator investigated each patient's eating habits in detail. The number of units consumed was determined by describing quantities in household measures (such as pieces, slices, tea- or tablespoonfuls, cups, glasses, or portions), each of which corresponded to a specific weight. Dietary and nutrient intakes were then calculated using nutritional analysis software (BeBIS, Nutrition Information System, Version 8). The nutritional values included any additional oral nutritional supplements used, but did not include additional parenteral feeding.

Anthropometric Measurements

Body weight was measured at the beginning and end of the study, and BMI values were calculated in both groups using the formula "weight (kg)/height² (m²)."

Nutrition Education

Nutrition education was given to 21 patients by a specialist dietician once a week for approximately 45 minutes over three months. Since it was difficult for patients to visit the hospital or other site where education could be given every week, and because the study coincided with the COVID-19 pandemic, interviews with patients and nutrition education were both carried out by phone. The aim of the education was to provide adequate and balanced nutrition, as well as to offer solutions to nutritional problems caused by the disease and the treatments applied. In the first week of nutrition education, patients and their relatives were informed about providing adequate and balanced nutrition, daily energy and nutrient requirements, the nutrients and quantities thereof they should consume daily, common nutritional problems specific to their diseases (loss of appetite, nausea, vomiting, mouth sores, dry mouth, tooth and gum problems, changes in the senses of taste and smell, early satiety, difficulty in chewing and swallowing, diarrhea, and constipation), and suggested solutions. The importance of adequate and balanced nutrition was explained and reiterated at weekly meetings, and individual nutritional problems experienced by the patients were discussed separately, with suggested solutions being given.

Assessment of quality of life

All patients, whether or not they received nutrition education, were evaluated using SF-36 at the beginning of the study and the end of the 12th week.

Collection of Patient and Control Serum Samples

Detailed histories were taken from all patients, and all underwent physical examinations and routine biochemistry tests. Briefly, 5 ml of 13 x 100 mM blood samples was taken from the peripheral vessels of all participants on the day of diagnosis of CACS and at the end of the 12th week, when the nutrition education had been completed. These samples were kept in Vacutainer® tubes for approximately 30 minutes before being stored at 4° C. Serum samples were obtained by centrifugation at 3000 rpm for 10 minutes and then stored at -80° C until biochemical analysis.

Biochemical Measurements

Routine biochemical test results and cancer markers were determined in all patients from both groups (with and without nutrition training) at the beginning and the end of the third month at the Ordu University Training and Research Hospital medical biochemistry laboratory. Human irisin, TNF- α , IL-6, PIF, and ZAG (AZGP1) serum levels were determined using ELISA kits (reference nos. DZE201125328, DZE201120083, DZE201120091, DZE201125751, and DZE201123131, Sunred Biological Technology, Shanghai, China) in line with the manufacturer's instructions using the ELISA method at the Ordu University medical biochemistry research laboratory. Absorbance was measured at 450 nm using a BioTek Instrument EL800 microplate reader (Winooski,

VT, USA). The results were expressed as ng/mL, ng/L, ng/L, μ g/mL, and ng/mL respectively.

Statistical Analysis

The test results were analyzed on SPSS 13.0.1 statistical software (license n. 9069727) (Statistical Package for the Social Sciences; SPSS Inc. Chicago, IL, USA). Data were presented as mean \pm standard deviation (SD) for normally distributed variables and as median (25%-75%) in case of non-normal distribution. The distributions of biochemical parameters, irisin, TNF- α , IL-6, PIF, and ZAG (AZGP1) levels in both groups were calculated using the Shapiro-Wilk test. Group comparisons were performed using Student's t-test in case of normal distribution and the Mann-Whitney U test in case of non-normal distribution. Comparisons of the experimental and control groups before and after the nutrition education were evaluated with the paired-sample t-test in the parametric groups and with the Wilcoxon test in the non-parametric groups. Statistical significance was determined as $p < 0.05$.

Results

The CACS patients' quality of life scale subdimensions, summary scores, and serum parameters (routine biochemical and ELISA tests) are shown in tables 1-4. There was no statistically significant difference between the experimental and control groups in terms of quality of life scale results. The groups were also evaluated within themselves in the form of pre- and post-tests conducted initially and at the end of the 12th week, and no statistically significant difference was again observed (Table 1). There was also no statistically significant difference between two groups regarding cachectic factors, inflammatory cytokines, and irisin levels (Table 2). In the experimental group, PIF and IL-6 levels increased and BMI levels decreased significantly after nutrition education (post-test) compared to the beginning(pre-test) ($p=0.027$, $p=0.010$, $p=0.014$). At the end of the training, Hgb levels increased significantly in the experimental group compared to the control group ($p=0.041$). There was no significant difference between the experimental and control groups in terms of BUN, creatinine, ALT, AST, or amylase levels. Creatinine levels decreased significantly in the experimental group after training ($p=0.015$), while LDH level increased significantly after training ($p=0.027$). Iron levels increased after training in both groups. Albumin levels increased significantly after training in the experimental group ($p=0.020$). CA 15.3 level increased significantly after training in the experimental and control groups ($p=0.011$, $p=0.008$). There was no significant change in α -1 antitrypsin, fibrinogen, hepatitis markers, or thyroid hormones (TSH, T3, T4) in the experimental or control groups before and after training (Table 3). While energy and nutrient intakes were similar between the groups at the beginning of the study, they were higher in the experimental group at the conclusion ($p < 0.05$). When the groups were evaluated within themselves, an increase in energy, protein, fat, vitamin B12, sulfur, manganese, and fluoride was observed in the experimental group after the study ($p < 0.05$), while no change was detected in the control group ($p > 0.05$) (Table 4).

Table 1. Quality of Life Scale Subdimension and Summary Scores

| Quality of Life Scale Subdimension and Summary Scores | | Experimental group (n=23) | Control group (n=18) | Test value | p-value |
|---|-----------|---|--|------------|--------------------|
| Sub Dimensions | | | | | |
| General Health Perception | Pre-test | 36.74±19.45 | 35.88±17.99 | 0.143 | 0.887 ^a |
| | Post-test | 31.44±15.97 | 33.35±15.24 | -0.383 | 0.704 ^a |
| Pain | Pre-test | 49.61±25.49 t= 1.958 p=0.063 ^c | 45.22±22.61 t= 1.457 p= 0.165 ^c | 0.574 | 0.569 ^a |
| | Post-test | 52.00±23.39 | 40.88±22.85 | 1.500 | 0.142 ^a |
| Physical Function | Pre-test | 72.39±25.17 75.00 (60.00-95.00) | 65.00±30.91 72.50 (40.00-91.25) | -0.712 | 0.476 ^b |
| | Post-test | 66.74±27.20 75.00 (45.00-90.00) Z=-1.984 | 61.17±31.74 75.00 (25.00-87.5) Z=-1.373 | -0.577 | 0.564 ^b |
| Physical Role | Pre-test | 45.65±46.25 25.00 (0.00-100.00) | 31.94±45.21 0.00 (0.00-100.00) | -0.964 | 0.335 ^b |
| | Post-test | 39.13±40.47 25.00 (0.00-75.00) Z=-0.907 | 36.76±42.49 25.00 (0.00-87.50) Z=-0.816 | -0.273 | 0.785 ^b |
| Mental Role | Pre-test | 52.17±19.66 66.67 (33.33-66.67) | 48.14±17.04 33.33 (33.33-66.67) | -0.873 | 0.383 ^b |
| | Post-test | 53.63±21.88 66.67 (33.33-66.67) Z=-0.173 | 47.05±23.74 33.33 (33.33-66.67) Z=-0.447 | -1.045 | 0.296 ^b |
| Social Function | Pre-test | 77.71±21.96 87.50 (62.50-100.00) | 63.88±29.97 68.75 (46.87-87.50) | -1.469 | 0.142 ^b |
| | Post-test | 67.39±16.74 62.50 (62.50-75.00) Z=-2.549 | 61.02±29.93 62.50 (37.50-87.50) Z=-1.179 | -0.307 | 0.759 ^b |
| Vitality | Pre-test | 48.04±20.87 45.00 (25.00-70.00) | 46.38±19.83 45.00 (32.50-65.00) | -0.330 | 0.741 ^b |
| | Post-test | 51.30±14.94 50.00 (40.00-65.00) Z=-1.577 | 42.94±16.96 45.00 (27.50-57.00) Z=-1.968 | -1.564 | 0.118 ^b |
| Mental Function | Pre-test | 56.17±17.66 64.00 (36.00-72.00) | 59.55±14.76 62.00 (52.00-69.00) | -0.370 | 0.711 ^b |
| | Post-test | 58.26±14.21 60.00 (52.00-72.00) Z=-1.000 | 57.64±13.42 60.00 (52.00-68.00) Z=-1.264 | -0.442 | 0.658 ^b |
| SUMMARY SCORES | | | | | |
| Physical Health Score | Pre-test | 39.80±11.61 | 35.62±13.32 | 1.070 | 0.291 ^a |
| | Post-test | 37.08±11.74 t=1.825 p=0.082 ^c | 34.67±13.18 t=1.009 p=0.328 ^c | 0.609 | 0.546 ^a |
| Mental Health Score | Pre -test | 37.54±11.61 | 36.12±10.07 | 0.410 | 0.684 ^a |
| | Post-test | 37.81±8.75 t=-0.205 p=0.839 ^c | 35.09±9.59 t=1.598 p=0.130 ^c | 0.932 | 0.357 ^a |

^a Independent Sample T Test, ^b Mann Whitney U, ^c Paired- Sample T Test, ^d Wilcoxon
*p<0,05

Table 2. CACS Patients' Serum Parameters and BMI Values

| Serum Parameters | | Experimental group (n=23) | Control group (n=18) | Test value | p-value |
|-------------------------|-----------|--|---|------------|--------------------|
| PIF (µg/mL) | Pre-test | 0.999±0.83 0.66 (0.34-1.70) | 1.94±2.52 1.22 (0.34-2.85) | -0.827 | 0.409 ^b |
| | Post-test | 2.00±1.43 1.80 (0.46-3.35) Z=-2.215 p=0.027* ^d | 1.83±1.25 1.52 (0.52-3.28) Z=-0.267 p=0.790 | -0.358 | 0.720 ^b |
| ZAG (ng/mL) | Pre-test | 29.47±11.82 | 26.21±12.00 | 0.793 | 0.991 ^a |
| | Post-test | 25.76±8.31 25.83 (19.38-32.62) Z=-1.198 p=0.231 ^d | 30.93±6.24 31.55 (27.2-36.0) Z=-1.412 p=0.158 | -1.859 | 0.063 ^b |
| Irisin (ng/mL) | Pre-test | 5.98±1.191 5.33 (4.65-6.85) | 6.72±3.18 6.29 (4.35-8.12) | -0.303 | 0.762 ^b |
| | Post-test | 5.88±2.31 Z=-0.469 p=0.639 ^d | 6.76±1.89 Z=-0.220 p=0.826 | -1.208 | 0.236 ^a |
| IL-6 (ng/L) | Pre-test | 48.94±14.29 | 71.03±22.1 | -3.20 | 0.089 ^a |
| | Post-test | 74.07±29.72 t=-2.935 p=0.010* ^c | 65.32±18.9 t=0.140 p=0.892 | 0.897 | 0.341 ^a |
| TNF-α (ng/L) | Pre-test | 0.72±0.34 | 0.61±0.41 | 0.797 | 0.557 ^a |
| | Post-test | 0.73±0.49 t=-0.132 p=0.896 ^c | 0.85±0.51 t=-1.342 p=0.204 | -0.703 | 0.946 ^a |
| BMI(kg/m ²) | Pre-test | 26.72±6.49 24.45 (21.22-32.04) | 26.90±6.57 24.33 (22.91-28.22) | -0.223 | 0.823 ^b |
| | Post-test | 25.43±5.59 24.33 (21.30-28.60) Z=-2.451 p=0.014* ^d | 26.19±6.66 25.15 (21.91- 27.70) Z=-2.385 p=0.017 | -0.424 | 0.671 ^b |

^a Independent Sample T Test, ^b Mann Whitney U , ^c Paired- Sample T Test, ^d Wilcoxon *p<0.05

Table 3. Routine Biochemical Parameters

| Serum Parameters | | Experimental group (n=23) | Control group (n=18) | Test value | p-value |
|--------------------|-----------|---|--|------------|---------------------|
| Hgb (g/dL) | Pre-test | 12.65±2.32 13.5 (7.9 - 16) | 11.74±1.77 11.7 (8.1 - 14.4) | 1.328 | 0.193 ^a |
| | Post-test | 12.4±1.41 12.6 (10.4 - 16.3) t= -1.443 p= 0.149 ^d | 11.3±1.47 11.5 (8.1 - 13.6) t= 0.576 p= 0.573 ^c | 109.000 | 0.041* ^b |
| BUN (mg/dL) | Pre-test | 26.75±9.93 24.15 (13.9 - 57.4) | 30.73±13.66 30.7 (9.4 - 66.09) | 163.000 | 0.341 ^b |
| | Post-test | 28.26±11.07 25.55 (8.9 - 59.2) t= -0.608 p= 0.543 ^d | 31.39±12.94 32 (6.7 - 56.2) t= -0.177 p= 0.862 | 157.000 | 0.395 ^b |
| Creatinine (mg/dL) | Pre-test | 0.86±0.13 0.86 (0.62 - 1.11) | 0.88±0.23 0.84 (0.6 - 1.56) | -0.468 | 0.642 ^a |
| | Post-test | 0.79±0.13 0.78 (0.59 - 1.1) t= 2.654 p= 0.015 ^c | 0.85±0.24 0.78 (0.56 - 1.5) t= 0.847 p= 0.409 ^c | -0.912 | 0.372 ^a |
| ALT(U/L) | Pre-test | 16.92±8.76 13 (7.6 - 36) | 14.16±8.87 10.1 (6 - 31.2) | 0.945 | 0.351 ^a |
| | Post-test | 18.39±9.57 16.5 (6.3 - 46.2) t= -1.062 p= 0.302 ^c | 15.85±8.62 12.9 (7.8 - 42.5) t= 0.398 p= 0.691 ^d | 140.500 | 0.265 ^b |
| AST(U/L) | Pre-test | 18.65±7.27 16.2 (9.6 - 39.2) | 16.64±8.3 12.7 (10 - 42.2) | 129.500 | 0.150 ^b |

| | | | | | |
|----------------------|-----------|---|---|---------|----------------------|
| Amilase(U/L) | Post-test | 23.08±8.71 22.2 (11.1 - 42.3) t= -1.654 p= 0.115 ^c | 19.81±10.54 15.25 (9.2 - 43.9) t= -0.568 p= 0.570 | 124.500 | 0.128 ^b |
| | Pre-test | 64.8±18.22 62 (26 - 96) | 61.93±24.49 56 (29 - 109) | 91.000 | 0.541 ^b |
| LDH(U/L) | Post-test | 78.17±27.58 72.5 (44 - 152) t= -1.719 p= 0.111 ^c | 70.33±27.82 64 (43 - 150) t= -1.119 p= 0.263 ^d | 107.000 | 0.311 ^b |
| | Pre-test | 178.56±46.44 169.5 (128 - 289) | 183.57±39.68 167 (140 - 280) | 95.500 | 0.492 ^b |
| Iron(µg/dL) | Post-test | 205.6±41.6 199.5 (140 - 318) t= -2.215 p= 0.027 ^d | 200.47±39.74 189 (141 - 288) t= -1.630 p= 0.134 ^c | 0.368 | 0.715 ^a |
| | Pre-test | 57.14±30.89 48 (10.3 - 129.2) | 34.28±24.88 28.75 (11.1 - 115.8) | 2.419 | 0.021 ^{*,a} |
| Total Protein (g/dL) | Post-test | 60.8±34.12 59.7 (16.2 - 127.2) t= -1.028 p= 0.318 ^c | 52.48±26.63 57.15 (18.6 - 102.9) t= -2.649 p= 0.019 ^c | 0.806 | 0.426 ^a |
| | Pre-test | 43.82±32.59 63.45 (6.5 - 75.9) | 48.16±28.81 63.1 (6.4 - 73) | -0.328 | 0.746 ^a |
| Albumine(g/dL) | Post-test | 68.26±6.38 69 (59.2 - 79) t= -1.433 p= 0.225 ^c | 65.97±5.57 66.15 (53.2 - 74.4) t= -1.243 p= 0.269 ^c | 0.877 | 0.392 ^a |
| | Pre-test | 33.17±17.47 40.4 (3.8 - 48.9) | 34.89±16.34 42.8 (3.4 - 47) | 143.000 | 0.815 ^b |
| CRP (mg/dL) | Post-test | 43.12±3.68 42.5 (36 - 50.7) t= -2.543 p= 0.020 ^c | 41.85±4.58 42.5 (29.4 - 49.1) t= -1.083 p= 0.279 ^d | 0.935 | 0.356 ^a |
| | Pre-test | 5.58±6.74 2.67 (0.13 - 23.69) | 4.76±4.33 3.4 (0.51 - 16.41) | 163.000 | 0.831 ^b |
| Anti-HCV (IU/mL) | Post-test | 5.08±4.75 3.69 (0.62 - 17.23) t= 0.113 p= 0.911 ^c | 3.32±4.33 1.7 (0.6 - 16.28) t= -1.193 p= 0.233 ^d | 112.000 | 0.058 ^b |
| | Pre-test | 0.05±0.04 0.04 (0.02 - 0.23) | 0.03±0.01 0.03 (0.02 - 0.05) | 141.500 | 0.277 ^b |
| HBsAg (mIU/mL) | Post-test | 0.05±0.05 0.04 (0.03 - 0.26) t= -0.922 p= 0.357 ^d | 0.04±0.01 0.04 (0.02 - 0.06) t= -1.100 p= 0.271 ^d | 0.805 | 0.426 ^a |
| | Pre-test | 0.59±0.13 0.61 (0.27 - 0.79) | 0.51±0.18 0.53 (0.2 - 0.84) | 110.500 | 0.066 ^b |
| Anti-HBs (mIU/mL) | Post-test | 0.61±0.12 0.62 (0.33 - 0.85) t= -0.735 p= 0.472 ^c | 0.57±0.15 0.61 (0.22 - 0.77) t= 0.738 p= 0.460 ^d | 1.031 | 0.310 ^a |
| | Pre-test | 31.83±67.27 1.7 (1.3 - 240.6) | 13.6±27.28 2.06 (1.4 - 106.6) | 138.000 | 0.482 ^b |
| Anti-HIV | Post-test | 42.96±92.83 1.7 (1 - 342.2) t= -0.831 p= 0.406 ^d | 43.62±110.81 1.75 (1.4 - 389) t= -1.482 p= 0.138 ^d | 105.500 | 0.442 ^b |
| | Pre-test | 0.2±0.03 0.21 (0.14 - 0.25) | 0.19±0.03 0.18 (0.15 - 0.25) | 0.594 | 0.556 ^a |
| CA-125(U/mL) | Post-test | 0.2±0.04 0.19 (0.11 - 0.3) t= 0.454 p= 0.655 | 0.19±0.04 0.18 (0.14 - 0.27) t= 0.526 p= 0.607 ^c | 0.634 | 0.530 ^a |
| | Pre-test | 13.11±7.43 11.2 (4.89 - 30.2) | 12.18±4.42 11.9 (6.69 - 24.1) | 123.000 | 0.865 ^b |

| | | | | | |
|-----------------------------|-----------|--|---|---------|---------------------|
| CA15-3 (U/mL) | Post-test | 16.64±9.9 14.3 (6.17 - 41.8) t= -2.685 p= 0.017 ^c | 13.18±7.66 10.7 (7.46 - 34.2) t= -0.157 p= 0.875 ^d | 1.130 | 0.266 ^a |
| | Pre-test | 16.17±8.16 14.85 (4.85 - 40.5) | 17.57±7.19 16 (5.81 - 35) | 160.000 | 0.444 ^b |
| CA 19-9 (U/mL) | Post-test | 20.98±9.77 18.9 (7.42 - 44.6) t= -2.555 p= 0.011 ^d | 23.19±9.94 23.75 (7.88 - 42.9) t= -3.189 p= 0.008 ^c | -0.656 | 0.516 ^a |
| | Pre-test | 11.39±10.5 7.13 (2 - 43.5) | 13.88±14.22 10.3 (2 - 56.86) | 131.500 | 0.536 ^b |
| CEA (U/mL) | Post-test | 10.03±7.43 7.14 (2 - 29) t= 0.555 p= 0.586 ^c | 14.87±9.91 13.15 (2 - 33.7) t= -0.711 p= 0.477 ^d | -1.721 | 0.094 ^a |
| | Pre-test | 3.91±4.51 2.17 (0.83 - 16.4) | 6.26±5.12 5.52 (0.59 - 18.3) | 124.000 | 0.110 ^b |
| Folate (ng/mL) | Post-test | 2.99±2.22 2.39 (0.91 - 8.79) t= -0.713 p= 0.476 ^d | 2.71±2.08 2.26 (1.36 - 9.58) t= -2.417 p= 0.016 ^d | 169.000 | 0.917 ^b |
| | Pre-test | 9.06±3.72 8.4 (4.08 - 18.5) | 7.16±3.98 6.47 (3.13 - 19.7) | 63.500 | 0.044 ^{*b} |
| Ferritin (ng/mL) | Post-test | 9.57±5.43 6.73 (3.48 - 20) t= -0.284 p= 0.776 ^d | 9.71±5.94 7.97 (3.53 - 20) t= -1.020 p= 0.308 ^d | 162.000 | 0.854 ^b |
| | Pre-test | 79.77±48.45 66.95 (24.3 - 169) | 31.6±22.86 25.8 (10.3 - 71.8) | 18.000 | 0.006 ^{*b} |
| TSH(mIU/L) | Post-test | 129.58±129.75 73.9 (10.3 - 472) t= -1.423 p= 0.155 ^d | 96.47±125.7 40.4 (8.45 - 437) t= -0.700 p= 0.484 ^d | 116.000 | 0.183 ^b |
| | Pre-test | 1.46±1.09 1.13 (0.16 - 4.67) | 2.49±2.08 1.57 (0.3 - 7.57) | -1.868 | 0.075 ^a |
| Free T4 (ng/dL) | Post-test | 1.7±1.22 1.45 (0.09 - 4.86) t= -2.183 p= 0.041 ^c | 1.33±0.98 1.18 (0.08 - 3.21) t= 1.510 p= 0.157 ^c | 0.968 | 0.340 ^a |
| | Pre-test | 1.24±0.21 1.19 (0.86 - 1.72) | 1.22±0.16 1.21 (0.99 - 1.58) | 185.000 | 0.955 ^b |
| Free T3(ng/dL) | Post-test | 1.24±0.25 1.26 (0.79 - 1.77) t= -0.260 p= 0.795 ^d | 1.19±0.17 1.14 (0.96 - 1.51) t= -0.408 p= 0.683 ^d | 0.709 | 0.483 ^a |
| | Pre-test | 2.83±0.47 2.92 (2.04 - 3.74) | 2.66±0.48 2.58 (1.98 - 3.36) | 93.000 | 0.337 ^b |
| Alpha-1 Antitrypsin (mg/dL) | Post-test | 2.68±0.68 2.65 (1.45 - 3.49) t= -0.051 p= 0.959 ^d | 2.53±0.34 2.45 (2.19 - 2.98) t= -0.716 p= 0.506 ^c | 20.000 | 0.191 ^b |
| | Pre-test | 174.27±36.3 165 (122 - 238) | 175.85±46.04 173 (118 - 279) | 70.000 | 0.931 ^b |
| Fibrinogen (mg/dL) | Post-test | 164.15±33.4 158 (91.1 - 253) t= -0.357 p= 0.721 ^d | 170.27±41.08 171 (107 - 264) t= -0.663 p= 0.508 ^d | -0.492 | 0.626 ^b |
| | Pre-test | 380.09±64.81 379.5 (280 - 479) | 405.4±50.77 418 (328 - 512) | 121.500 | 0.178 ^b |
| Fibrinogen (mg/dL) | Post-test | 386.35±71.46 380.5 (246 - 523) t= -0.411 p= 0.681 ^d | 380.6±99.52 350 (244 - 622) t= -1.503 p= 0.133 ^d | 0.199 | 0.843 ^b |

^a Independent Sample T Test, ^b Mann Whitney U, ^c Paired- Sample T Test, ^d Wilcoxon
*p<0.05

Table 4. Energy and Nutrients

| Energy and Nutrients | | Experimental group (n=23) | Control group (n=18) | Test value | p-value |
|----------------------|-----------|--|---|------------|---------------------|
| Energy ((kcal) | Pre-test | 1211.37±550.23 1216.71 (375.08 - 2440.34) | 1314.28±522.21 1305.8 (595.94 - 2512.07) | -0.608 | 0.547 ^a |
| | Post-test | 1526.93±422.11 1446.83 (701.85 - 2316.6) t= -2.466 p= 0.022* ^c | 1071.07±477.35 1038.24 (539.35 - 2257.09) t= 1.918 p= 0.072 ^c | 3.240 | 0.002* ^a |
| Protein (g) | Pre-test | 48.83±25.65 50.21 (13.84 - 98.11) | 52.39±20.66 49.69 (21.89 - 85.02) | -0.478 | 0.635 ^a |
| | Post-test | 67.11±26.76 67.53 (24.53 - 135.7) t= -2.396 p= 0.026* ^c | 47.44±24.94 44.58 (18.55 - 104.86) t= 0.773 p= 0.450 ^c | 2.406 | 0.021* ^a |
| Fat (g) | Pre-test | 47.89±27.22 44.99 (11.9 - 110.62) | 57.3±32.93 47.45 (17.27 - 155.71) | -1.003 | 0.322 ^a |
| | Post-test | 64.03±24.69 60.6 (26 - 128.83) t= -2.414 p= 0.025* ^c | 44.82±23.79 38.6 (17.38 - 99.39) t= 1.651 p= 0.117 ^c | 2.512 | 0.016* ^a |
| Carbohydrate (g) | Pre-test | 141.15±63.71 126.68 (37.61 - 263.38) | 143.34±51.15 157.87 (50.03 - 213.33) | -0.119 | 0.906 ^a |
| | Post-test | 165.68±53.54 157.24 (83.32 - 323.34) t= -1.517 p= 0.144 ^c | 116.32±51.52 103 (51.17 - 246.74) t= 1.867 p= 0.079 ^c | 2.978 | 0.005* ^a |
| Vitamin A(µg) | Pre-test | 898.89±1118.2 575.88 (216.16 - 5602.51) | 859.13±475.4 960.24 (151.6 - 1814.12) | -0.972 | 0.331 ^b |
| | Post-test | 1186.21±1232.71 645.6 (293.68 - 4949.65) t= -0.152 p= 0.879 ^d | 650.83±901.7 372.32 (201.85 - 4147.52) t= -1.851 p= 0.248 ^d | -2.338 | 0.019* ^b |
| Vitamin E (mg) | Pre-test | 7.42±7.2 6.45 (1.02 - 30.3) | 9.85±8.05 8.53 (2.04 - 33.26) | -1.314 | 0.189 ^b |
| | Post-test | 10.21±7.04 6.95 (2.35 - 26.21) Z= -0.821 p= 0.411 ^d | 6.07±5.61 4.12 (1.91 - 24.78) Z= -1.764 p= 0.078 ^d | -2.509 | 0.012* ^b |
| Thiamine (mg) | Pre-test | 0.62±0.32 0.65 (0.17 - 1.34) | 0.64±0.28 0.66 (0.2 - 1.11) | -0.221 | 0.826 ^a |
| | Post-test | 0.8±0.33 0.84 (0.26 - 1.49) t= -1.901 p= 0.070 ^c | 0.52±0.26 0.5 (0.21 - 1.01) t= 1.681 p= 0.111 ^c | 2.951 | 0.005* ^a |
| Riboflavin (mg) | Pre-test | 1.04±0.62 1.07 (0.21 - 2.14) | 1.18±0.49 1.16 (0.47 - 1.89) | -0.786 | 0.436 ^a |
| | Post-test | 1.29±0.44 1.21 (0.44 - 2.12) t= -1.589 p= 0.126 ^c | 0.98±0.46 0.98 (0.39 - 1.94) t= 1.500 p= 0.152 ^c | 2.174 | 0.036* ^a |
| Niacin (mg) | Pre-test | 8.41±8.09 4.83 (1.84 - 38.56) | 7.83±4.15 7.03 (1.61 - 20.16) | -0.709 | 0.478 ^b |
| | Post-test | 11.66±6.06 10.95 (1.71 - 25.76) Z= -1.734 p= 0.083 ^d | 7.62±5.88 6.54 (1.31 - 25.7) Z= -0.936 p= 0.349 ^d | -2.391 | 0.017* ^b |
| Biotin (µg) | Pre-test | 35.37±21.86 35.14 (4.96 - 81.52) | 41.54±18.31 39.97 (17.04 - 78.78) | -0.961 | 0.343 ^a |
| | Post-test | 43.22±13.07 42.51 (23.99 - 68.6) t= -1.279 p= 0.214 ^c | 30.99±12.2 26.78 (16.26 - 56.89) t= 2.081 p= 0.053 ^c | 3.061 | 0.004* ^a |
| Vitamine B12 (µg) | Pre-test | 3.04±2.09 2.59 (0 - 6.67) | 4.25±2.3 3.65 (1.42 - 8.82) | -1.759 | 0.086 ^a |
| | Post-test | 5.05±3.96 | 2.94±2.07 | 2.200 | 0.035* ^a |

| | | | | | |
|----------------|-----------|--|---|--------|---------------------|
| Potassium (mg) | Pre-test | 4.46 (1.03 - 18.03) t= -2.140 p= 0.044* ^c 2033.7±1047.77 1966.25 (779.38 - 4402.53) | 2.44 (0.85 - 8.84) t= 1.760 p= 0.096 ^c 2127.57±882.97 2194.88 (543.74 - 3472.35) | -0.604 | 0.546 ^b |
| | Post-test | 2526.22±909.75 2672.79 (971.96 - 4394.8) Z= -1.369 p= 0.171 ^d | 1686.91±872.92 1516.26 (653.43 - 4128.85) Z= -1.459 p= 0.145 ^d | 2.838 | 0.005* ^b |
| Magnesium (mg) | Pre-test | 193.42±94.45 203.69 (59.32 - 408.7) | 206.04±93.46 216.58 (70.44 - 334.82) | -0.426 | 0.672 ^a |
| | Post-test | 243.8±84.51 236.7 (95.54 - 398.35) t= -2.038 p= 0.054 ^c | 167.84±80.6 158.85 (72.89 - 342.7) t= 1.608 p= 0.126 ^c | 2.914 | 0.006* ^a |
| Sulfur | Pre-test | 279.59±167.32 270.79 (40.8 - 722.54) | 226.11±178.66 168.55 (34.06 - 677.9) | 0.986 | 0.330 ^a |
| | Post-test | 758.77±294.45 740.35 (287.81 - 1571.76) t= -2.103 p= 0.047* ^c | 564.5±300.23 515.29 (241.51 - 1288.61) t= 0.531 p= 0.602 ^c | 2.079 | 0.044* ^a |
| Iron (mg) | Pre-test | 6.88±3.58 5.89 (1.99 - 14.65) | 7.08±3.42 5.62 (2.99 - 11.57) | -0.173 | 0.863 ^a |
| | Post-test | 8.81±3.8 8.08 (2.38 - 16.34) t= -1.927 p= 0.067 ^c | 5.72±3.2 5.52 (2.39 - 14.5) t= 1.563 p= 0.136 ^c | 2.761 | 0.009* ^a |
| Copper (mg) | Pre-test | 1.05±0.62 0.96 (0.31 - 2.41) | 1.06±0.6 0.87 (0.29 - 2.24) | -0.171 | 0.864 ^b |
| | Post-test | 1.55±0.98 1.22 (0.36 - 4.2) Z=-1.867 p= 0.062 ^d | 1.07±0.87 0.72 (0.28 - 3.15) t= -0.008 p= 0.994 ^c | -2.286 | 0.022* ^b |
| Manganese (mg) | Pre-test | 2.32±1.45 2.08 (0.57 - 6.31) | 2.61±1.78 2.18 (0.74 - 6.45) | -0.079 | 0.937 ^b |
| | Post-test | 3.55±1.85 3.45 (1.05 - 6.92) Z=-2.768 p= 0.006* ^d | 2.06±1.6 1.53 (0.72 - 5.85) t= 1.217 p= 0.240 ^c | -2.786 | 0.005* ^b |
| Fluorine (µg) | Pre-test | 424.11±270.7 354.7 (130.48 - 1255.17) | 416.32±227.35 344.14 (119.56 - 865.67) | -0.079 | 0.937 ^b |
| | Post-test | 556.5±207.44 569 (244.64 - 935.33) Z=-2.190 p= 0.029* ^d | 430.76±258.78 318.26 (160.7 - 1018.45) t= -0.199 p= 0.845 ^c | -2.102 | 0.036* ^b |
| Omega-3(g) | Pre-test | 1.13±0.67 0.88 (0.42 - 2.88) | 1.23±0.71 0.92 (0.41 - 3.1) | -0.486 | 0.630 ^a |
| | Post-test | 1.26±0.44 1.22 (0.46 - 2.42) t= -0.815 p= 0.424 ^c | 0.94±0.35 0.85 (0.54 - 1.69) t= 1.691 p= 0.109 ^c | 2.541 | 0.015* ^a |
| Omega-6(g) | Pre-test | 4.92±4.19 3.83 (1.12 - 18.37) | 6.69±5.38 4.91 (1.13 - 19.81) | -1.196 | 0.232 ^b |
| | Post-test | 6.23±3.79 4.81 (2.02 - 15.12) Z=-1.247 p= 0.212 ^d | 4.44±4.07 2.5 (1.38 - 16.05) Z=-1.851 p= 0.064 ^d | -2.233 | 0.026* ^b |

^a Independent Sample T Test, ^b Mann Whitney U , ^c Paired- Sample T Test, ^d Wilcoxon
*p<0.05

Discussion

Cachexia is the second most important cause, after sepsis, of cancer-related deaths. The condition progresses with weight loss in cancer patients and cannot be reversed using normal nutritional support. Research in recent years has shown that cancer cachexia is associated with

anorexia, a condition known as CACS [10]. The etiology of CACS involves numerous factors. Its most important feature is the deterioration of fat tissue along with muscle loss. Cancer patients experience loss of appetite in addition to involuntary weight loss, and this reduces their

quality of life. In its advanced stages, CACS cannot be completely treated by eating more or using nutritional supplements. Increasing weight loss over time causes patients with this syndrome to experience difficulties in meeting their daily needs. The occurrence of the syndrome depends on the type of cancer involved. Weight loss is observed before starting treatment in 50-85% of patients with gastrointestinal, pancreatic, lung, and colorectal cancers. Cancer anorexia-cachexia is responsible for 20% of all cancer-related deaths, and is most commonly seen in gastrointestinal and lung malignancies [11]. When a tumor is present, numerous metabolic factors are produced by the host and the tumor, and these play a critical role in tissue (fat and skeletal muscle) wastage in cancer cachexia. Although the mechanism has not been fully elucidated, lipid mobilizing factor (LMF)/Zinc- α -2-glycoprotein (ZAG), TNF- α , IL-1, and IL-6, which are excessively synthesized by the body, are released from tumor cells in cancer cachexia, while factors that cause cancer cachexia continue to be investigated [12]. PIF is secreted by the tumor and causes a severe decrease in muscle mass by increasing protein catabolism, primarily by activating the ATP-dependent proteolytic system. PIF has particularly been detected in the urine of patients experiencing body weight loss. LMF increases lipolysis in adipose tissue, and both are secreted by the tumor [13,14]. A previous study reported elevated serum TNF- α levels in 36.5% of 63 patients with pancreatic cancer, in patients with metastatic pancreatic tumors, and others [15]. TNF- α , IL-1, IL-6, and IFN- γ levels were not found to be associated with weight loss in advanced and terminal cancer patients in another study [16]. Pro-inflammatory mediators are released by the tumor microenvironment and systemically. This leads to reduced muscle protein synthesis by downregulating the mammalian target of rapamycin (mTOR) and increased muscle degradation by upregulating atrogen-1 and muscle RING-finger protein-1 (MuRF-1) [17]. The connection between irisin, known as weight loss myokine, and "cancer and weight loss" (irisin and cancer, irisin and weight loss) has to date been reported in immunological and biochemical studies. Strong irisin immunoreactivity was observed in cancerous tissues of the gastrointestinal system (the pancreas, liver, spleen, and stomach) in one study [18]. Other research examined the tumor tissues of mice with gastric cancer, and no expression of irisin precursor FNDC5 was found in stomach tumor tissue. However, it was suggested that some factors released from the tumor tissue activate numerous unknown signaling pathways, first stimulating the expression of FNDC5 in white and brown adipose tissue, thus increasing the level of irisin in the circulation. Irisin then causes weight loss by increasing the degradation of triacylglycerol in white adipose tissue [19]. An in vitro study conducted with lung, breast, prostate, osteosarcoma, and pancreatic cancer cell lines showed that irisin applied at a specific dose and duration exhibited an inhibitory effect. However, the application of irisin did not affect endometrial, colon, thyroid, or esophageal cell

lines [20]. Another in vivo study suggested that serum irisin may represent a potential diagnostic marker for breast, renal, colon, and rectal cancers [21].

No statistically significant differences were observed between the experimental and control groups in this study in terms of cachectic factors, inflammatory cytokines, or irisin levels (Table 2). However, in the experimental group, PIF and IL-6 levels increased and BMI levels decreased significantly following nutrition education (post-test) compared to baseline (pre-test) ($p=0.027$, $p=0.010$, and $p=0.014$, respectively). Interestingly, nutrition education given to the experimental group may have raised those patients' stress levels, resulting in a decrease in BMI and increased inflammation (IL-6 and TNF- α).

Quality of life is one of the indicators of health. In addition to traditional mortality and morbidity measures, it is also an important factor that should be taken into consideration when determining individuals' health levels. Since the creation and development of quality of life scales is a labor-intensive process, the use of scales developed by others and whose validity and reliability have already been tested is becoming increasingly widespread. SF-36 was developed and has been made available by the Rand Corporation for the evaluation of quality of life [22,23]. It was translated into Turkish and validated by Koçyiğit et al. [24]. The scale is short, easy to apply, and highly versatile [25]. The Turkish version of SF-36 can be used in cancer research in Türkiye [8]. The scale subdimensions and summary scores are shown in Table 1. There was no statistically significant difference between the experimental and control groups in terms of SF-36. The groups were also evaluated within themselves in the form of pre- and post-tests initially and at the end of the 12th week, and no statistically significant difference was again observed. Studies have reported that nutrition education increases patients' quality of life [26-32]. In a study examining the effect of education on the quality of life of breast cancer patients, a comprehensive education program, including nutrition education, was applied to an experimental group for 12 weeks. At the end of the study period, a significant increase was observed in the quality of life of the patients who received education [26]. Lee et al. observed that three-week nutrition education given to postgastroectomy patients increased their quality of life and improved their nutritional status [27]. Another study examining the effect of nutrition education on eating habits and quality of life in stomach cancer patients, significant improvements were determined in serum albumin and hematocrit levels after education, together with a significant increase in the amount of food consumed and quality of life [28]. Research involving other patients (with kidney, liver, and lung diseases) has also reported that nutrition education improves quality of life [29-32]. We attribute the fact that nutrition education in the present study did not produce significant changes in the quality of life of cancer patients to the data collection process coinciding with the Covid-19 pandemic, cancer patients losing hope as a result of the pandemic,

and physical activity being limited by the curfew. All these factors had an adverse psychological effect on cancer patients. Additionally, the nutrition education given to the experimental group may have caused an increase in patients' concerns about the disease and therefore a decrease in BMI. Studies have reported that nutrition education exhibits positive effects on biochemical parameters [28,33-38]. Studies examining the effectiveness of nutrition education in cancer patients are scarce. However, the effects of nutrition education on biochemical parameters in other diseases have been investigated. Both a desired change in dietary intake and BMI and improvements in biochemical parameters were observed in a previous study involving patients with metabolic syndrome [33]. Another study found that nutrition education given to obese and hypertensive patients resulted in a decrease in BMI and an improvement in biochemical parameters [34]. Studies conducted with other patient groups have similarly reported positive effects of nutrition education on BMI and biochemical parameters [35-38].

Chao et al. reported a significant increase in nutritional intake following nutrition education in their study of 444 oncology patients [39]. Nutrition education (information about both good management of the treatment process and nutrition) given to gastric cancer patients receiving chemotherapy has been shown to result in a significant increase in food intake [40]. Similarly to the results of other studies, nutrition education in the present study produced a desired change in nutrient intake. Nutrition education given to malnourished patients with cancer can produce the desired change in nutritional intake.

At the end of the training, Hgb levels increased significantly in the experimental group compared to the control group ($p=0.041$). However, no significant differences were observed between the two groups in terms of BUN, creatinine, ALT, AST, or amylase levels. Creatinine levels decreased significantly in the experimental group after training, while LDH levels increased significantly ($p=0.015$ and $p=0.027$, respectively). Iron levels increased after training in both groups. Albumin levels increased significantly after training in the experimental group ($p=0.020$). CA 15.3 levels increased significantly after training in the experimental and control groups ($p=0.011$ and $p=0.008$, respectively). There was no significant change in α -1 antitrypsin, fibrinogen, hepatitis markers, or thyroid hormones (TSH, T3, T4) in either group before and after training.

Limitations

The COVID-19 epidemic throughout the study period caused our project duration to be extended, and additional time was requested for completion (12 months). The majority of patients with CACS died due to cachexia before completing the nutrition education period. Our project commenced with CACS patients

diagnosed with stomach cancer, but due to the prolonged nature of the pandemic and our inability to reach the desired number of samples, patients with colon and rectal cancers were also included. Differences in the stages and types of cancer may have affected the results. The effects of chemotherapeutic drugs taken for 12 weeks on biochemical parameters, quality of life, and nutrition should not be ignored. Patients with the same type of cancer, with similar stages, and using similar drugs for treatment should be selected when planning further studies. Once the nutrition education has been completed, in order to observe the long-term results, blood specimens should be collected at the end of the third, sixth, and ninth months

Conflict of Interest

The authors declare that they have no conflicting interests.

Acknowledgments

The authors are indebted to the study participants, to the hospital staff, and to the laboratory staff for their assistance in performing the investigations.

Ethical Approval

This study was reviewed and approved by the Institutional Review Board at Ordu University, Türkiye, and was performed in agreement with the principles of the Declaration of Helsinki. Ethics committee approval was received from the Ordu University local ethical committee under reference no. 2021/184.

References

- [1] Anand U., Dey A., Chandel A.K.S., Sanyal R., Mishra A., Pandey D.K., De Falco V., Upadhyay A., Kandimalla R., Chaudhary A., Dhanjal J.K., Dewanjee S., Vallamkondu J., Pérez de la Lastra J.M., Cancer chemotherapy and Beyond Current status, drug candidates, associated risks, and progress in targeted therapeutics, *Genes Dis.*, 18(1) (2023) 1367-1401.
- [2] Aoyagi T., Terracina K.P., Raza A., Matsubara H., Takabe K., Cancer cachexia, mechanism, and treatment, *World J. Gastrointest Oncol.*, 7(4) 17-29.
- [3] Law M.L., Cancer Cachexia: Pathophysiology and Association with Cancer-Related Pain, *Front. Pain Res. (Lausanne Switz)*, 3 (2022) 971295
- [4] Siddiqui J.A., Pothuraju R., Jain M., Batra S.K., Nasser M.W., Advances in cancer cachexia: Intersection between affected organs, mediators, and pharmacological interventions, *Biochim Biophys Acta Rev Cancer.*, 1873(2) (2020) 188359.
- [5] Muliawati Y., Haroen H., Rotty L.W., Cancer anorexia - cachexia syndrome, *Acta Med Indones.*, 44(2) (2012) 154-62.
- [6] Us Altay D., Onder S., Etgu F., Uner A., Noyan T., A newly identified myokine: irisin, and its relationship with chronic spontaneous urticaria and inflammation, *Arch Dermatol Res.*, 315(2023) 437-442.

- [7] Arhire L.I., Mihalache L., Covasa M., Irisin: A Hope in Understanding and Managing Obesity and Metabolic Syndrome, *Front Endocrinol (Lausanne)*, 2(10) (2019) 524.
- [8] Pinar R., Reliability and construct validity of the SF-36 in Turkish cancer patients, *Qual Life Res.*, 14(1) (2005). 259-64.
- [9] Kondrup J., Rasmussen H.H., Hamberg O., Stanga Z., Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials, *Clinical Nutrition*, 22 (3) (2003) 321-336.
- [10] Ezeoke C.C., Morley J.E., Pathophysiology of anorexia in the cancer cachexia syndrome, *J Cachexia Sarcopenia Muscle*, 6(4) (2015) 287-302.
- [11] Ni J., Zhang L., Cancer Cachexia: Definition, Staging, and Emerging Treatments, *Cancer Manag Res.*, 9 (2020) 12 5597-5605.
- [12] Işık G., Demirezen S., Beksac M.S., Tümör Nekroz Faktör ve Servikal Kanser Bağlantısı, *Türk Bilimsel Derlemeler Dergisi*, 1 (2) (2008) 55-61.
- [13] Lecker S.H., Solomon V., Mitch W.E., Goldberg A.L., Muscle protein breakdown and the critical role of the ubiquitin-proteasome pathway in normal and disease states, *J Nutr.*, 129 (1999) 227-237.
- [14] Todorov P.T., McDevitt T.M., Meyer D.J., Purification and characterization of a tumor lipid mobilizing factor, *Cancer Res.*, 58 (1998) 235.
- [15] Karayiannakis A.J., Syrigos K.N., Polychronidis A., Pitakoudis, M, Bounovas A, Simppoulos K., Serum levels of tumor necrosis factor- α and nutritional status in pancreatic cancer patients, *Anti Cancer Res.*, 21 (2001) 1355-1358.
- [16] Maltoni M., Fabbri L., Nanni O., Scarpi E., Pezzi L., Flamini E., Rittobon A., Derni S., Pallotti G., Amadori D., Serum levels of tumor necrosis factor and other cytokines do not correlate with weight loss and anorexia in cancer patients, *Support Care Cancer*, 5 (1997) 130-135.
- [17] Vendrell I., Macedo D., Alho I., Dionísio M.R., Costa L., Treatment of Cancer Pain by Targeting Cytokines, *Mediat. Inflamm.*, (2015) 984570.
- [18] Aydın S., Kuloglu T., Ozercan M.R., Albayrak S., Aydın S., Bakal U., Irisin İmmüno histochemistry In Gastrointestinal System Cancers, *Biotech Histochem.*, 91 (2015) 242-250.
- [19] Us Altay D., Keha E.E., Ozer Yaman S., Ince I., Alver A., Erdogan B., Canpolat S., Cobanoglu U., Mentese A., Investigation of the expression of irisin and some cachectic factors in mice with experimentally induced gastric cancer, *QJM: An International Journal of Medicine*, 109 (12) (2016) 785-790.
- [20] Moon H.S., Mantzoros C.S., Regulation of cell proliferation and malignant potential by irisin in endometrial, colon, thyroid, and esophageal cancer cell lines, *Metabolism*, 63(2) (2014) 188-93.
- [21] Maalouf G.E., Khoury D., Exercise-induced irisin, the fat browning myokine, as a potential anticancer agent, *Journal of Obesity*, 1 (2019) 6561726.
- [22] Ware J.E., Sherbourne C.D., The MOS 36-item short-form health survey (SF-36) 1: conceptual framework and item selection, *Medical Care*, 30 (1992) 473-483.
- [23] Taft C., Karlsson J., Sullivan M., Do SF-36 summary component scores accurately summarize scores?, *Quality of Life Research*, 10 (2001) 395-404.
- [24] Koçyiğit H., Aydemir Ö., Fisek G., Ölmez N., Memiş A., Kısa form-36 (KF-36)'nın Türkçe versiyonunun güvenilirliği ve geçerliliği, *İlaç ve Tedavi Dergisi*, 12 (1999) 102-106.
- [25] Bozdemir H., Karaciğer Transplantasyonu Uygulanan Hastalarda Yaşam Kalitesinin İncelenmesi, Yüksek lisans tezi, İzmir Dokuz Eylül Üniversitesi, Sağlık Bilimleri Enstitüsü, 2006.
- [26] Çınar D., Meme kanserli kadınlarda e-mobil eğitimin yaşam kalitesine etkisi, Doktora tezi, İzmir Ege Üniversitesi, Sağlık Bilimleri Enstitüsü, 2019.
- [27] Lee H.O., Han S.R., Choi S.I., Lee J.J., Kim S.H., Ahn H.S., Lim H., Effects of intensive nutrition education on nutritional status and quality of life among postgastrectomy patients, *Annals of Surgical Treatment and Research*, 90(2) (2016) 79-88.
- [28] Jung Y., Lee J., Effect of Nutrition Education on the Eating Habits and Quality of Life of Gastric Cancer Outpatients Undergoing Gastrectomy, *Korean Journal of Community Nutrition*, 23(2) (2018) 162-173.
- [29] Ebrahimi H., Sadeghi M., Amanpour F., Dadgari A., Influence of nutritional education on hemodialysis patients' knowledge and quality of life, *Saudi Journal of Kidney Diseases and Transplantation*, 27(2) (2016) 250-255.
- [30] Brug J., Schols A., Mesters I., Dietary change, nutrition education, and chronic obstructive pulmonary disease, *Patient Education and Counseling*, 52(3), 249-257.
- [31] Anderson C.A., Nguyen H.A., Nutrition education in the care of patients with chronic kidney disease and end-stage renal disease, *In Seminars in Dialysis*, 31(2) (2018) 115-121
- [32] Alavinejad P., Hajiani E., Danyae B., Morvaridi M., The effect of nutritional education and continuous monitoring on clinical symptoms, knowledge, and quality of life in patients with cirrhosis, *Gastroenterology and hepatology from bed to bench*, 12 (1) (2019) 17-24.
- [33] Kim J., Bea W., Lee K., Han J., Kim S., Kim M., Sohn C Effect of the telephone-delivered nutrition education on dietary intake and biochemical parameters in subjects with metabolic syndrome, *Clinical nutrition research*, 2(2) (2013) 115-124.
- [34] Gajewska D., Kucharska A., Kozak M., Wunderlich S., Niegowska J., Effectiveness of individual nutrition education compared to group education, in improving anthropometric and biochemical indices among hypertensive adults with excessive body weight: a randomized controlled trial, *Nutrients*, 11(12), (2019), 2921.
- [35] Pivi G.A., da Silva R.V., Juliano Y., Novo N.F., Okamoto I.H., Brant C.Q., Bertolucci P.H., A prospective study of nutrition education and oral nutritional supplementation in patients with Alzheimer's disease, *Nutrition Journal*, 10(1) (2011) 1-6.
- [36] Luisi M.L.E., Biffi B., Gheri C.F., Sarli E., Rafanelli E., Graziano E., Macchi C., Efficacy of a nutritional education program to improve diet in patients attending a cardiac rehabilitation program: outcomes of a one-year follow-up, *Internal and Emergency Medicine*, 10 (2015) 671-676.
- [37] Kim H.J., Hong J.I., Mok H.J., Lee K.M., Effect of workplace-visiting nutrition education on anthropometric and clinical measures in male workers, *Clinical Nutrition Research*, 1(1) (2012) 49-57.
- [38] Kim S.Y., Kim S.B., Effects of Nutrition Education at a Community Health Center on Overweight and Obese Middle-aged Women in Jeonbuk Area-Focused on Personalized Daily Energy Requirement and Food Exchange Units, *Korean Journal of Community Nutrition*, 22(4) (2017) 307-322.
- [39] Chao P.C., Chuang H.J., Tsao L.Y., Chen P.Y., Hsu C.F., Lin H.C., Lin C.F., The Malnutrition Universal Screening Tool (MUST) and a nutrition education program for high-risk cancer patients: strategies to improve dietary intake in cancer patients, *Biomedicine*, 5 (2015) 1-6
- [40] Xie F.L., Wang Y.Q., Peng L.F., Lin F.Y., He Y.L., Jiang Z.Q., The beneficial effect of educational and nutritional intervention on the nutritional status and compliance of gastric cancer patients undergoing chemotherapy: a randomized trial, *Nutrition and Cancer*, 69(5) (2017) 762-771.

A Bibliometric and Visual Analysis of Publications on Low-Density Lipoprotein Cholesterol Estimating Equations

Serkan Bolat ^{1,a,*} Serra İlayda Yerlitaş ^{2,3,b}, Ahu Cephe ^{4,c}, Necla Koçhan ^{5,d}, Gözde Ertürk Zararsız ^{2,3,e}, Halef Okan Doğan ^{1,f}, Gökmen Zararsız ^{1,2,6,g}

¹Department of Medical Biochemistry, School of Medicine, Sivas Cumhuriyet University, Sivas, Türkiye

²Department of Biostatistics, Erciyes University School of Medicine, 38039 Kayseri, Türkiye

³Drug Application and Research Center (ERFARMA), Erciyes University, 38280 Kayseri, Türkiye

⁴Institutional Data Management and Analytics Unit, Erciyes University Rectorate, Kayseri, Türkiye

⁵Department of Mathematics, Izmir University of Economics, 35330, Izmir, Türkiye

⁶Erciyes Teknopark, Hematiner Biotechnology and Health Products Inc., 38010, Kayseri, Türkiye

*Corresponding author

Research Article

History

Received: 15/03/2024

Accepted: 16/11/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

The concentration of low-density lipoprotein cholesterol (LDL-C) is recognized as a crucial factor in cardiovascular health. This study aims to conduct a comprehensive bibliometric and visual analysis to provide a comprehensive review of current research trends and patterns in this research area. We retrieved the publications from the Web of Science (WoS) database and conducted the bibliometric analyses using VOSviewer software, bibliometrix R package, and biblioshiny web tool. The analysis was conducted on 620 original articles and review papers published between 1990 and 2023 from institutions located in 62 different countries and published in 329 journals. The countries with the most articles were the US, China, and Japan. The most productive journals were Clinica Chimica Acta, Clinical Biochemistry, and Clinical Chemistry; while the most cited journals were Clinical Chemistry, Circulation, and JAMA-Journal of the American Medical Association. The co-occurrence network visualizations of keywords and terms provided a global overview of LDL-C estimating equations. The study presented provides valuable insights into potential research avenues within the examined field, serving as a foundation for future research initiatives.

Keywords: Bibliometric analysis, Cardiovascular disease, Low-density lipoprotein cholesterol, LDL-C, VOSviewer.

^adrsolat@gmail.com

^bahucephe.87@gmail.com

^cgozdeerturk@erciyes.edu.tr

^dgokmen.zararsiz@gmail.com

^e<https://orcid.org/0000-0002-8669-8782>

^f<https://orcid.org/0000-0001-9374-4495>

^g<https://orcid.org/0000-0002-5495-7540>

^h<https://orcid.org/0000-0001-5801-1835>

ⁱserrayerlitas@erciyes.edu.tr

^jnecla.kayaalp@gmail.com

^khalefokan@gmail.com

^l<https://orcid.org/0000-0003-2830-3006>

^m<https://orcid.org/0000-0003-2355-4826>

ⁿ<https://orcid.org/0000-0001-8738-0760>

Introduction

Cardiovascular diseases (CVD) cover a wide range of conditions affecting the heart and blood vessels, posing a significant global health challenge [1]. Several risk factors contribute to CVD development, including high blood pressure, diabetes, and dyslipidemia [2]. Dyslipidemia refers to irregularities in lipid levels, specifically an increase in low-density lipoprotein cholesterol (LDL-C). LDL-C concentration is recognized as a crucial factor in cardiovascular health, with higher levels associated with an increased risk of atherosclerosis and coronary artery disease. Several studies emphasize the importance of managing LDL-C levels to mitigate cardiovascular risk. As an essential component of lipid profiles, understanding and controlling LDL-C concentration play a pivotal role in cardiovascular risk assessment and preventive strategies [3, 4].

The gold standard for LDL-C measurement is ultracentrifugation followed by beta-quantification, which accurately isolates the LDL-C fraction [5]. However, this method is impractical for routine use due to its high cost, time requirements, and the need for large sample volumes and specialized equipment. The Friedewald formula, introduced in 1972, has become a global

standard for LDL-C estimation in clinical practice due to its convenience and cost-effectiveness [6]. Despite its widespread use, the Friedewald formula has several limitations. It is sensitive to triglyceride levels, leading to inaccuracies, particularly in high triglyceride concentrations [7]. It tends to be less accurate in cases with low LDL-C levels, potentially resulting in unreliable LDL-C estimates [8]. Additionally, its fixed triglyceride (TG)—very low-density lipoprotein (VLDL-C) ratio does not account for individual variations, resulting in further inaccuracies in LDL-C estimation [9]. This may lead to overestimation of hypercholesterolemia prevalence compared to direct measurement methods [10]. Alternative methods like the Martin/Hopkins and Sampson formulas provide improved accuracy, particularly in cases with high triglycerides and non-fasting samples, making them valuable alternatives in clinical settings [6, 7, 11].

In Turkey, there has been a significant increase in studies focused on the validation of LDL-C formulas [6, 12]. Notably, there is a growing trend towards developing new methodologies utilizing machine models, including linear regression, artificial neural networks, Bayesian-

regularized neural networks, random forests, decision trees, and gradient-boosted trees [13-15]. These investigations indicate that machine learning algorithms offer more reliable predictions and classifications of LDL-C levels compared to traditional formulas, thereby facilitating personalized healthcare interventions with more accurately assessed cardiovascular risks in clinical applications.

Bibliometric analysis is a quantitative method for assessing scientific publications within a specific field. By applying statistical techniques to bibliographic data, it reveals patterns, trends, and relationships among academic works. This method provides insights into scholarly communication dynamics and evaluates the impact, productivity, and influence of authors, journals, institutions, and research topics. Ultimately, bibliometric analysis serves as a valuable tool for exploring scholarly publications and offers critical insights for researchers, policymakers, and institutions [16-18].

Only one systematic review and meta-analysis on LDL-C estimation has been identified in the literature. Ephraim et al. (2023) evaluated existing equations for LDL-C estimation using a random-effects meta-analysis approach [19]. To our knowledge, no bibliometric analysis specifically addressing LDL-C estimation equations exists in the literature. This study aims to provide a global overview of LDL-C estimating equations through bibliometric analysis, focusing on i) trends in article and citation volume over time, ii) prevalent research areas, iii) the most productive countries and institutions, iv) leading journals, v) highly cited papers and author collaborations, and vi) frequently used keywords and their co-occurrence networks within the field.

Materials and Methods

Data Collection and Search Strategy

The data for the analysis was retrieved from the Web of Science (WoS) database on January 1, 2024. The query employed during the scan of publications is outlined below:

(AB=("LDL-C" AND "equation") OR AB=("LDL-C Estimation") OR AB=("LDL-C estimating equation") OR AB=("Martin-Hopkins" AND "formula") OR AB=("Martin-Hopkins" AND "equation") OR AB=("Sampson" AND "formula") OR AB=("Sampson" AND "equation") OR AB=("LDL-C calculation") OR AB=("novel formula" AND "LDL-C") OR AB=("novel equation" AND "LDL-C") OR AB=("calculation of LDL-C") OR AB=("estimation of LDL-C") OR AB=("Friedewald's") OR QMTS=("Friedewald Formula") OR QMTS=("Friedewald Equation") OR QMTS=("Friedewald S Formula") OR QMTS=("Friedewald S Equation") OR QMTS=("Sampson Equation") OR QMTS=("Sampson Formula") OR QMTS=("Martin Hopkins Equation") OR QMTS=("Anandaraja S Formula"))

The pertinent data from the literature was exported to a Plain Text File (.txt) by selecting the record content, specifically, 'Full Record and Cited References'.

Bibliometric Analysis

As a result of the search query, we obtained 734 records published in the literature on LDL-C estimating equations between 1990 and 2023. Of these records, 74 records in the document types meeting abstracts, proceeding papers, editorial materials, early access, letters, book chapters and reviews, data papers, and notes were filtered. Of the remaining 660 original research and review articles, 40 were filtered because they were not written in the English language. After filtering, the remaining 620 original research and review articles in English were included in the bibliometric analyses (Figure 1). These articles had 12,347 citations, each article was cited an average of 25.6 times, and the H-index was 59. Figure 2 shows the changes in the number of articles and citations in this field over the years. It is seen that there is a general increase in both the number of articles and citations, and peak values were reached in 2022-2023. An important finding is that 316 articles were published between 2016 and 2023, accounting for 51% of the total number of articles.

For analyses such as research trends, content, and keyword analyses, we utilized biblioshiny 4.1.4 [20], a web interface for the bibliometrix R package, and VOSviewer 1.6.20 software [21].

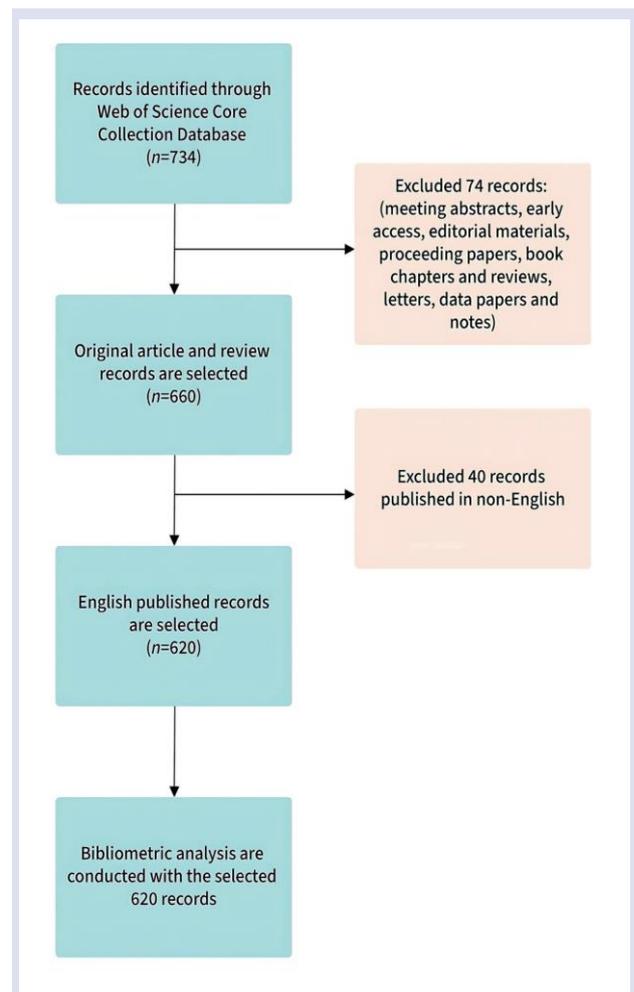


Figure 1. Flow chart of study selection process.

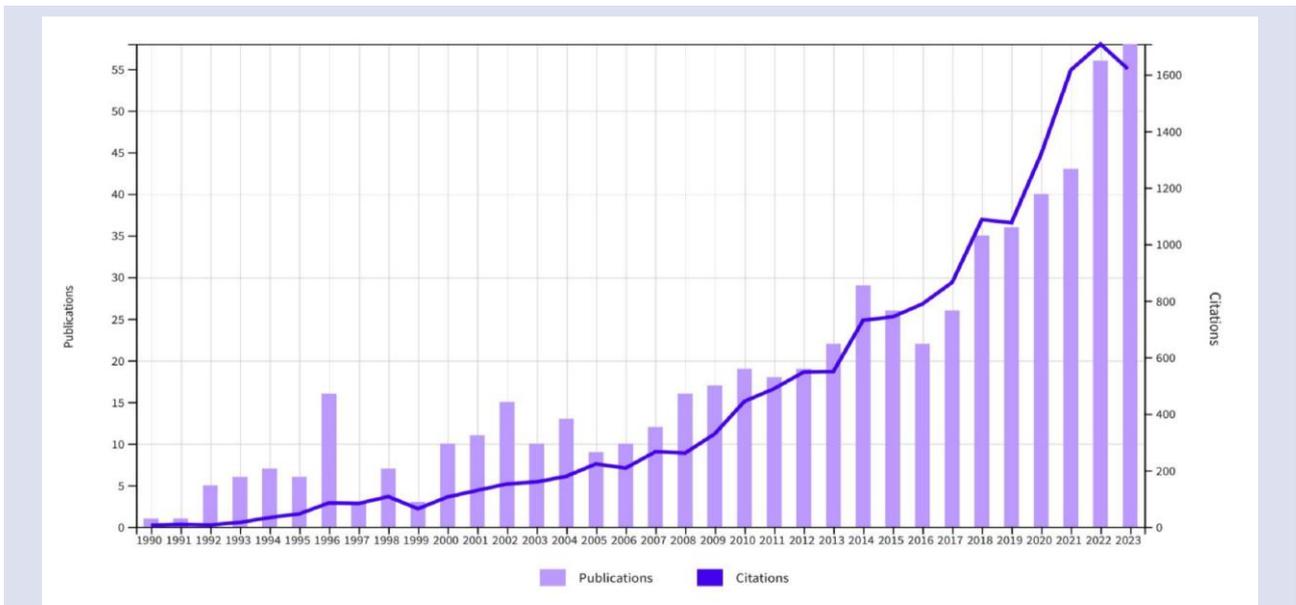


Figure 2. The number of publications per year between 1990 and 2023.

Results

Distribution of Research Areas

The 620 articles included in the bibliometric analyses were related to 57 WoS research areas. The 10 research areas with the highest number of articles were medical laboratory technology (n=120, 19.4%), cardiac & cardiovascular systems (n=114, 18.4%), medicine, general & internal (n=103, 16.6%), endocrinology & metabolism (n=62, 10.0%), nutrition & dietetics (n=47, 7.6%), pharmacology & pharmacy (n=38, 6.1%), biochemistry & molecular biology (n=36, 5.8%), medicine, research & experimental (n=29, 4.7%), science, technology & other topics (n=18, 7.6%), and public, environmental & occupational health (n=17, 2.7%).

Analysis of Main Countries and Top Country Institutions

To better understand the current state of research on LDL-C estimating equations, we conducted country- and institution-level analyses. 620 studies between 1990 and 2023 were conducted by institutions in 62 different countries. The countries with the highest number of publications were the USA (n=436), China (n=248) and Japan (n=224), while the countries with the highest average citations were the USA (n=42.8), UK (n=36.0) and Japan (n=31.5) (Table 1). When analyzed by institutions, the most productive institutes of the countries in Table 1 were Johns Hopkins University (USA, n=48), Yonsei University (South Korea, n=21), and Tehran University of Medical Sciences (Iran, n=20). When ranked based on all institutes, the top 5 universities were as follows: Johns Hopkins University (USA, n=48), Harvard University (USA, n=24), Yonsei University (South Korea, n=21), Tehran University of Medical Sciences (Iran, n=20), and National Taiwan University (Taiwan, n=17).

Table 1. Top ten countries and institutions with the highest number of publications on LDL-C estimating equations (1990-2023)

| Country | Number of Articles | Total Citations | Average Article Citations | Top Country Institution | Top Institution Articles |
|-------------|--------------------|-----------------|---------------------------|---------------------------------------|--------------------------|
| USA | 436 | 4,750 | 42.8 | Johns Hopkins University | 48 |
| China | 248 | 667 | 10.8 | Central South University | 10 |
| Japan | 224 | 1,450 | 31.5 | Kyoto University | 11 |
| Spain | 108 | 422 | 17.6 | Universitat Autònoma de Barcelona | 8 |
| Iran | 101 | 353 | 13.1 | Tehran University of Medical Sciences | 20 |
| India | 95 | 314 | 7.5 | Carmel Medical Centre | 6 |
| South Korea | 89 | 343 | 13.7 | Yonsei University | 21 |
| UK | 88 | 720 | 36.0 | Imperial College London | 13 |
| Canada | 77 | 602 | 27.4 | Laval University | 10 |
| Finland | 61 | 87 | 14.5 | University of Turku | 9 |

LDL-C, low-density lipoprotein cholesterol.

Most Productive and Cited Journals

The conducted analysis showed that the articles were published in 329 different journals. According to Bradford’s law, journals in a particular field undergo ranking based on their publication count. Subsequently, these journals are categorized into three segments, each holding an identical number of publications. The segment with the fewest journals is denoted as the core. The core group houses the most productive journals, whereas in the remaining segments, the number of journals rises while maintaining an equivalent publication count to the core group. The Bradford Law Graph, demonstrating the most productive journals located in the core group, is given in Figure 3. According to this plot, the number of core sources with the largest number of LDL-C estimating

equations publications is 17, according to 5.2% of the entire literature.

When the citations in these articles were evaluated, the 10 journals with the highest number of citations were as follows: Clinical Chemistry, Circulation, JAMA-Journal of the American Medical Association, Atherosclerosis, Journal of the American College of Cardiology, Lancet, New England Journal of Medicine, Diabetes Care, Clinica Chimica Acta, and European Heart Journal. The number of

records of the most productive and the most cited journals, as well as journal impact factors (JIF), JIF quartiles, and journal citation reports (JCR) research categories, are shown in Table 2. Of these 17 journals in the core group, 2 (12%) had an impact factor above 5, and 13 (76%) were among the top 50%. The journal with the highest number of articles was Clinica Chimica Acta (n=28, 4.5%), while the journal with the highest impact factor was Metabolism-Clinical and Experimental.

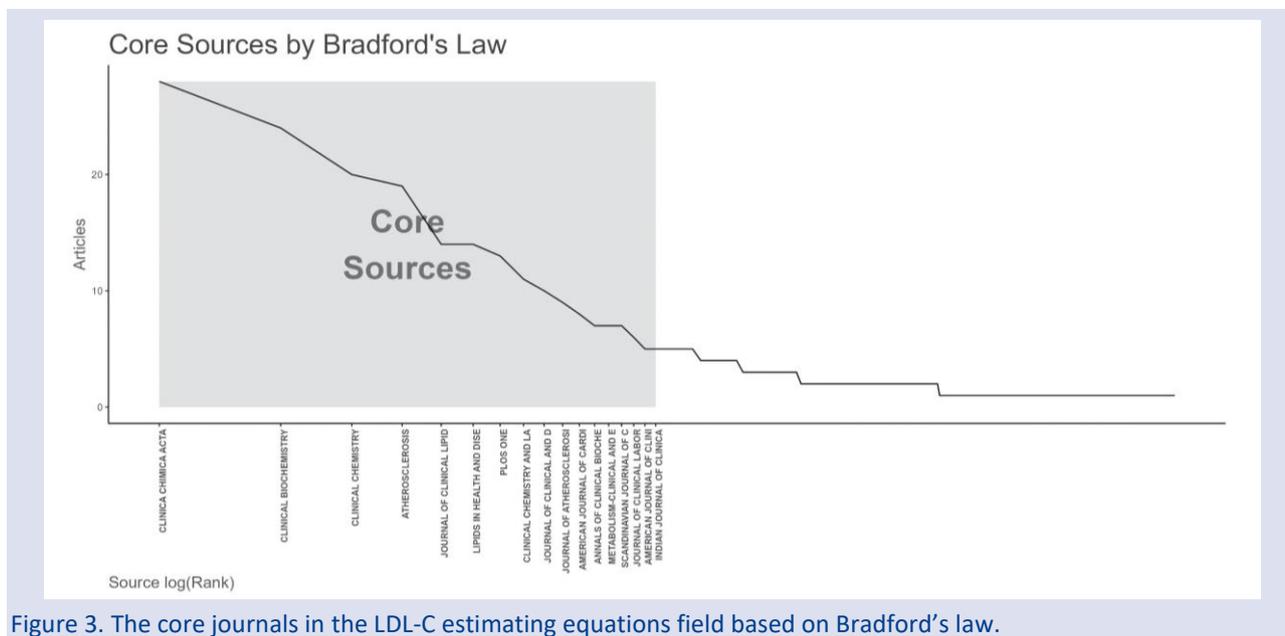


Figure 3. The core journals in the LDL-C estimating equations field based on Bradford’s law.

Table 2. The most productive and cited journals for research publications on LDL-C Estimating Equations (1990-2023)

| Journal Name | Number of Records | JIF (2023) | JIF Quartile (2023) | JCR Category |
|---|-------------------|------------|---------------------|---|
| Most Productive Journals | | | | |
| Clinica Chimica Acta | 28 art. | 3.2 | Q2 | Medical Laboratory Technology |
| Clinical Biochemistry | 24 art. | 2.5 | Q2 | Medical Laboratory Technology |
| Clinical Chemistry | 20 art. | 7.1 | Q1 | Medical Laboratory Technology |
| Atherosclerosis | 19 art. | 4.9 | Q1 | Cardiac & Cardiovascular Systems, Peripheral Vascular Disease |
| Journal of Clinical Lipidology | 14 art. | 3.6 | Q2 | Pharmacology & Pharmacy |
| Lipids in Health and Disease | 14 art. | 3.9 | Q2 | Biochemistry & Molecular Biology, Nutrition & Dietetics |
| Most Cited Journals | | | | |
| Plos One | 13 art. | 2.9 | Q1 | Multidisciplinary Sciences |
| Clinical Chemistry and Laboratory Medicine | 11 art. | 3.8 | Q1 | Medical Laboratory Technology |
| Journal of Clinical and Diagnostic Research | 10 art. | 0.2 | Q4 | Medicine, General & Internal |
| Journal of Atherosclerosis and Thrombosis | 9 art. | 3.0 | Q2 | Peripheral Vascular Disease |
| American Journal of Cardiology | 8 art. | 2.3 | Q2 | Cardiac & Cardiovascular Systems |
| Annals of Clinical Biochemistry | 7 art. | 2.1 | Q3 | Medical Laboratory Technology |
| Metabolism-Clinical and Experimental | 7 art. | 10.8 | Q1 | Endocrinology & Metabolism |
| Scandinavian Journal of Clinical & Laboratory Investigation | 7 art. | 1.3 | Q4 | Medicine, Research & Experimental |
| Journal of Clinical Laboratory Analysis | 6 art. | 2.6 | Q2 | Medical Laboratory Technology |
| American Journal of Clinical Pathology | 5 art. | 2.3 | Q2 | Pathology |
| Indian Journal of Clinical Biochemistry | 5 art. | 1.5 | Q4 | Biochemistry & Molecular Biology |
| Most Cited Journals | | | | |
| Clinical Chemistry | 1,930 cit. | 7.1 | Q1 | Medical Laboratory Technology |
| Circulation | 857 cit. | 35.5 | Q1 | Cardiac & Cardiovascular Systems, Peripheral Vascular Disease |
| JAMA-Journal of the American Medical Association | 709 cit. | 63.1 | Q1 | Medicine, General & Internal |
| Atherosclerosis | 551 cit. | 4.9 | Q1 | Cardiac & Cardiovascular Systems, Peripheral Vascular Disease |
| Journal of the American College of Cardiology | 485 cit. | 21.7 | Q1 | Cardiac & Cardiovascular Systems |
| Lancet | 466 cit. | 98.4 | Q1 | Medicine, General & Internal |
| New England Journal of Medicine | 419 cit. | 96.2 | Q1 | Medicine, General & Internal |
| Diabetes Care | 400 cit. | 14.8 | Q1 | Endocrinology & Metabolism |
| Clinica Chimica Acta | 382 cit. | 3.2 | Q2 | Medical Laboratory Technology |
| European Heart Journal | 331 cit. | 37.6 | Q1 | Cardiac & Cardiovascular Systems |

LDL-C, low-density lipoprotein cholesterol; JIF, journal impact factor; JCR, journal citation reports; art., articles; cit., citations.

Highly Cited Papers and the Analysis of Main Authors

The most cited 10 studies are presented in Table 3. Of these 10 studies, 7 were published in Q1 ranked and 3 in Q2 ranked journals. It is noteworthy that the lowest journal impact factor of the Q1 ranked journals is 7.1. Seven of the most cited articles were from the USA and 3 articles were published in Clinical Chemistry. In the study published in JAMA-Journal of the American Medical Association, which has the highest impact factor (63.1) and the highest number of citations (481), Martin et al. (2013) proposed an alternative formula to the Friedewald

formula, which uses a constant coefficient of 5 for the TG:VLDL-C ratio, and validated the formula in data including 1,350,908 children, adolescents and adults in whom actual LDL-C measurements were made by ultracentrifugation [22]. Another interesting study is the study by Sampson et al. published in JAMA Cardiology where the normalized citations per year statistic is the highest [11]. In this study, Sampson et al. (2020) developed a new formula that works specifically for hypertriglyceridemia and/or a low level of low-density lipoprotein cholesterol and compared their results to the Martin-Hopkins formula.

Table 3. Top 10 most highly cited documents on LDL-C estimating equations (1990-2023)

| Authors | Article Name | Published Year | Country | Total Citations | Citations per Year | Normalized Citations per Year | Journal | JIF 2023 | JIF Quartile 2023 |
|----------------------|--|----------------|---------|-----------------|--------------------|-------------------------------|---|----------|-------------------|
| Martin et al. [22] | Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile | 2013 | USA | 481 | 40.08 | 8.91 | JAMA-Journal of the American Medical Association | 63.1 | Q1 |
| Warnick et al. [23] | Estimating low-density lipoprotein cholesterol by the Friedewald equation is adequate for classifying patients on the basis of nationally recommended cutpoints | 1990 | USA | 398 | 11.37 | 1.00 | Clinical Chemistry | 7.1 | Q1 |
| Nauck et al. [24] | Methods for measurement of LDL-cholesterol: a critical assessment of direct measurement by homogeneous assays versus calculation | 2002 | Germany | 271 | 11.78 | 4.38 | Clinical Chemistry | 7.1 | Q1 |
| Martin et al. | Friedewald-estimated versus directly measured low-density lipoprotein cholesterol and treatment implications | 2013 | USA | 263 | 21.92 | 4.87 | Journal of the American College of Cardiology | 21.7 | Q1 |
| McNama | Change in LDL particle size is associated with change in plasma triglyceride concentration | 1992 | USA | 240 | 7.27 | 2.99 | Arteriosclerosis, Thrombosis and Vascular Biology | 7.4 | Q1 |
| Sampson et al. [11] | A New Equation for Calculation of Low-Density Lipoprotein Cholesterol in Patients With Normolipidemia and/or Hypertriglyceridemia | 2020 | USA | 203 | 40.6 | 16.41 | JAMA Cardiology | 14.8 | Q1 |
| Tremblay et al. [26] | Validation of the Friedewald formula for the determination of low-density lipoprotein cholesterol compared with beta-quantification in a large population | 2004 | Canada | 167 | 7.95 | 6.17 | Clinical Biochemistry | 2.5 | Q2 |
| Frost & Havel [27] | Rationale for use of non-high-density lipoprotein cholesterol rather than low-density lipoprotein cholesterol as a tool for lipoprotein cholesterol screening and assessment of risk and therapy | 1998 | USA | 148 | 5.48 | 2.66 | American Journal of Cardiology | 2.3 | Q2 |
| Chen et al. [28] | A modified formula for calculating low-density lipoprotein cholesterol values | 2010 | China | 140 | 9.33 | 5.26 | Lipids in Health and Disease | 3.9 | Q2 |
| Rifai et al. [29] | Measurement of low-density-lipoprotein cholesterol in serum: a status report | 1992 | USA | 132 | 4.00 | 1.65 | Clinical Chemistry | 7.1 | Q1 |

LDL-C, low-density lipoprotein cholesterol; JIF, journal impact factor.

The authors' productivity over time is shown in Figure 4. It can be seen that these two works have had a significant impact on the literature on LDL-C estimating equations. When we evaluate each article specifically, it can be said that most of the recent studies conducted validation and performance comparison studies of Martin-Hopkins and Sampson formulas in different populations.

The scientific collaborations between the main authors are given in Figure 5. The connections among the main authors represent co-authorships, with the thickness of these connections indicating the different frequencies of collaboration. Wider links signify a higher level of collaboration between the two main authors.

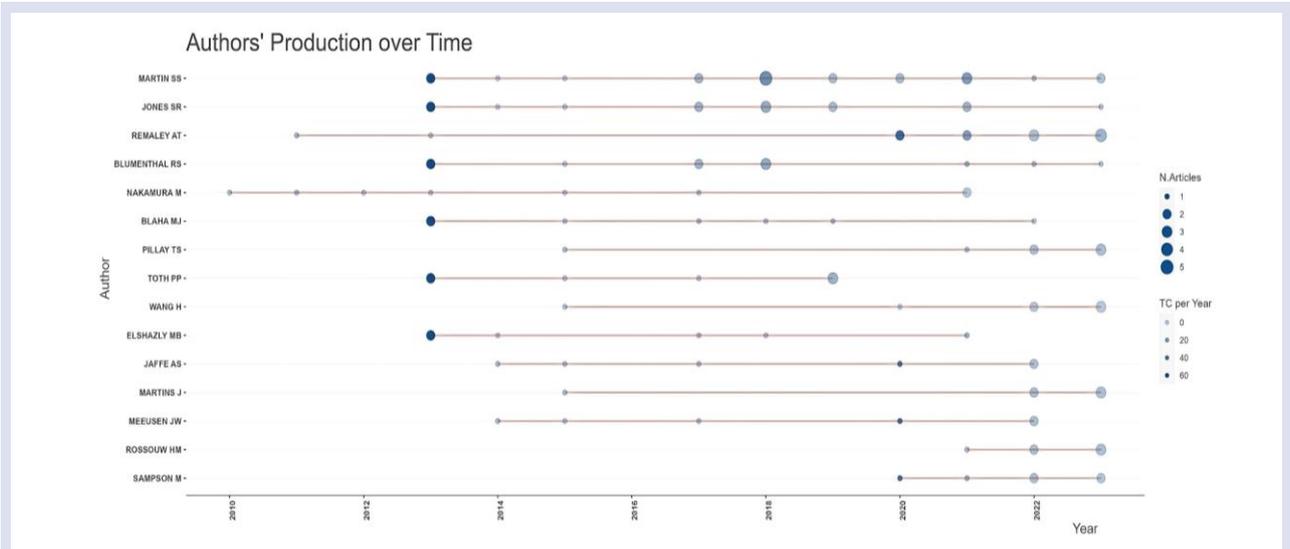


Figure 4. Top authors' production over time (1990-2023).

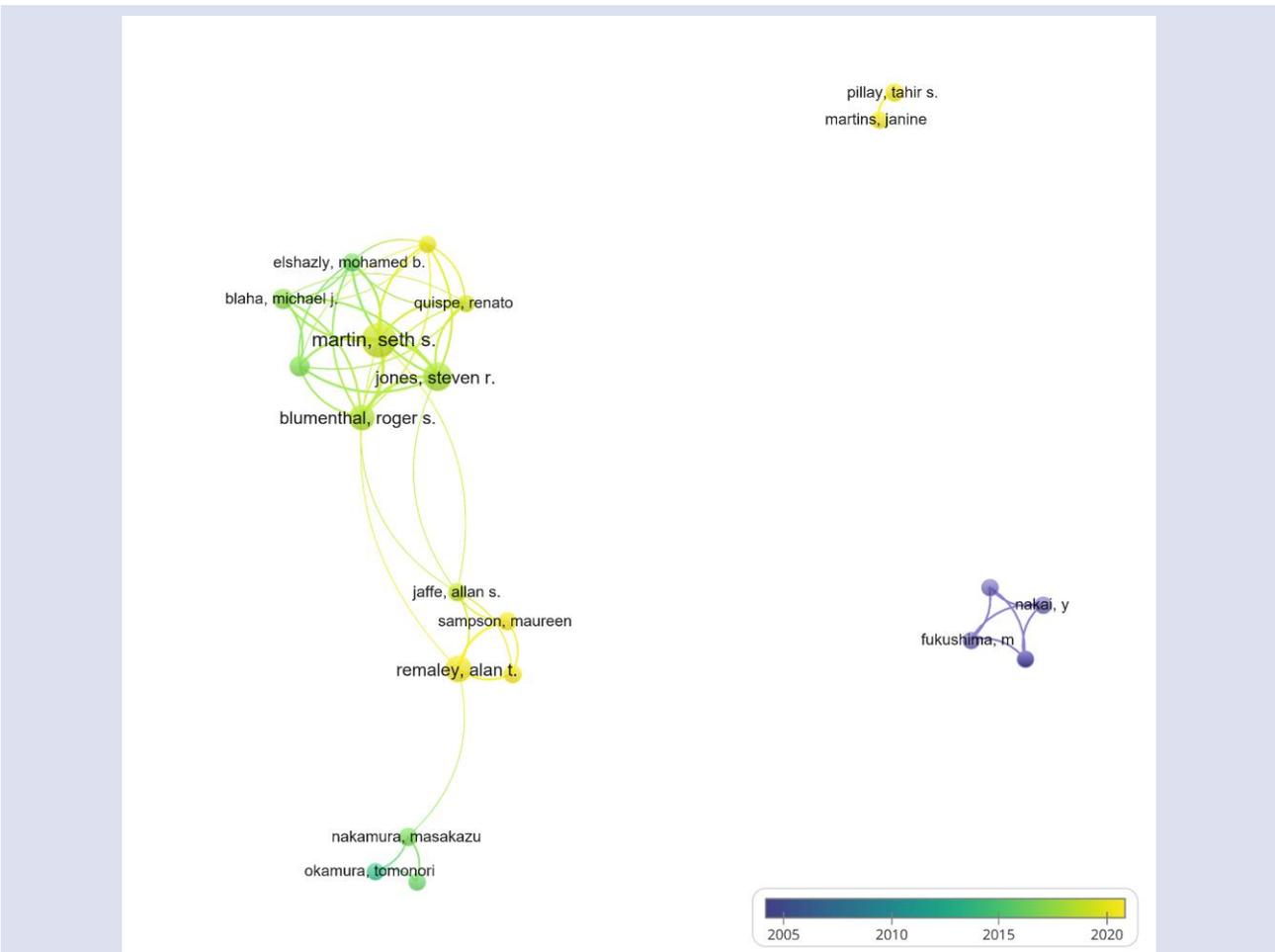


Figure 5. The authors' scientific collaboration network.

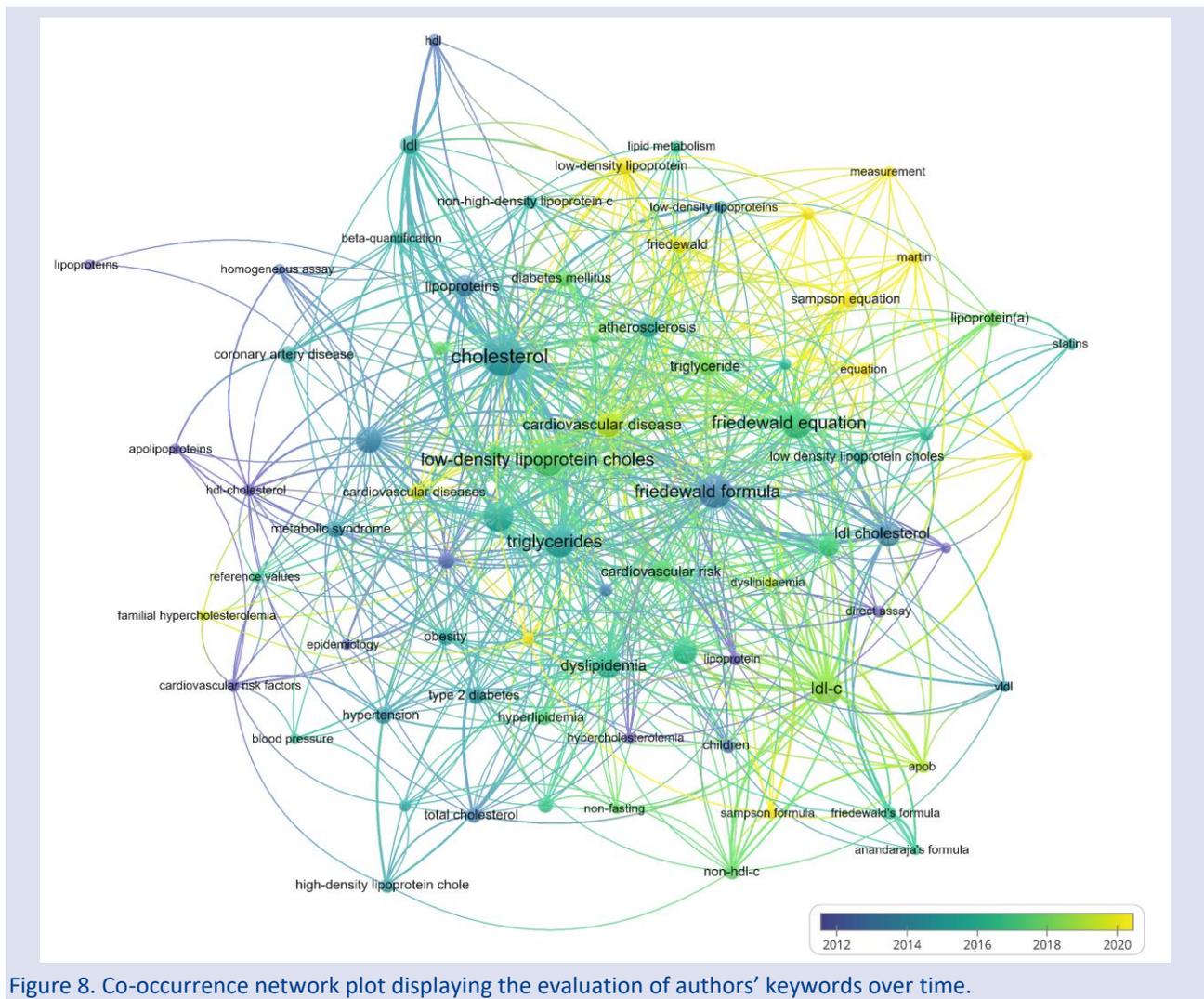


Figure 8. Co-occurrence network plot displaying the evaluation of authors' keywords over time.

Discussion

In this study, a bibliometric analysis of LDL-C estimation equations was performed. This bibliometric analysis was carried out using VOSviewer and Bibliometrix, it is aimed to understand the current state of the literature on LDL-C estimation equations and to present the studies conducted until January 1, 2024 in a global and accessible way. The conducted bibliometric analysis resulted in 620 studies in the WoS database between 1990 and 2023. It is noteworthy that the number of publications in this field has increased by 290% in the last 10 years and 179% in the last 5 years. The WoS research categories of the journals in which these studies were published were medical laboratory technology, cardiac & cardiovascular systems, medicine, general & internal, endocrinology & metabolism, nutrition & dietetics, pharmacology & pharmacy, biochemistry & molecular biology, and medicine, research & experimental. These research areas are the main areas of knowledge where the calculation of LDL-C levels is of importance. The studies were mostly published in journals with high-impact factors in the fields of biochemistry and cardiology such as Journal of the American College of Cardiology, JAMA Cardiology, Clinica Chimica Acta, Clinical

Biochemistry, and Clinical Chemistry. The role of LDL-C in the increased risk of cardiovascular disease has been shown by scientific studies [30, 31]. These analyzed studies focused on the development and validation of the most accurate and simple equations to measure LDL-C levels for this problem.

As a result of the study, the top 10 list of articles on LDL-C estimating equations is shared. The top three countries and the top country institutions are the USA (John Hopkins University), China (Central South University), and Japan (Kyoto University). According to the Centers for Disease Control and Prevention, cardiovascular disease is the leading cause of death, with one death from cardiovascular disease every 33 seconds (1 in 5 deaths in total) in the USA [32]. Since LDL-C level is an important risk factor for cardiovascular disease, we believe that studies on LDL-C estimation equations are mostly conducted in the USA. It is observed that LDL-C estimations and subsequent validation studies for different populations are common in the literature.

SS Martin, the author with the highest number of publications on this topic, is also the main author of the most cited study in the field. In this study, Martin et al. (2013) proposed their Martin-Hopkins formula instead of the Friedewald formula, which has an important place in

clinical practice, and validated the formula in data from 1,350,908 children, adolescents, and adults living in the USA [22]. In the other most cited studies, the validity of the Friedewald formula, which is widely accepted in the literature, was investigated and its compatibility with direct methods was evaluated [7, 23, 24, 26]. McNamara et al. (1992) demonstrated the variation of LDL-C particle numbers according to triglyceride levels [25]. The other most cited study is the development of the Sampson formula for low LDL-C levels and/or hypertriglyceridemia (TG levels, ≤ 800 mg/dL) [11].

When the co-occurrence network graph displaying the evaluation of the authors' keywords over time is analyzed, it is seen that keywords such as "atherosclerosis", "dyslipidemia", "triglycerides", and "cardiovascular risk" used in previous years were replaced by words such as "Sampson equation", "Martin", and "equation". Therefore, it can be said that studies on LDL-C prediction equations have intensified in recent years, and studies on investigating the validity of Martin-Hopkins and Sampson equations in different populations and comparing their performance (including the Friedewald formula) have increased.

One limitation of the study is that only the Web of Science database was used for exporting data. This could have resulted in studies being influenced by bias and having incomplete inclusion.

Conclusions

To the extent of our knowledge, this is the first bibliometric analysis study in the field of LDL-C estimation equations. In this study, we analyzed the sources, countries and institutions, authors, keywords, and themes of LDL-C estimation equations articles published between 1990 and 2023. In addition, the trends in the literature on the topic are shared. The results showed that LDL-C estimating equations research captured significant interest from researchers, especially in the last seven years. In summary, studies on LDL-C estimating equations may receive more and more attention, new equations may be developed or validation of existing equations in different populations or systems may be investigated. This study can guide researchers in this field to quickly understand the knowledge structures in this area.

Conflicts of interest

There are no conflicts of interest in this work.

Ethical Approval Statement

Ethics committee approval is not required as there is no human or animal research.

References

- [1] Mensah G.A., Roth G.A., Fuster V., The global burden of cardiovascular diseases and risk factors: 2020 and beyond, *J. Am. Coll. Cardiol.*, 74 (20) (2019) 2529-32.
- [2] Hedayatnia M., Asadi Z., Zare-Feyzabadi R., Yaghooti-Khorasani M., Ghazizadeh H., Ghaffarian-Zirak M. et al, Dyslipidemia and cardiovascular disease risk among the MASHAD study population, *Lipids Health Dis.*, 19 (1) (2020) 1-11.
- [3] Mortensen M.B., Dzaye O., Bøtker H.E., Jensen J.M., Maeng M., Bentzon J.F., et al, Low-Density Lipoprotein Cholesterol Is Predominantly Associated With Atherosclerotic Cardiovascular Disease Events in Patients With Evidence of Coronary Atherosclerosis: The Western Denmark Heart Registry, *Circulation*, 147 (14) (2023) 1053-1063.
- [4] Upadhyay R.K., Emerging Risk Biomarkers in Cardiovascular Diseases and Disorders, *J. Lipids*, 2015 (2015) 1-50.
- [5] Steyn N., Muller Rossouw H., Pillay T.S., Martins J., Comparability of calculated LDL-C with directly measured LDL-C in selected paediatric and adult cohorts, *Clin. Chim. Acta*, 537 (2022) 158-166.
- [6] Ertürk Zararsız G., Bolat S., Cephe A., Kochan N., Yerlitas S.I., Dogan H.O., et al, Validation of Friedewald, Martin-Hopkins and Sampson low-density lipoprotein cholesterol equations, *PLoS One*, 17 (5) (2022) e0263860.
- [7] Martin S.S., Blaha M.J., Elshazly M.B., Brinton E.A., Toth P.P., McEvoy J.W., et al, Friedewald-estimated versus directly measured low-density lipoprotein cholesterol and treatment implications, *J. Am. Coll. Cardiol.*, 62 (8) (2013) 732-739.
- [8] Samuel C., Park J., Sajja A., Michos E.D., Blumenthal R.S., Jones S.R., et al, Accuracy of 23 Equations for Estimating LDL Cholesterol in a Clinical Laboratory Database of 5,051,467 Patients, *Glob. Heart*, 18 (1) (2023) 36.
- [9] Chung S., Usefulness of the Martin Method for Estimation of Low-Density Lipoprotein Cholesterol in Coronary Atherosclerosis, *Med. Princ. Pract.*, 27 (1) (2018) 8-14.
- [10] Kapoor R., Chakraborty M., Singh N., A Leap above Friedewald Formula for Calculation of Low-Density Lipoprotein-Cholesterol, *J. Lab. Physicians*, 7 (01) (2015) 011-016.
- [11] Sampson M., Ling C., Sun Q., Harb R., Ashmaig M., Warnick R., et al, A New Equation for Calculation of Low-Density Lipoprotein Cholesterol in Patients with Normolipidemia and/or Hypertriglyceridemia, *JAMA Cardiol.*, 5 (5) (2020) 540-548.
- [12] Alpdemir M.F., Alpdemir M., Comparison of different equations for estimation of low-density lipoprotein (LDL)-cholesterol, *Turk. J. Biochem.*, 45 (5) (2020) 601-611.
- [13] Çubukçu H.C., Topcu D.İ., Estimation of low-density lipoprotein cholesterol concentration using machine learning, *Lab. Med.*, 53 (2) (2022) 161-171.
- [14] Hataysal E.P., Körez M.K., Yeşildal F., İşman F.K., A comparative evaluation of low-density lipoprotein cholesterol estimation: Machine learning algorithms versus various equations, *Clin. Chim. Acta*, 557 (2024) 117853.
- [15] Sezer S., Oter A., Ersoz B., Topcuoglu C., Bulbul H.İ., Sagioglu S., et al, Explainable artificial intelligence for LDL cholesterol prediction and classification, *Clin. Biochem.*, 130 (2024) 110791.
- [16] Ellegaard O., Wallin J.A., The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105 (3) (2015) 1809-1831.
- [17] Maniu I., Maniu G., Totan M., Clinical and Laboratory Characteristics of Pediatric COVID-19 Population—A Bibliometric Analysis, *J. Clin. Med.*, 11 (20) (2022) 5987.
- [18] Feng X.W., Hadizadeh M., Zheng L.H., Li W.H., A

- Bibliometric and Visual Analysis of Exercise Intervention Publications for Alzheimer's Disease (1998–2021), *J. Clin. Med.*, 11 (19) (2022) 5903.
- [19] Ephraim R.K.D., Ativi E., Ashie S.A., Abaka-Yawson A., Darkwah K.O., Assessment of estimated low-density lipoprotein-cholesterol (LDL-C) equations: a systematic review and meta-analysis, *Bull. Natl. Res. Cent.*, 47 (1) (2023) 71.
- [20] Aria M., Cuccurullo C., bibliometrix: An R-tool for comprehensive science mapping analysis, *J. Informetr.*, 11 (4) (2017) 959-975.
- [21] van Eck N.J., Waltman L., Software survey: VOSviewer, a computer program for bibliometric mapping, *Scientometrics*, 84 (2) (2010) 523-538.
- [22] Martin S.S., Blaha M.J., Elshazly M.B., Toth P.P., Kwiterovich P.O., Bluementhal R.S., et al, Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile, *JAMA*, 310 (19) (2013) 2061-2068.
- [23] Warnick G.R., Knopp R.H., Fitzpatrick V., Branson L., Estimating low-density lipoprotein cholesterol by the Friedewald equation is adequate for classifying patients on the basis of nationally recommended cutpoints, *Clin. Chem.*, 36 (1) (1990) 15-19.
- [24] Nauck M., Warnick G.R., Rifai N., Methods for measurement of LDL-cholesterol: A critical assessment of direct measurement by homogeneous assays versus calculation, *Clin. Chem.*, 48 (2) (2002) 236-254.
- [25] McNamara J.R., Jenner J.L., Li Z., Wilson P.W.F., Schaefer E.J., Change in LDL particle size is associated with change in plasma triglyceride concentration, *Arterioscler. Thromb. Vasc. Biol.*, 12 (11) (1992) 1284-1290.
- [26] Tremblay A.J., Morrissette H., Gagné J.M., Bergeron J., Gagné C., Couture P., Validation of the Friedewald formula for the determination of low-density lipoprotein cholesterol compared with β -quantification in a large population, *Clin. Biochem.*, 37 (9) (2004) 785-790.
- [27] Frost P.H., Havel R.J., Rationale for use of non-high-density lipoprotein cholesterol rather than low-density lipoprotein cholesterol as a tool for lipoprotein cholesterol screening and assessment of risk and therapy, *Am. J. Cardiol.*, 81 (4 A) (1998) 26B-31B.
- [28] Chen Y., Zhang X., Pan B., Jin X., Yao H., Chen B., et al, A modified formula for calculating low-density lipoprotein cholesterol values, *Lipids Health Dis.*, 9 (1) (2010) 1-5.
- [29] Rifai N., Warnick G.R., McNamara J.R., Belcher J.D., Grinstead G.F., Ivan D Frantz J., Measurement of low-density-lipoprotein cholesterol in serum: a status report, *Clin. Chem.*, 38 (1) (1992) 150-160.
- [30] Domanski M.J., Tian X., Wu C.O., Reis J.P., Dey A.K., Gu Y., et al, Time Course of LDL Cholesterol Exposure and Cardiovascular Disease Event Risk, *J. Am. Coll. Cardiol.*, 76 (13) (2020) 1507-1516.
- [31] Stanculescu L.A., Scafa-Udriste A., Dorobantu M., Exploring the Association between Low-Density Lipoprotein Subfractions and Major Adverse Cardiovascular Outcomes—A Comprehensive Review, *Int. J. Mo.l Sci.* 24 (7) (2023) 6669.
- [32] Centers for Disease Control and Prevention. Available at: <https://www.cdc.gov/heartdisease/facts.htm>. Retrieved January 3, 2024.

Comparative Evaluation of Chemical Compositions and Antioxidant Activities of *Passiflora edulis* Sims Extracts

Emine Kılıçkaya Selvi^{1,a,*}

¹ Department of Chemistry and Chemical Processing Technologies, Mustafa Cikirkioglu Vocational School, Kayseri University, Türkiye

*Corresponding author

Research Article

History

Received: 30/06/2024

Accepted: 03/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Passiflora members have an important quality in both conventional and modern medicine. In the present study *Passiflora edulis* Sims from Mediterranean Turkey (Antalya-Gazipaşa) was investigated from the aspects of phenolic compounds and antioxidant activities. The plant material (stems, leaves) used as a whole and chloroform, ethyl acetate and methanol were used for the extraction. High performance liquid chromatography was used to analyse the phenolic compounds of the extracts. Besides, the antioxidant potential of the extracts was characterised by the total phenolic content, the total flavonoid content, the DPPH (2,2-diphenyl-1-picrylhydrazyl), the CUPRAC (Cupric Ion Reducing Antioxidant Capacity) and the FRAP (ferric reducing antioxidant power assay) tests. Methanol extract was found to be more active than the other extracts in the all antioxidant experiment. DPPH radical scavenging activity of methanol extract was 0.028 ± 0.001 mg/mL, FPAP assay was 0.511 ± 0.012 $\mu\text{molGA/g}$ and CUPRAC assay was 3.728 ± 0.150 $\mu\text{molGA/g}$. As a result, methanol extract of *Passiflora edulis* was established ferulic acid (175.12 mg std/g) and *p*-coumaric acid (116.88 mg std/g) were determined as primary phenolic compounds in different amounts along with protocatechuic acid, chlorogenic acid, caffeic acid, apigenin and isorhanmetin. The results showed that *Passiflora edulis* deserves to be studied both for its use as a food and for its therapeutic properties.

Keywords: Antioxidant activity, HPLC-DAD, *Passiflora edulis*, Turkey.

^a emineselvi@kayseri.edu.tr

^{ORCID} <https://orcid.org/0000-0003-0291-5362>

Introduction

Reactive oxygen species (ROS) play an important role in human diseases and disorders. Accordingly, ROS level/concentration is an important factor in the growth, progression and establishment of various diseases [1]. In a normal healthy cell, ROS in cells are balanced within the cell by the antioxidant system, which is a variety of enzymatic and non-enzymatic compounds. However, it is also known that in disease or pathological conditions, ROS levels increase, antioxidants are depleted and ultimately oxidative stress causes deterioration. Antioxidant supplementation is important to reduce oxidative stress and therefore related diseases [2-4].

Passiflora belongs to the family of Passifloraceae, comprises approximately 600 species distributed in the tropical and subtropical regions around the world [5-6]. *Passiflora* species occupy an important place in both conventional and modern medicine; they are good source of biologically active substances containing antioxidants and immune modulators with tremendous therapeutic potential [7-9]. Various phytochemical investigations have revealed the presence of secondary metabolite among them phenolics, flavonoids, alkaloids, saponins and terpenes [8-14], for *Passiflora* species.

Moreover, this species is also known to possess many biological activities including anticancer, antioxidant, antiproliferative, sedative, antihypertensive, analgesic [7,10,13,15-17]. Extracts of leaves of different species of *Passiflora* plant are used in traditional medicine for the

treatment of disorders of the nervous system such as migraine and insomnia, anxiolytic, whooping cough, bronchitis and asthma, and also as sedative, anticonvulsant and analgesic [6,8,19].

Passiflora edulis has been studied for its antimicrobial, anticancer and antioxidant potential. However, studies on phytochemical content and antioxidant activity of the plant in solvents with different polarity are limited. Therefore, with this study was conducted for two purposes: (1) to identify and quantify phenolic compounds in the chloroform, methanol and ethyl acetate extracts of *P. edulis* using HPLC-DAD, (2) to evaluate the antioxidant activity of the extracts using 2,2-diphenyl-1-picrylhydrazil (DPPH), FRAP and CUPRAC, total phenolic content and total flavonoid content antioxidant tests.

Materials and Methods

Chemicals and Solvents

The chemicals and reagents, ethanol, 2,2-diphenyl-1-picrylhydrazyl (DPPH), neocuproin, CuCl_2 , 2,4,6-Tris(2-pyridyl)-s-triazine, Iron (III) Chloride Hexahydrate, Acetic Acid Glacial, HCl, Sodium Hydroxide, were purchased from Merck (Darmstadt, Germany) and Aldrich (Milwaukee, WI, USA). All chemicals were used in analytical pure form. HPLC-DAD analyses was performed using HPLC grade phenolic standards and solutions.

Sample Preparation

Passiflora edulis was collected from Antalya-Gazipaşa, in Türkiye the coordinates of 36°16'0.428"N, 32°19'37.674". The collected plant specimens were identified by Vagif Atamov at the Biology Department of RTE University. The plant material (stems, leaves) used as a whole were dried in the shade at room temperature. They were then ground in a blender. Three different solvents were used for extraction. Five grams of the dried sample were put into three flasks and extracted with 50 mL of chloroform, ethyl acetate and methanol in an ultrasonic bath (Heidolph, Germany) at 40 °C for 60 min. Each extract was centrifuged at 10000 rpm, for 10 min. The extracts were transferred into new flasks and evaporated using a rotary evaporator. The dried extracts were dissolved in methanol. The solution was stored at -18 °C until analysed. The extraction procedure was modified according to the method developed by Selvi et al. (2018) [19].

Methanol the most preferred solvents for phenolic compounds because the polyphenols are mostly well soluble and stable in methanol and ethanol [20-23]. While chloroform is preferred for non-polar compounds, ethyl acetate is mostly suitable for flavonoid extraction [24]. Polyphenols in plants have a wide polarity range and the use of solvents with different polarities in extraction is more suitable for the extraction of phenolic compounds. Therefore, extraction was performed using three extraction solvents to elucidate the phenolic content of *P. edulis*.

Determination of Phenolic Compounds by HPLC-DAD

HPLC analyses were performed with Thermo Ultimate 3000 series HPLC system. Chromatographic separation was performed on Agilent C18 column (4.6 × 150 mm, 5 µm) by using a gradient elution at a flow rate of 1 mL/min, the column temperature was 30 °C and injection volume was 20 µL. Gradient elution was used for HPLC analyses using Selvi et al (2024) [25]. Detection wavelengths were set at 280 and 315 nm. Gallic acid, protocatechuic acid, catechin, chlorogenic acid (3-caffeoylquinic acid), caffeine, caffeic acid, vanillic acid, rutin, p-coumaric acid, ferulic acid, o-coumaric acid, quercetin, apigenin, kaempferol, and isorhamnetin were used for phenolic standards.

Determination of Total Phenolic Content (TPC)

TPC of the extracts of *P. edulis* were analysed with Folin-Ciocalteu's phenol reagent. Gallic acid was used to generate a standard curve in a range from 0.0020 and 1.00 mg.mL⁻¹ ($r^2 = 0.999$) [26]. All the experiments were carried out in triplicate and the absorbance of the mixture was measured using a UV-Vis spectrophotometer at 760 nm (Labomed Inc. Culver City-USA). The concentrations of total phenolic compounds were given as mg of gallic acid equivalent (GAE) per g of dry weight (dw).

Determination of Total Flavonoid Contents (TFC)

The aluminum complexation method, as reported by Marcucci et al. (1998) [27], was used to determine the

total flavonoid concentration. Using this procedure, 0.5 mL of plant extract was mixed with 0.1 mL of 10% aluminum nitrate, 0.1 mL of 1 M potassium acetate, and 4.3 mL of 80% ethyl alcohol. After 40 minutes of room temperature incubation, the samples were subjected to UV-Vis (Labomed Inc., Culver City, USA) absorbance measurement at 415 nm. Using quercetin as the standard, a calibration curve in the range of 0.00195 to 0.5 mg.mL⁻¹ ($r^2 = 0.999$) was produced. The total flavonoid content was represented as mg of quercetin equivalent (QE) per g of dry weight (dw), based on the mean of three readings.

Free Radical Scavenging Activity Assay (DPPH)

The scavenging activity of the extracts against the 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical was determined by spectrophotometric method at 517 nm [28]. Extracts were prepared at 2.0 mg/mL concentrations and diluted to different concentrations in methanol in the range of 1.00-0.025 mg/mL. Briefly, 0.75 mL of 0.1 mM DPPH in methanol was added to 0.75 mL of plant extract. Gallic acid and quercetin were used as standards to measure radical scavenging activity. Results were expressed as SC₅₀ values, which indicate the sample concentration required to scavenge 50% of DPPH free radicals (SC₅₀; mg sample per mL methanol). All analyses were performed in triplicate.

Ferric Reducing Antioxidant Power Assay (FRAP)

The FRAP assay was used to assess the extract's antioxidant capability [29]. The FRAP reagent was prepared by mixing 25 mL of acetate buffer (300 mM, pH 3.6), 2.5 mL of 10 mM 2,4,6-tris(2-pyridyl)-s-triazine solution in 40 mM hydrochloric acid, and 2.5 mL of 20 mM iron (III) chloride hexahydrate solutions. Gallic acid was used as a standard ($r^2 = 0.999$). Results were given as µmol gallic acid equivalent per gram of the extract.

Cupric Ion Reducing Antioxidant Capacity (CUPRAC)

The cupric reducing antioxidant capacity of the extracts of *P. edulis* were determined using the method of Apak et al. (2004) [30]. First, 1 mL of Copper (II) chloride, neocuproin and ammonium acetate (1 M, pH=7) taken into a tube, the plant extract was added and the solutions were incubated for 30 minutes at room conditions. After 30 minutes, absorbance was measured at 450 nm. Gallic acid (Sigma Chemical Co) was used as a standard. Results of CUPRAC assay was presented as µmol gallic acid equivalent of one g extract.

Results and Discussion

Identification of Phenolic Compounds in the Extracts

The HPLC-DAD method was used to qualitatively and quantitatively analyse 15 phenolic standards in the extracts of *P. Edulis*. Ferulic acid and p-coumaric acid were determined as primary phenolic compounds in different

amounts in the three extracts. Protocatechuic acid, chlorogenic acid, caffeic acid, apigenin and isorhamnetin were also determined (Table 1). Gallic acid, vanillic acid, rutin, myricetin, quercetin and kaempferol were not detected in any of the plant extracts. The most abundant phenolic content was in the methanol extract, with the highest levels of ferulic acid (Table 1) and *p*-coumaric acid (Figure 1) than the other extracts. Methanol extract was the highest phenolic content among chloroform, ethyl acetate extracts.

Ferulic acid and *p*-coumaric acid, the major phenolic compounds in chloroform, ethyl acetate and methanol extracts of *P. edulis*, were reported as an important component of phenolic compounds with antioxidant properties and the ability to eliminate reactive oxygen [16; 31-33]. Six phenolic compounds such as isoorientin, isovitexin, orientin, vitexin, coumaric acid, rutin and quercetin have previously been determined in ethanolic extracts of *P. edulis* and *P. vitifolia* [13]. In another study, Lourith *et al.* [31] extracted *P. edulis* seeds with 40% methanol: water and fractionated with n-hexane and ethyl acetate. Kojic acid, gallic acid, chlorogenic acid, caffeic acid, ferulic acid, rosmarinic acid and quercetin were used as phenolic standards. Chlorogenic acid, caffeic acid, ferulic acid, and rosmarinic acid were identified in the ethyl acetate fraction. Rotta *et al.* [33] studied phenolic compounds and antioxidant potential of *P. edulis*, *P. alata* and *P. ligularis* fruits. 4-hydroxybenzoic acid, chlorogenic acid, vanillic acid, caffeic acid, *p*-

coumaric acid, ferulic acid, rutin, quercetin and trans-cinnamic acid were found in *P. edulis* fruit. The results were similar to the HPLC-DAD analysis in this work (Table 1). Another study determined phenolic profile of ethyl acetate extracts of fresh fruits of *P. cincinnata* and *P. edulis*. Gallic acid, catechin, epicatechin, epicatechin gallate, epigallocatechin-2,3,4-13C3 gallate, myricetin, rutin, quercetin, resveratrol, caffeic acid, chlorogenic acid, *p*-coumaric acid, and syringic acid were observed in different amounts in the two *Passiflora* species [34].

Table 1. Phenolic composition of *Passiflora edulis* extracts

| Retention Time (Minute) | Standards | Extracts | | |
|-------------------------|-------------------------|-------------------------------|----------------------------------|-----------------------------|
| | | Chloroform (mg std/g extract) | Ethyl acetate (mg std/g extract) | Methanol (mg std/g extract) |
| 6.97 | Protocatechuic acid | 3.24 | 11.83 | 20.88 |
| 7.83 | Chlorogenic acid | 0.21 | 0.48 | 1.32 |
| 9.83 | Caffeic acid | 1.32 | 17.11 | 49.73 |
| 18.54 | <i>p</i> -coumaric acid | 45.87 | 86.30 | 116.88 |
| 19.20 | Ferulic acid | 99.14 | 132.79 | 175.12 |
| 28.51 | Apigenin | 24.61 | 34.77 | 85.60 |
| 29.72 | Isorhamnetin | 2.73 | 7.07 | 32.59 |

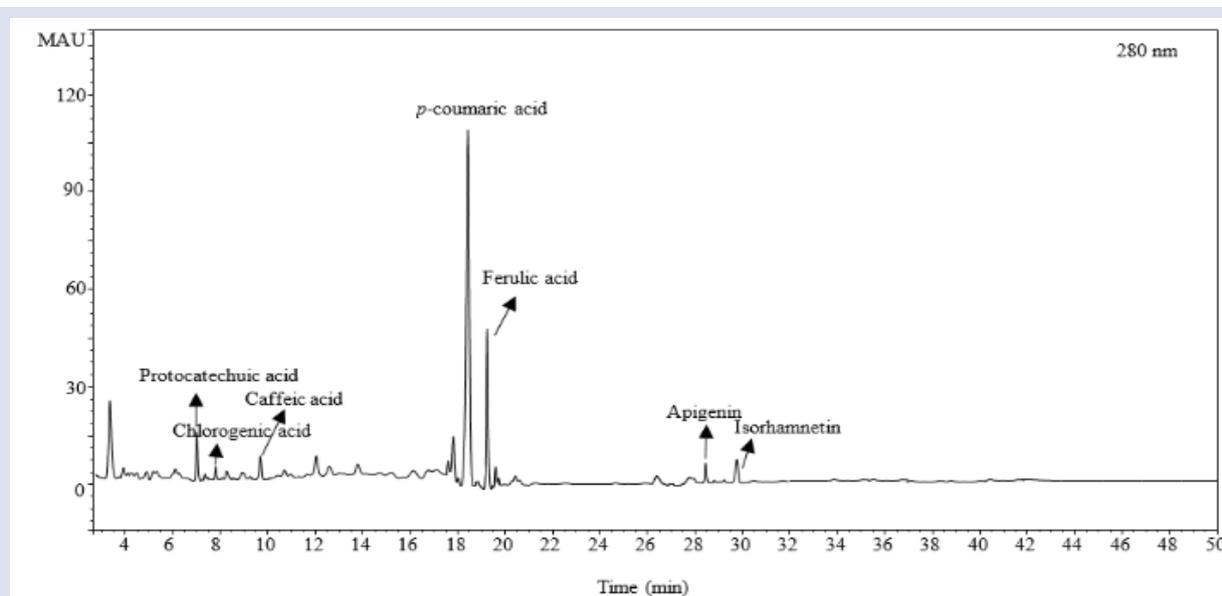


Figure 1. HPLC-DAD chromatogram of methanol extract of *Passiflora edulis*

Total phenolic (TPC) and Total flavonoid content (TFC)

TPC in the extracts of *P. edulis* were determined spectroscopically. According to experimental results, the highest TPC was obtained in the methanol extract of *P. edulis* as 825.09 ± 1.86 and the lowest was obtained in the chloroform extract as 142.57 ± 2.38 mg GAE/g (Table 2).

Total flavonoid content measured with aluminum chloride reagent and quercetin was used as a standard. TFC were in the extracts of *P. edulis* are presented in Table 2. According to the results of study, the highest TFC was obtained in the methanol extract as 502.78 ± 0.90 and the lowest was obtained in the chloroform extract as 106.76 ± 1.22 mg QE/g.

Table 2. Total phenolic and flavonoid contents of the extracts from *Passiflora edulis**

| | Chloroform | Ethyl acetate | Methanol |
|---------------|---------------|---------------|---------------|
| TPC (mgGAE/g) | 142.57 ± 2.38 | 382.74 ± 0.95 | 825.09 ± 1.86 |
| TFC (mgQE/g) | 106.76 ± 1.22 | 377.78 ± 0.52 | 502.78 ± 0.90 |

*: Results were given as $\bar{x} \pm$ standard deviation

Lourith *et al.* [31] investigated total phenolic contents of ethyl acetate and water fractions of *P. edulis* seeds. The total phenolic contents were 58.3 g GAE/100 g for ethyl acetate fraction and 2.7 g GAE/100 g for water fraction, respectively. Colomeu *et al.* [35] investigated methanol, ethanol and aqueous extracts of four *Passiflora* spp. They reported that TPC in those plants ranges from 0.023 to 0.228 mg GAE/g. Santos *et al.* [34] have used two Brazilian

passion fruit species: *P. cincinnata* and *P. edulis*. Gallic acid used as the standard. The results of their study showed that the highest TPC was calculated *P. edulis* (476.1 mg GAE/kg) followed by that from *P. cincinnata* (365 mg GAE/kg). Another study reported the total phenolic and total flavonoid contents of ethanol extracts of *P. vitifolia* and *P. edulis* seeds. Results of that study showed the ethanol extract of *P. edulis* seeds were rich in total phenolic and total flavonoids [13].

Antioxidant Activity

In order to determine antioxidant activity, three extracts were analysed using radical scavenging (DPPH), reducing power (FRAP, CUPRAC) assays. Results of antioxidant assay were presented in Table 3.

Table 3. Antioxidant activities of the extracts from *Passiflora edulis**

| Extracts | CUPRAC $\mu\text{molGA/g}$ | FRAP $\mu\text{molGA/g}$ | DPPH, SC50 mg/mL |
|---------------|-------------------------------|-----------------------------|---------------------|
| Methanol | 3.728 ± 0.150 | 0.511 ± 0.012 | 0.028 ± 0.001 |
| Ethyl acetate | 0.783 ± 0.051 | 0.358 ± 0.012 | 0.036 ± 0.002 |
| Chloroform | 0.161 ± 0.028 | 0.333 ± 0.018 | 0.345 ± 0.017 |
| Gallic acid | - | - | 0.004 ± 0.000 |
| Quercetin | - | - | 0.003 ± 0.000 |

*: Results were given as $\bar{x} \pm$ standard deviation

Lourith *et al.* [31] investigated antioxidant activity of *P. edulis* seed obtained from the residue after fruit juice production. Seeds of *P. edulis* was extracted with methanolic water and fractionated with n-hexane and ethyl acetate. Also, DPPH, FRAP and ABTS antioxidant tests were used. The ethyl acetate fraction showed the most potent antioxidant activity than other fractions. Additionally, the antioxidant activity of ethyl acetate extracts of *P. edulis* and *P. cincinnata* fruits was confirmed by Santos *et al.* [34] using DPPH, ABTS and FRAP methods. According to the results of their study, the fruits of *P. edulis* extract was higher than that of the *P. cincinnata*. Another study [32] have investigated antioxidant activity of aqueous, ethanol and methanol/acetone extracts of *P. edulis* fruits by using an ABTS, DPPH and FRAP methods. As a result of the study, the aqueous and ethanol extracts showed higher antioxidant activity when compared to methanol/acetone extract.

Conclusions

In this study, chloroform, ethyl acetate and methanol extracts of *Passiflora edulis* were investigated for the phenolic composition and antioxidant activities. Methanol extract of *P. edulis* was found to be the richest in total phenolic content and total flavonoid content. These findings were also supported by the results of HPLC-DAD analysis. For the phytochemical analysis, the methanol extract had the highest phenolic content among the chloroform, ethyl acetate extracts. Similarly, DPPH, FRAP and CUPRAC antioxidant tests showed that methanol extract had the highest antioxidant activity. In accordance with the results of the HPLC-DAD analysis, the extracts contain significant amounts of *p*-coumaric acid and ferulic acid, which are known to have potential biological activity.

This study provided an important contribution to the existing knowledge on secondary metabolites of *P. edulis*. In conclusion, the fact that the plant has high antioxidant activity and is a good source of phenolic compounds may be encouraging in terms of expanding the biological properties of the plant. These studies may enable the plant to be used as a therapeutic agent in the future.

Conflict of interests

The author declare that she has no conflicts of interest.

Acknowledgments

The author would like to thank Erciyes University, Drug Application and Research Center (ERFARMA) for laboratory and device support.

References

- [1] Sevindik M., Mohammed F.S., Uysal I., Autism: Plants With Neuro-Psychopharmacotherapeutic Potential, *Prospect. Pharm. Sci.*, 21(3) (2023) 38-48.
- [2] El-Chaghaby G.A., Mohammed F.S., Rashad S., Uysal I., Koçer O., Lekeşiz Ö., Sevindik M., Genus *Hypericum*: General Properties, Chemical Contents and Biological Activities, *Egypt. J. Bot.*, 64(1) (2024) 1-26.
- [3] Moloney J.N., Cotter T.G., ROS Signalling in the Biology of Cancer, *Semin Cell Dev. Biol.*, 80 (2018) 50–64.
- [4] Khan A.Q., Rashid K., AlAmodi A.A., Agha M.V., Akhtar S., Hakeem I., Raza S.S., Uddin S., Reactive Oxygen Species (ROS) in Cancer Pathogenesis and Therapy: An Update on the Role of ROS in Anticancer Action of Benzophenanthridine Alkaloids, *Biomed. Pharmacother.*, 143 (2021) 112142.
- [5] Ayres A.S., Araujo A.L., Soares T.C., *et al.* Comparative Central Effects of the Aqueous Leaf Extract of Two

- Populations of *Passiflora edulis*, *Rev. Bras. Farmacogn.*, 25 (2015) 499-505.
- [6] Wosch L., Santos K.C., Imig D.C., Santos C.A.M., Comparative Study of *Passiflora* Taxa Leaves: II. A Chromatographic Profile, *Rev. Bras. farmacogn.*, 27 (2017) 40-49.
- [7] Dhavan K., Dhavan S., Sharma A., *Passiflora*: a Review Update, *J. Ethnopharmacol.*, 94 (2004) 1-23.
- [8] Ingale S.P., Kasture S.B., Antioxidant and Antiparkinsonian Activity of *Passiflora incarnata* Leaves, *Orient Pharm. Exp. Med.*, 14(3) (2014) 231-6.
- [9] Uzunoğlu F., Özmen K., Toprak S., et al. The Effect of Different Pre-sowing Treatments on Seedling Emergence, Quality and Development in *Passiflora edulis* Seeds, *Erwerbs-Obstbau.*, 65 (2023) 2509-2516.
- [10] Alves J.S., Marques J.I., Demarque D.P., Costa L.R., Amaral J.G., Lopes N.P., da Silva-Júnior A.A., Soares L.A.L., Gavioli E.C., Ferreira L.S., Zucolotto S.M., Involvement of Isoorientin in the Antidepressant Bioactivity of a Flavonoid-rich Extract From *Passiflora edulis* f. *flavicarpa* Leaves, *Rev. Bras. Farmacogn.*, 30 (2020) 1-11.
- [11] Alvarez-Rivera G., Ballesteros-Vivas D., Parada-Alfonso F., Ibañez E., Cifuentes A., Recent Applications of High Resolution Mass Spectrometry for the Characterization of Plant Natural Products, *TrAC Trends Anal. Chem.*, 112 (2019) 87-101.
- [12] Ballesteros-Vivas D., Alvarez-Rivera G., Ibáñez E., Parada-Alfonso F., Cifuentes A., Integrated Strategy for the Extraction and Profiling of Bioactive Metabolites from *Passiflora mollissima* seeds Combining Pressurized-liquid Extraction and Gas/Liquid Chromatography-high Resolution Mass Spectrometry, *J. Chromatogr. A.*, 1595 (2019) 144-57.
- [13] Rodríguez A.J., Arteaga J.M., Arango W.M., Pabon M.F., Vasodilator Effect of Ethanolic Extracts of *Passiflora vitifolia* and *Passiflora edulis* f. *edulis* Seeds, *J. Appl. Pharm. Sci.*, 11(10) (2021) 61-69.
- [14] Viganó J., de Paula Assis B.F., Náthia-Neves G., dos Santos P., Meireles M.A.A., Veggi P.C., Martínez J., Extraction of Bioactive Compounds from Defatted Passion Fruit Bagasse (*Passiflora edulis* sp.) Applying Pressurized Liquids Assisted by Ultrasound, *Ultrason Sonochem.*, 64 (2020) 104999.
- [15] Aguillón J., Maldonado M., Loango N., Arango S., Landázuri P., Actividad Antioxidante Antiproliferativa de Extractos Etanólicos y Acuoso de las Hojas y el Jugo del Fruto de *Passiflora edulis*, *Perspect Nutr. Hum.*, 15 (1) (2013) 13.
- [16] Silva J.K., Cazarin C.B., Batista Â.G., Maróstica M., Effects of Passion Fruit (*Passiflora edulis*) Byproduct Intake in Antioxidant Status of Wistar Rats Tissues, *LWT Food Sci. Technol.*, 59 (2) (2014) 1213-1219.
- [17] Villada Ramos J.A., Aguillón Osma J., Restrepo Cortes B., Loango Chamarro N., Maldonado Celis M.E., Identification of Potential Bioactive Compounds of *Passiflora edulis* Leaf Extract Against Colon Adenocarcinoma Cells, *Biochem. Biophys. Rep.*, 28 (2023) 34-101453.
- [18] Zibadi S., Watson R.R., Passion Fruit (*Passiflora edulis*) Composition, Efficacy and Safety, *J. Evid. Based Integr. Med.*, 1 (2024) 183-187.
- [19] Selvi E.K., Turumtay H., Demir A., Turumtay E.A., Phytochemical Profiling and Evaluation of the Hepatoprotective Effect of *Cuscuta campestris* by High-performance Liquid Chromatography with Diode Array Detection, *Anal. Lett.*, 51 (10) (2018) 1464-1478.
- [20] Yao L., Jiang Y., Datta N., Singanusong R., Liu X., Duan J., Raymont K., Lisle A., Xu Y., HPLC Analyses of Flavanols and Phenolic Acids in the Fresh Young Shoots of Tea (*Camellia sinensis*) Grown in Australia, *Food Chem.*, 84 (2004) 253-63.
- [21] Dmitrienko S.G., Kudrinskaya V.A., Apyari V.V., Methods of Extraction, Preconcentration, and Determination of Quercetin, *J. Anal. Chem.*, 67(4) (2012) 299-311.
- [22] Mokrani A., Madani K., Effect of Solvent, Time and Temperature on the Extraction of Phenolic Compounds and Antioxidant Capacity of Peach (*Prunus persica* L.) Fruit, *Sep. Purif. Technol.*, 162 (2016) 68-76.
- [23] Sukeksi L., Sarah M., Characterizations and Extraction of Polyphenols from Residual Pulp of Pink guava as Source of Antioxidants, *J. Appl. Eng. Sci.*, 11 (2016) 5209-16.
- [24] Stalikas C.D. Extraction, Separation, and Detection Methods for Phenolic acids and Flavonoids, *J. Sep. Sci.*, 30 (2007) 3268-3295. doi: 10.1002/jssc.200700261.
- [25] Selvi E.K., Güler N., Güven S., Makbul S., Determination of Chemical Compositions, Antioxidant, Dna Cleavage and Binding Properties of *Vincetoxicum tmoaleum* Extract, *Studia UBB Chemia*, 1 (2024) 163-174.
- [26] Singleton V.L., Orthofer R., Lamuela R.M., Analysis of Total Phenols and Other Oxidation Substrates and Antioxidants by Means of Folin-ciocalteu Reagent, *Methods Enzymol.*, 299 (1999) 152-178.
- [27] Marcucci M., Woisky R., Salatino A., Uso de Cloreto de Alumínio na Quantificação de Flavonóides em Amostras de Própolis, *Mensagem Doce.*, 46 (1998) 3-9.
- [28] Molyneux P., The Use of the Stable Free Radical Diphenylpicrylhydrazyl (DPPH) for Estimating Antioxidant Activity, *SJST.*, 26 (2) (2004) 211-19.
- [29] Benzie I.F.F., Strain J.J., Ferric Reducing/Antioxidant Power Assay: Direct Measure of Total Antioxidant Activity of Biological Fluids and Modified Version for Simultaneous Measurement of Total Antioxidant Power and Ascorbic Acid Concentration, *Meth. Enzymol.*, 299 (1999) 15-27.
- [30] Apak R., Guclu K., Ozyurek M., Karademir SE., Novel Total Antioxidant Capacity Index for Dietary Polyphenols and Vitamins C and E, Using Their Cupricionreducing Capability in the Presence of Neocuproine: CUPRAC method, *J. Agric. Food Chem.*, 52 (26) (2004) 7970-7981.
- [31] Lourith N., Kanlayavattanakul M., Antioxidant Activities and Phenolics of *Passiflora edulis* Seed Recovered From Juice Production Residue, *J. Oleo Sci.*, 62(4) (2013) 235-240.
- [32] Sova M., Sosa L., Natural Sources, Pharmacokinetics, Biological Activities and Health Benefits of Hydroxycinnamic Acids and Their Metabolites, *Nutrients*, 12(8) (2020) 2190.
- [33] Rotta E.M., Rodrigues C.A., Jardim I., Maldaner L., Visentainer J.V., Determination of Phenolic Compounds and Antioxidant Activity in Passion Fruit Pulp (*Passiflora* spp.) Using a modified QuEChERS Method and UHPLC-MS/MS, *LWT.*, 100 (2019) 397-403.
- [34] Santos T.B., Araujo F.P., Neto A.F., Freitas S.T., Souza A.J., Vilar O.S.B., Araújo A.J.B., Lima M.S., Phytochemical Compounds and Antioxidant Activity of the Pulp of Two Brazilian Passion Fruit Species: *Passiflora cincinnata* Mast. And *Passiflora edulis* Sims, *Int. J. Fruit Sci.*, 21(1) (2021) 255-269.
- [35] Colomeu T.C., Figueiredo D., Zollner R.L., Comparison of Antioxidant and Antiproliferative Effect Among Four *Passiflora* Spp, *JALS.*, 4 (2) (2017) 1-

A Comparative study of DNA Alignment Algorithms and Boosting Performance Using Different Compilation Strategies

Osman Doluca^{1,a,*}

¹ Department of Biomedical Engineering, Izmir University of Economics, Türkiye

*Corresponding author

Research Article

History

Received: 06/06/2024

Accepted: 03/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

With the development of next generation sequencing technologies, the requirement of higher performance from DNA and Protein sequence alignment algorithms has become even greater. This work is a systematic comparison of different compilation strategies for two common DNA or Protein sequence alignment algorithms, Needleman-Wunsch and Smith-Waterman, using Python programming language. It aims to investigate the performance benefits of already widely used Biopython's pairwise alignment module versus different compilation approaches of an in-house software. It is shown that using Numba just-in-time compiler provide greater performance overall in comparison to PyPy and Cython compilers or the Biopython module. This work may increase the efficiency of software prototyping where large-scale sequence alignment is necessary.

Keywords: DNA/Protein sequence alignment, Python, Just-in-time compilers.

 osman.dolucas@ieu.edu.tr

 <https://orcid.org/0000-0003-0412-6148>

Introduction

With the development of next generation sequencing (NGS) the amount of available biological data, in terms of genome or exome sequences, has been dramatically increasing. The alignment methods have been becoming more and more relevant as the necessity to process the raw NGS data or to investigate the similarities between sequences for the discovery of their functional properties has been increasing. As the current NGS sequencing technologies are producing relatively short fragments of sequence information, ranging between 50 to 1000 bp, their alignment is quite computation exhausting process since a single stretch of continuous DNA sequence may be up to hundreds of millions bp long. [1] While the NGS technologies improve the count of the fragments read, the reading length tend to get even shorter, relying even more to the computation for their alignment to produce the genome [2] with the exception of nanopore sequencing which is yet to become applicable. [3]

Any solution to this problem has one of two approaches. One being, increasing the computational power, has long been market-driven endeavor. Unfortunately, as the Moore's law is approaching its limits and a stagnation in the increase of the processing power of chips is imminent, any hardware-based solution is shifting towards multi-core chips instead of faster chips. [4] The other solution is implementing alternative algorithms to the alignment problem. The two most commonly used algorithms, Needleman Wunsch (global) [5] and Smith Waterman (local) [6], and their variations [7] are known to provide the best alignment, but also the slowest as their implementations make use of multiple 2D matrices. New heuristic approaches are also developed

however these approaches does not match the sensitivity of these two algorithms or their variations. [8-10]

New approaches for processing biological data often requires software prototyping and testing, and the majority of the data scientists rely on high-level programming languages such as Python. The less time spent coding and high readability of such high-level languages is the foremost reason behind this preference. [11] In comparison, while the low-level programming languages such as C provide better the execution speed, the high complexity deters the scientists from using it. However, especially when it comes to NGS data analysis, speed may be more of an issue than many other factors due to sheer size of alignments required. The solution is often sought in accelerating the alignment algorithms. [12-14] Speed improvement is possible at the interpreter level, either through using modules that exploit pre-compiled libraries or compilation of the code on the go, also known as just-in-time compilation. [15] There has been a number of approaches for improving the performance of SW algorithm, mostly through parallelism by performing calculations on FPGAs or GPU. [16-18] However, these are not available as modules for mainstream programming languages.

Here we implement the global and local alignment algorithms including the affine gap extension developed by Gotoh et al. [5-7] We compare the performance improvements when rewritten using Numpy module, or compiled using Cython, Numba or PyPy. The performance improvements were also compared with pairwise2 module of Biopython library. The code is made available online at <http://github.com/odolucas/Fast-NW-and-SW-Pairwise-alignment-using-numba-JIT/>

Method

The affine gap penalty [7] variation of original global [5] and local [6] alignment algorithms was written using python 3.6+. For each local and global algorithms two other variations were written, one of which uses single scores for matches and mismatches or a substitution matrix. Substitution matrix is especially necessary for protein alignments as transition between amino acids do not have equal probabilities. Each of the four methods were rewritten in two forms; one for discovery of only the highest score while the other uses backtracking to reports the best (or one of the best) alignment(s). Together these variations yielded eight different methods. (Table 1)

Each method was interpreted or compiled using different approaches. 1) Pure python approach uses only python 3.6+ syntax and built-in types. No external module was used. 2) Numpy approach incorporates ndarrays from Numpy module (1.14.2) for all matrices. 3) Cython (v0.29.12) was used to compile pure python or Numpy-using code using cythonize method and "build_ext" argument to build all extensions. 4) Just-In-Time compilation using Numba module (v0.44.1) was used with both pure python or with Numpy. Because the Numba does not accept strings in "nopython" mode, all methods were written to accept two lists of integers as sequences to be aligned instead of lists of characters, as in strings, where each character indicates nucleotide or amino acid residues. An additional method is written to convert any protein, DNA or RNA sequence into a list of integers. All numba methods were compiled just-in-time in "nopython" and "cached" modes for optimum performance using "@jit(nopython=True, cached=True)" decoration. 5) finally, a 32-bit pypy compiler (v7.1.1) was used with pure python code to compare. All algorithms were written in two variations, "score only" and "backtrack". The first one is where only scores are calculated, and the latter is where the best alignment is constructed by tracing back the path. Backtracking requires three additional matrices to keep track of the path. Additionally, Biopython's pairwise2 alignment was used for comparison to view the performance improvements. Biopython was also tested in "score only" mode as well as "alignment" mode for equivalent comparison.

Performance was measured using timeit module, aligning sequences with varying percentages of similarity and varying sizes of sequences. For each pairwise alignment, a sequence was generated randomly at first. The other sequence was obtained by introducing a number of mutations until a given percent similarity is obtained. Each mutation was introduced with 80% chance for the substitution, 10% for the insertions and deletions each. The percent similarity was calculated as a ratio of global alignment score of the alignment of the two sequences to the alignment of max possible sequence of any two sequence with the same lengths. The global alignment was performed using +1 for matches, -1 for mismatches and gaps. For each category a thousand

sequence pairs were aligned and total processing times were found as summation. All tests were performed using timeit module with garbage collection off to increase accuracy. All module imports or any preprocessing is left out of performance testing and measurements were done only during sequence alignments. Every performance test was repeated five times and the best of five was reported. All tests were performed at AMD 1950X machine equipped with 128 GB ECC RAM with ECC-mode on and locked to the same core. All algorithms were previously run using random sequences and compared with Biopython's pairwise2 module to confirm that the same results were produced before performance testing.

Results and Discussion

With varying features included in the algorithm eight different methods were written and tested in this work. The list of these methods and their features are listed in Table 1. Briefly substitution matrix feature enables different penalties for substitution between different residues. This is preferred especially if the mutations between particular residues is more common or expected than others. Another feature is called "backtracking" which enables production of a final alignment of the two sequences. Alternative, "score only" mode reports only the score of the best alignment which may be used as a measure of sequence similarity. This is often useful for construction of phylogenetic trees. Backtracking requires keeping of three additional matrices with a size of (n x m) with n and m being the lengths of the two sequences.

Table 1. Methods used in this work and their abbreviations.

| algorithm | substitution matrix | backtracking | method abbreviation |
|-----------|---------------------|--------------|---------------------|
| global | No | No | globalms |
| local | No | No | localms |
| global | Yes | No | globalds |
| local | Yes | No | localds |
| global | No | Yes | globalms |
| local | No | Yes | localms |
| global | Yes | Yes | globalds |
| local | Yes | Yes | localds |

The methods were written and executed with varying modules and compilers. Not all compilers were compatible, as a result, we have tested seven combinations of modules and compilers/interpreters. (Table 2., Figure 1.) As Cython or standard Python interpreter proved to be much slower in all cases, their performance evaluation is omitted, focusing on Biopython, PyPy and Numba.

The effect of sequence length. The biggest impact on the performance was observed to be the sequence length. In all the cases there was an exponential increase, close to the order of two as the size of the matrices (n x m) also

increases in the order of two. This was independent of the algorithm or modules that were used. Interestingly only with Numba the order of the power was close to 1.5 which

indicates that the matrix processing is not a bottleneck for Numba.

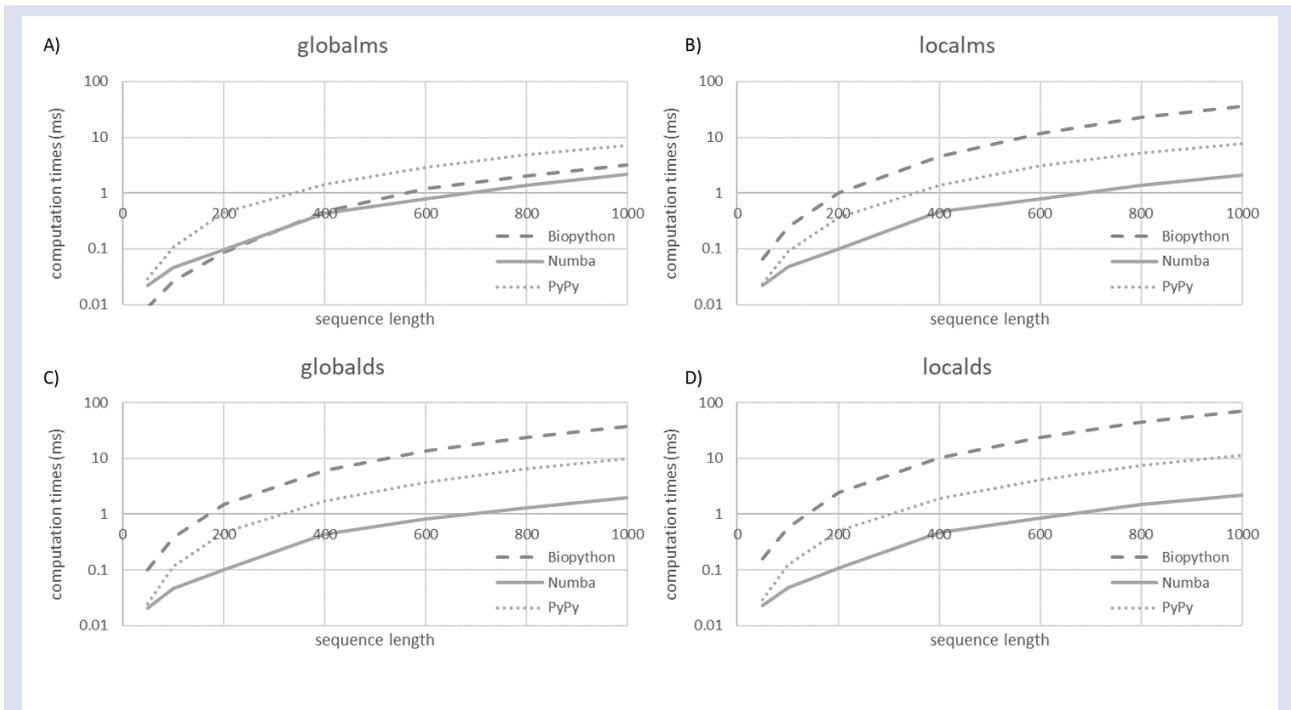


Figure 1. The computation times in milliseconds (ms) versus sequence length for globalms (A), localms (B), globalds (C) and localds (D) alignment methods without backtracking, using various compilation methods, Biopython using standard Python (dashed), In-house method with Numba compiler (light gray) and In-house method with PyPy compiler (dotted).

The effect of sequence similarity. On the other hand sequence similarities ranging from 20% to 80% did not seem to have an effect on the computation time. (Data not shown) In most cases the differences were too small to consider significant. Although some difference was expected between methods using backtrack to produce an alignment since the final alignments would be longer when there is low similarity due to increased amounts of gap.

“Score only” versus “backtracking” modes. When a comparison between “score only” and “backtracking” modes, there has been small variations in the computation time in spite of the choice of the compiler/interpreter since backtracking requires three additional matrices to be constructed. Between “score_only” modes and “backtracking” modes of Biopython there is a huge impact on the performance. This impact drops down for larger sequences for all methods with Biopython. On the other hand, with Numba and PyPy the impact is significantly less and ranging only between 1.5 and 2. The difference is mostly related to the way the Biopython’s pairwise2 module is executed rather than the choice of the compiler/interpreter. Local vs global. No significant difference was observed between local and global algorithms. The only exception was Biopython methods where the difference varies between 1.3 to 10 times. In case of Numpy or PyPy no significant difference was apparent.

Substitution matrix choice. Oddly, Numpy exploiting methods when compiled with Cython proved to be even slower when using methods that include a substitution matrix. Considering that these methods are generally less efficient than Cython compiled pure Python code, an overall Cython Numpy uncomplimentary was apparent.

Between the methods a dramatic difference was observed when Biopython module was used. Only for Biopython, while globalms method was fastest, localds and localds with backtracking were significantly slower than their counterparts.

Using standard Python interpreter, Biopython showed greater overall performance in comparison to Numpy module or pure Python. However, with the introduction of an alternative compiler performance dramatically improves. When overall performance is considered, Numba assisted just-in-time compilation proves to be optimal in almost all cases. On average, Numba provided 15 times faster computation time in “score only” mode and 18 times faster computation time in “backtracking” mode than Biopython. Biopython showed better performance only for globalms method when aligning sequences shorter than ~200 base pairs. At the same time, PyPy achieved greater performance only for localms method and for sequences of a length of ~50 bp. Comparison of PyPy and BioPython showed that PyPy was around 3.5 times faster on average in “score only” mode while comparable in “backtracking” mode.

Table 2. Computation times in milliseconds of various sequence alignment methods, globalms, localms, globalds and localds, with and without backtracking, using different compilation strategies. All methods were tested with varying sequence lengths and 20% sequence similarity. Best performances of each series of sequence lengths are reported in bold.

| module | method: sequence length: interpreter/compiler | globalms without backtracking | | | | | | |
|-----------|---|-------------------------------|-------|--------|---------|--------|--------|---------|
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 0.009 | 0.026 | 0.087 | 0.470 | 1.212 | 2.048 | 3.225 |
| numpy | Numba | 0.022 | 0.046 | 0.096 | 0.433 | 0.795 | 1.363 | 2.159 |
| - | PyPy | 0.029 | 0.108 | 0.449 | 1.433 | 2.918 | 4.887 | 7.194 |
| numpy | Cython | 0.139 | 0.553 | 2.236 | 10.120 | | | |
| - | Cython | 0.631 | 2.475 | 9.878 | 40.824 | | | |
| numpy | standard Python 3.6+ | 1.001 | 3.916 | 15.898 | | | | |
| - | standard Python 3.6+ | 0.475 | 1.980 | 7.548 | | | | |
| module | method: interpreter/compiler | localms without backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 0.065 | 0.248 | 1.019 | 4.583 | 11.708 | 22.448 | 36.291 |
| numpy | Numba | 0.022 | 0.047 | 0.098 | 0.465 | 0.793 | 1.383 | 2.125 |
| - | PyPy | 0.024 | 0.090 | 0.367 | 1.396 | 3.102 | 5.203 | 7.766 |
| numpy | Cython | 0.145 | 0.569 | 2.258 | 10.025 | | | |
| - | Cython | 0.748 | 2.883 | 11.559 | 47.241 | | | |
| numpy | standard Python 3.6+ | 1.149 | 4.614 | 17.805 | | | | |
| - | standard Python 3.6+ | 0.494 | 1.958 | 7.694 | | | | |
| module | method: interpreter/compiler | globalds without backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 0.100 | 0.381 | 1.501 | 6.082 | 13.618 | 24.038 | 37.705 |
| numpy | Numba | 0.021 | 0.047 | 0.100 | 0.445 | 0.827 | 1.318 | 1.949 |
| - | PyPy | 0.025 | 0.114 | 0.466 | 1.715 | 3.757 | 6.488 | 9.768 |
| numpy | Cython | 0.459 | 1.825 | 7.354 | 30.445 | | | |
| - | Cython | 0.602 | 2.362 | 9.399 | 38.945 | | | |
| numpy | standard Python 3.6+ | 0.998 | 3.967 | 15.507 | | | | |
| - | standard Python 3.6+ | 0.879 | 3.314 | 13.300 | | | | |
| module | method: interpreter/compiler | localds without backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 0.157 | 0.588 | 2.481 | 10.386 | 24.193 | 44.655 | 70.487 |
| numpy | Numba | 0.023 | 0.048 | 0.107 | 0.466 | 0.865 | 1.475 | 2.205 |
| - | PyPy | 0.029 | 0.124 | 0.511 | 1.928 | 4.086 | 7.518 | 11.400 |
| numpy | Cython | 0.520 | 2.089 | 8.506 | 35.492 | | | |
| - | Cython | 0.735 | 2.853 | 11.417 | 46.367 | | | |
| numpy | standard Python 3.6+ | 1.172 | 4.542 | 18.131 | | | | |
| - | standard Python 3.6+ | 0.962 | 3.784 | 14.622 | | | | |
| module | method: interpreter/compiler | globalms with backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 1.761 | 2.190 | 4.022 | 7.949 | 9.164 | 18.490 | 20.475 |
| numpy | Numba | 0.023 | 0.050 | 0.177 | 0.681 | 1.573 | 2.667 | 4.226 |
| - | PyPy | 0.058 | 0.214 | 0.840 | 3.495 | 7.218 | 12.612 | 18.929 |
| numpy | Cython | 0.317 | 1.253 | 4.951 | 22.217 | | | |
| - | Cython | 1.622 | 6.361 | 25.222 | 103.363 | | | |
| numpy | standard Python 3.6+ | 2.308 | 9.126 | 36.224 | | | | |
| - | standard Python 3.6+ | 0.969 | 3.581 | 14.521 | | | | |
| module | method: interpreter/compiler | localms with backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 0.169 | 0.616 | 2.496 | 10.177 | 24.115 | 43.812 | 69.278 |
| numpy | Numba | 0.023 | 0.049 | 0.126 | 0.684 | 1.612 | 2.799 | 4.327 |
| - | PyPy | 0.050 | 0.190 | 0.728 | 3.042 | 6.380 | 11.616 | 17.473 |
| numpy | Cython | 0.293 | 1.148 | 4.554 | 20.728 | | | |
| - | Cython | 1.653 | 6.310 | 25.439 | 102.680 | | | |
| numpy | standard Python 3.6+ | 2.366 | 9.056 | 36.411 | | | | |
| - | standard Python 3.6+ | 0.903 | 3.558 | 13.804 | | | | |
| module | method: interpreter/compiler | globalds with backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 1.399 | 2.377 | 3.429 | 8.613 | 16.305 | 27.459 | 41.009 |
| numpy | Numba | 0.026 | 0.058 | 0.158 | 0.867 | 1.850 | 3.252 | 4.909 |
| - | PyPy | | | | | | | |
| numpy | Cython | 0.962 | 3.837 | 15.267 | 63.191 | | | |
| - | Cython | 1.522 | 5.965 | 23.866 | 96.943 | | | |
| numpy | standard Python 3.6+ | 2.329 | 8.859 | 35.338 | | | | |
| - | standard Python 3.6+ | 1.640 | 6.794 | 26.157 | | | | |
| module | method: interpreter/compiler | localds with backtracking | | | | | | |
| | | 50 | 100 | 200 | 400 | 600 | 800 | 1000 |
| Biopython | standard Python 3.6+ | 1.444 | 3.545 | 6.627 | 18.518 | 38.599 | 68.369 | 105.336 |
| numpy | Numba | 0.027 | 0.064 | 0.180 | 0.945 | 2.028 | 3.541 | 5.458 |
| - | PyPy | | | | | | | |
| numpy | Cython | 1.018 | 3.999 | 16.154 | 68.273 | | | |
| - | Cython | 1.679 | 6.457 | 25.711 | | | | |
| numpy | standard Python 3.6+ | 2.392 | 9.485 | 37.360 | | | | |
| - | standard Python 3.6+ | 1.670 | 6.669 | 27.091 | | | | |

Conclusion

While Python provide a versatile development environment for prototyping new ideas for data scientists, it remains to be one of slower executed programming languages. With the development of next generation sequencing technology and the increase in the amount data to be processed, massive alignment problems such as de novo genome construction or multiple sequence alignments (MSA) arise and bioinformaticians find themselves needing faster solutions for testing their ideas that require sequence alignments. Here, in order to boost the performance of these alignment algorithms, the two most commonly used alignment algorithms were written using Python language and tested for various compilation strategies. Our findings show that the choice of compiler may have significant impact on the speed of execution. The overall optimal approach was found to be using just-in-time Numba compiler while followed by PyPy just-in-time compiler. Biopython also proved to be a decent option, considering certain methods at certain sequence lengths it may give equivalent performance. In conclusion, it is important to consider compilers, no matter the choice of the compiler is, as they would give up to 200 times higher performance than standard Python interpreter.

On the other hand, further improvements may still be possible if compilation strategy is combined with parallelism. However, the overhead cost of setting up parallelism needs to be considered depending on the number and length of the query sequences and the available hardware, such as GPUs. [16-18] For that reason, parallelism might not be a feasible solution to improve the performance in all situations. None the less, the compilation strategy ensures better performance in any Python environment given that necessary modules are installed and independent of the hardware setting.

Conflicts of interest

There are no conflicts of interest in this work.

Acknowledgment

This work was funded by TÜBİTAK project no 118Z226.

References

- [1] Zhang J., Chiodini R., Badr A., Zhang G., The impact of next-generation sequencing on genomics, *Journal of Genetics and Genomics*, 38 (2011) 95-109.
- [2] McPherson J.D., Next-generation gap, *Nature Methods*, 6 (2019) S2-S5.
- [3] Branton D., Deamer D. The Development of Nanopore Sequencing, *Nanopore Sequencing*, (2019) 1-16
- [4] Theis T.N., Wong P.H.S., The End of Moore's Law: A New Beginning for Information Technology, *Computing in Science & Engineering*, 19 (2017) 41-50.
- [5] Needleman S.B., Wunsch C.D., A general method applicable to the search for similarities in the amino acid sequence of two proteins, *Journal of Molecular Biology*, 48 (1970) 443-453.
- [6] Smith T.F., Waterman M.S., Identification of common molecular subsequences, *Journal of Molecular Biology*, 147 (1981) 195-197.
- [7] Gotoh O., An improved algorithm for matching biological sequences, *Journal of Molecular Biology*, 162 (1982) 705-708.
- [8] Marco-Sola, S., Moure, J. C., Moreto, M., Espinosa, A., Fast Gap-Affine Pairwise Alignment Using the Wavefront Algorithm, *Bioinformatics*, 37 (2020) 456-463.
- [9] Song Y.-J., Ji D. J., Seo H., Han G.-B., Cho D.-H., Pairwise Heuristic Sequence Alignment Algorithm Based on Deep Reinforcement Learning, *IEEE Open Journal of Engineering in Medicine and Biology*, 2 (2021) 36-43.
- [10] Rashed A. E. E.-D., Amer H. M., El-Seddek M., Moustafa H. E.-D., Sequence Alignment Using Machine Learning-Based Needleman-Wunsch Algorithm, *IEEE Access*, 9 (2021) 109522-109535.
- [11] Nagpal A., Gabrani G., Python for Data Analytics, Scientific and Technical Applications, 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, (2019).
- [12] Mondal S., Khatua S., Accelerating Pairwise Sequence Alignment Algorithm by MapReduce Technique for Next-Generation Sequencing (NGS) Data Analysis, *Advances in Intelligent Systems and Computing*, (2019) 213-220.
- [13] Marçais G., Delcher A.L., Phillippy A.M., Coston R., Salzberg S.L., Zimin A., MUMmer4: A fast and versatile genome alignment system, *PLoS Computational Biology*, 14 (2018) e1005944.
- [14] Tarasov A., Vilella A.J., Cuppen E., Nijman I.J., Prins P., Sambamba: fast processing of NGS alignment formats, *Bioinformatics*, 31 (2015) 2032-2034.
- [15] Marowka A., Python accelerators for high-performance computing, *The Journal of Supercomputing*, 74 (2018) 1449-1460.
- [16] Haghi A., Marco-Sola S., Alvarez L., Diamantopoulos D., Hagleitner C., Moreto M., An FPGA Accelerator of the Wavefront Algorithm for Genomics Pairwise Alignment, 31st International Conference on Field-Programmable Logic and Applications (FPL), Dresden, (2021).
- [17] Rognes, T., Faster Smith-Waterman database searches with inter-sequence SIMD parallelization, *BMC Bioinformatics*, 12 (2011) 1.
- [18] Liu Y., Maskell D. L., Schmidt, B., CUDASW++: optimizing Smith-Waterman sequence database searches for CUDA-enabled graphics processing units. *BMC Research Notes*, 2, 1 (2009) 73.

Inflammatory and Biochemical Concepts in Children with Specific Learning Disorders: A Comparative Study

Cansu Mercan Isik ^{1,a,*}, Ayla Uzun Cicek ^{1,b}, Melih Ugur Karaarslan ^{1,c}, Serdar Akkus ^{2,d}, Dilara Ülger Özbek ^{3,e}

¹ Department of Child and Adolescent Psychiatry, Cumhuriyet University Faculty of Medicine, Sivas, Türkiye

² Department of Child and Adolescent Psychiatry, Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Türkiye

³ Department of Biochemistry, Cumhuriyet University Faculty of Medicine, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 24/06/2024

Accepted: 04/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This study aimed to investigate the role of inflammation and biochemical parameters in children with Specific Learning Disorders (SLD) and to evaluate these parameters based on the severity of SLD. The study was planned as a retrospective. 39 children diagnosed with SLD and 32 healthy controls aged 6-16 years, who had hemogram and biochemistry tests performed at their admission, were included in the study. Diagnoses were based on DSM-5 criteria, clinical interviews, family interviews, and assessments of reading, writing, and math skills. The study received ethical approval from the Cumhuriyet University Non-Invasive Clinical Research Ethics Committee. Lymphocyte count and WBC were significantly higher in the SLD group ($p = 0.003$, $p = 0.006$, respectively). A significant difference was detected between the groups regarding platelet/lymphocyte ratio (PLR) ($p = 0.047$). No significant differences were found in blood parameters (erythrocyte, neutrophil, lymphocyte, WBC, platelet, T4, TSH, folic acid, vitamin B12, PLR, neutrophil/lymphocyte ratio (NLR)) when evaluated according to the severity of SLD. The results of our study suggest that inflammation may play a role in SLD, but further research with larger sample sizes, longitudinal designs, and comprehensive assessments of inflammatory markers is needed to better understand these associations.

Keywords: SLD, PLR, NLR, Inflammation, Biochemical markers.

^a dr.cansumercan@gmail.com

^b <https://orcid.org/0000-0001-9437-3024>

^c melihugurkaraarslan@gmail.com

^d <https://orcid.org/0009-0006-8810-0109>

^e dilaraulger@cumhuriyet.edu.tr

^e <https://orcid.org/0000-0002-2189-6876>

^b dr.f.ayla@hotmail.com

^d <https://orcid.org/0000-0003-2274-3457>

^d serdar.akkus93@gmail.com

^e <https://orcid.org/0000-0002-9978-6861>

Introduction

Neurodevelopmental disorders constitute a heterogeneous group of inherited medical conditions that primarily impact social communication, language, attention, impulsivity, learning, perception, and motor coordination. The condition has a detrimental impact on both the individual and the family, persisting from childhood to adulthood [1]. There is mounting evidence that inflammatory processes are involved in the etiology of neurodevelopmental disorders, including autism, specific learning disorders, attention deficit hyperactivity disorder, emotional disorders such as depression and bipolar disorder, and tic disorders. Neuroinflammation has been linked to alterations in brain development, including modifications to synaptic plasticity and synaptogenesis [2]. Specific learning disorder (SLD) is a neurodevelopmental disorder that is characterized by a failure to achieve the expected level of proficiency in reading, mathematics, and written expression skills when considering the individual's chronological age, educational level, and intelligence. Despite its prevalence in childhood, the etiology of SLD remains unclear [3]. Proposed mechanisms include neuronal damage and degeneration, increased oxidative stress, decreased neurotrophic support, glial activation, changes in neurotransmitter metabolism, and disruption of the blood-brain barrier [4]. Consequently, the potential contribution of inflammation to neurodevelopmental disorders is being subjected to further scrutiny. The

hematological system is a vital organ for human immune defense and plays a central role in the inflammatory process and sepsis [5]. Leukocytes, endothelium, platelets (PLTs), and numerous other components are responsible for the activation of the immune system. An increase in the inflammatory response triggered by environmental factors in the early stages of development may result in structural and/or functional alterations in brain development, which may subsequently lead to the onset of neurodevelopmental disorders such as specific learning disabilities (SLD) and attention deficit hyperactivity disorder (ADHD) [6]. The neutrophil/lymphocyte ratio (NLR), mean platelet volume (MPV), monocyte/lymphocyte ratio (MLR), and red blood cell distribution width (RDW) are frequently employed as straightforward peripheral inflammation markers that can reflect the underlying state of systemic inflammation and can be readily and rapidly measured in whole blood [10]. The utilization of these cells as markers has facilitated the identification of alterations in inflammatory status across a spectrum of psychiatric disorders [11-16]. These psychiatric disorders include schizophrenia, bipolar disorder, autism spectrum disorders, and attention deficit hyperactivity disorder [4, 7-9]. Furthermore, vitamins, minerals, and hormones are implicated in the etiology, disease process, and treatment of psychiatric disorders [10]. It is assumed that B12, folic acid and thyroid hormones exert a direct effect on cognitive functions,

especially on the energy metabolism of neurons and glial cells, the synthesis of neurotransmitters and receptor binding. [11].

A review of the literature reveals a paucity of studies examining the relationship between SLD and inflammation. Nevertheless, no studies have examined the role of SLD and inflammation and biochemical parameters together, nor have they evaluated these factors according to SLD severity. In light of these findings, the objective of our study was to examine the role of SLD and inflammation and biochemical parameters in patients with specific learning disorders, evaluating them according to SLD severity.

Materials and Methods

The records of the child and adolescent mental health and diseases outpatient clinic between May 2023 and May 2024 were retrospectively examined and a diagnosis of SLD was made as a result of a clinical interview based on DSM-5, family interview, psychiatric examination including assessment of reading, writing and mathematics skills appropriate for the age and grade level, and teacher information form evaluation. Thirty-nine children aged between 6 and 16 years with no accompanying psychiatric, neurological and medical diseases, no mental retardation and no mental retardation, who underwent haemogram and biochemistry examinations, and 32 age- and gender-matched healthy children who were not diagnosed at their first presentation to child and adolescent mental health and diseases and who underwent haemogram and biochemistry examinations were included as the control group. Routine blood tests were requested from patients at the first application in order to organize diagnosis and treatment decisions and to make a differential diagnosis. Another purpose was to see the current metabolic status and to observe the side effects of drugs if any medication was to be started. The severity of SLD was determined as a result of a clinical interview based on DSM-5. Sociodemographic

characteristics registered in the system were used. The minimum sample size was determined by power analysis. The study was started after approval was obtained from the Cumhuriyet University Non-interventional Clinical Research Ethics Committee (dated 16.05.2024, numbered 2024/05-14). When $\beta = 0.10$ and $\alpha = 0.05$ $1 - \beta = 0.90$ was taken in the study, it was decided to include 39 individuals in the patient group and 32 individuals in the control group. The power of the test was found as $p = 0.90498$. Hormone and vitamin results; Roche Cobas 8000 system (Roche Diagnostics, Mannheim, Germany). Complete Blood Count (CBC); CBC MINDRAY 5200 (China) was analyzed in the hospital biochemistry laboratory.

Statistical Methods

The data obtained from our study were evaluated with SPSS 23.0 program. The normality of the data was checked with Kolmogorov-Smirnov test. Since the data provided parametric conditions, it was analyzed with independent sample t test for two independent groups and F test (ANOVA) for more than two groups. When ANOVA was used in more than two group comparisons, Tukey's T2 test was used for those that provided the homogeneity assumption and Tamhane's T2 test was used for those that did not provide the homogeneity assumption to determine which group was different from the others. Chi-square test was used for categorical variables. The error level will be taken as 0.05.

Results

Demographic Characteristics

This study was conducted with 39 children with SLD and 32 healthy control groups. Table 1 shows the socio-demographic and clinical characteristics of the SLD groups and controls. There was no significant difference between the groups in terms of age, age groups, gender, place of residence, family income level, and mother and father's education level (all p values >0.05) (Table 1).

Table 1. Sociodemographic characteristics of the sample

| | SLD group (N=39) | Control group (N=32) | p-value* |
|---|---------------------|-------------------------|----------|
| Age (mean-years±SD) | 10.41±2.67 | 11.03±2.86 | 0.348 |
| Gender (n,%) | | | 0.934 |
| Male | 24 (61.5) | 20 (62.5) | |
| Female | 15 (38.5) | 12 (37.5) | |
| Place of residence (n,%) | | | 0.994 |
| Urban | 28 (71.8) | 23 (71.9) | |
| Rural | 11 (28.2) | 9 (28.1) | |
| Family income level (n,%) [†] | | | 0.712 |
| The minimum wage/less than minimum wage | 20 (51.3) | 15 (46.9) | |
| Above the minimum wage | 19 (48.7) | 17 (53.1) | |
| Education level of mother (n,%) | | | 0.292 |
| Primary education and lower | 30 (76.9) | 21 (65.6) | |
| Upper primary education | 9 (23.1) | 11 (34.4) | |
| Education level of father (n,%) | | | 0.621 |
| Primary education and lower | 23 (59.0) | 17 (53.1) | |
| Upper primary education | 16 (41.0) | 15 (46.9) | |

*The chi-square test for categorical variables and the Independent-samples t-test for continuous variables were used to test group differences.

Bold font indicates statistical significance: $p < 0.05$

[†]The level of income was determined by the minimum wage value on the date of the study.

Abbreviations: SD, Standard Deviation; SLD, Specific Learning Disorder.

Comparison of Hemogram and Biochemical Parameters Between Two Groups

Table 2. Comparison of hemogram and biochemical parameters between two groups.

| | SLD group (N=39) | Control group (N=32) | p-value* |
|---|---------------------|-------------------------|--------------|
| Erythrocyte (10 ³ /uL) (mean±SD) | 4,98±0,48 | 5,00±0,41 | 0,957 |
| Neutrophil (10 ³ /uL) (mean±SD) | 4,17±1,68 | 3,64±1,12 | 0,127 |
| Lymphocyte (10 ³ /uL) (mean±SD) | 2,96±0,95 | 2,32±0,72 | <i>0,003</i> |
| WBC (10 ³ /uL) (mean±SD) | 7,91±2,37 | 6,54±1,41 | <i>0,006</i> |
| Platelet(10 ³ /uL) (mean±SD) | 337,21±67,86 | 308,72±66,33 | 0,080 |
| T4 (mcg/dl) (mean±SD) | 1,28±0,16 | 1,29±0,35 | 0,879 |
| TSH (mIU/L) (mean±SD) | 2,31±2,42 | 2,42±1,58 | 0,729 |
| Folic acid(ng/ml) (mean±SD) | 9,31±3,17 | 8,37±2,14 | 0,160 |
| Vitamin B12 Levels(pg/ml) (mean±SD) | 415,46±158,89 | 366,25±163,02 | 0,204 |
| NLR | 1,48±0,58 | 1,74±0,95 | 0,157 |
| PLR | 121,96±39,14 | 142,79±47,48 | <i>0,047</i> |

Independent-samples t-test for continuous variables were used to test group differences.
 Italic font indicates statistical significance: $p < 0.05$
 Abbreviations: SD, Standard Deviation; SLD, Specific Learning Disorder; TSH, Thyroid Stimulating Hormone; WBC, White Blood Cell; NLR, Neutrophil lymphocyte ratio; PLR, Platelet lymphocyte ratio.

*Independent-samples t-test

When blood parameters were evaluated, no significant difference was found between the groups in terms of erythrocyte, neutrophil, platelet, T4, TSH, folic acid, vitamin B12, and NLR values. Lymphocyte count and WBC were significantly higher in the SLD group than in the control group ($p = 0.003$, $p = 0.006$, respectively). A significant difference was detected between the groups regarding PLR ($p = 0.047$). General characteristics and biochemical values of the groups are shown in Table 2. (Table 2)

Comparison of Parameters according to SLD Severity

When blood parameters were evaluated according to the severity of SLD, no significant difference was found between the groups in terms of erythrocyte, neutrophil, lymphocyte, WBC, platelet, T4, TSH, folic acid, vitamin B12, PLR, and NLR values. General characteristics and biochemical values of the groups are shown in Table 3. (Table 3)

Table 3. Comparison of parameters according to SLD severity

| | Mild(n=7) | Moderate(n=13) | Severe (n=19) | p-value |
|--|---------------|----------------|---------------|---------|
| Erythrocyte(10 ³ /uL) (mean±SD) | 5,25±0,19 | 5,15±0,56 | 4,79±0,43 | 0,006 |
| Neutrophil(10 ³ /uL) (mean±SD) | 4,68±2,35 | 4,10±1,49 | 4,03±1,58 | 0,681 |
| Lymphocyte(10 ³ /uL) (mean±SD) | 3,42±1,14 | 2,70±1,07 | 2,97±0,76 | 0,281 |
| WBC (10 ³ /uL) (mean±SD) | 8,96±3,33 | 7,53±2,57 | 7,78±1,80 | 0,428 |
| Platelet (10 ³ /uL) (mean±SD) | 379,71±78,19 | 325,62±64,12 | 329,47±63,88 | 0,188 |
| T4 (mcg/dl) (mean±SD) | 1,33±0,08 | 1,31±0,23 | 1,25±0,13 | 0,425 |
| TSH (mIU/L) (mean±SD) | 2,67±1,34 | 2,10±0,45 | 2,32±1,11 | 0,481 |
| Folic acid (ng/ml) (mean±SD) | 9,74±1,82 | 10,26±4,53 | 8,49±2,23 | 0,286 |
| Vitamin B12 Levels (pg/ml) (mean±SD) | 445,57±209,86 | 412,38±161,618 | 406,47±144,08 | 0,860 |
| NLR | 1.34±0.35 | 1.64±0.62 | 1.43±0.62 | 0.478 |
| PLR | 120.54±47.42 | 132.70±48.05 | 115.14±28.52 | 0.470 |

One-way ANOVA for continuous variables were used to test group differences.
 Bold font indicates statistical significance: $p < 0.05$
 Abbreviations: SD, Standard Deviation; SLD, Specific Learning Disorder; TSH, Thyroid Stimulating Hormone; WBC, White Blood Cell; NLR, Neutrophil lymphocyte ratio; PLR, Platelet lymphocyte ratio.

Discussion

The present study sought to compare inflammatory markers and biochemical parameters between a specific learning disorder (SLD) group and a control group. The findings indicated that no notable discrepancy was observed between the two groups concerning erythrocyte, neutrophil, platelet, T4, TSH, folic acid,

vitamin B12, and NLR values. Nevertheless, the lymphocyte count and WBC values were observed to be markedly elevated in the SLD group in comparison to the control group. Additionally, a notable discrepancy was observed between the two groups about PLR.

The literature review highlighted a limited number of studies examining B12 and folic acid levels in individuals with SLD. These studies largely corroborate our findings.

For instance, Esnafoğlu's study reported no significant difference in vitamin B12 levels between SLD and control groups, although folic acid levels were notably lower in the SLD group [12]. Similarly, a 2016 study investigating B12 and zinc levels found no significant difference in B12 levels [13]. Deficiencies in vitamin B12 and folate have been associated with learning difficulties, psychosomatic symptoms, and anxiety in children with attention-deficit/hyperactivity disorder (ADHD) [14, 15]. In some cases, children with asymptomatic deficiencies in these nutrients may present with inattention or developmental delays, warranting an evaluation for potential neurodevelopmental disorders. Consequently, clinicians should consider assessing micronutrient levels when there is clinical suspicion of such deficiencies. The varying results observed in SLD, despite its classification as a neurodevelopmental disorder, may stem from differences in socioeconomic status and nutritional practices within the studied population.

Another parameter analyzed in this study was thyroid function. Thyroid hormones, including free triiodothyronine (fT3), free thyroxine (fT4), and thyroid-stimulating hormone (TSH), play a vital role in regulating motor, cognitive, and emotional processes. Extensive evidence underscores the adverse effects of maternal thyroid dysfunction on fetal brain development [17]. However, a review of the literature revealed no previous studies investigating thyroid function specifically in the context of SLD. In this study, thyroid function levels did not differ significantly between the SLD and control groups, potentially due to sample selection criteria or the limited sample size. Nonetheless, further research is necessary to clarify the role of thyroid function in brain development and its potential implications for SLD.

Systemic inflammation has been proposed as a potential contributor to the etiology of psychiatric disorders, including autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD). However, the precise nature of the relationship between neurodevelopmental disorders and systemic inflammation remains uncertain [18]. In our study, an increase in white blood cell (WBC) and lymphocyte levels may serve as an indicator of chronic inflammation in individuals with learning disabilities. Chronic inflammation, particularly in neurodevelopmental disorders, can result in a state of constant stress in the brain and body. It is established that chronic inflammation has a deleterious impact on synaptic plasticity and neurodevelopment in the brain [19]. This may contribute to the etiology of cognitive disorders such as learning disabilities.

Secondly, individuals with learning disabilities may exhibit disturbances in the regulation of the immune system. An increase in white blood cell (WBC) and lymphocyte levels may indicate that the body's immune system is hyperactive. Such outcomes may be attributable to environmental factors, including infections or maternal immune activation, to which children with learning disabilities are exposed, particularly during prenatal or

early childhood. Maternal immune activation may result in the triggering of inflammatory responses in the fetus, which may subsequently give rise to neurodevelopmental consequences [20].

Also, a low platelet-to-lymphocyte ratio (PLR) may signify an alternative aspect of the immune response in individuals with learning disabilities. A low PLR indicates a reduction in platelet count relative to lymphocyte count, which may contribute to an increased complexity of the inflammatory state. Consequently, low platelet counts may result in a weakened or altered inflammatory response, which can occur in several ways [21]. A low PLR may provide insights into the general impact of the immune system in individuals with learning disabilities, and the effects of this condition on cognitive functions warrant further investigation.

In the study conducted by Bilac and colleagues, a total of 64 participants were included, comprising 31 children with SLD and 33 healthy children. The study yielded findings indicating that the patient group exhibited elevated neutrophil and lymphocyte counts and NLR values in comparison to the control group. PLR values did not differ significantly between the two groups. Nevertheless, regression analyses revealed that the NLR value, age, and gender, which were identified as independent variables, did not exert a significant influence on SLD [22]. In a further study conducted in 2023, no significant difference was observed between the two groups in terms of hemoglobin (Hb), red cell distribution width (RDW), platelet crit (PCT), platelet distribution width (PDW), white blood cell (WBC) count, neutrophil-to-lymphocyte ratio (NLR) and thrombocyte count (TLO) levels. The platelet count was observed to be higher in the SLD group. This study aimed to investigate the relationship between SLD types and to determine whether a significant difference exists. The results demonstrated that no significant difference was found [23]. Unlike earlier research, our study categorized SLD by severity but found no significant differences in blood parameters across severity levels.

While our findings provide valuable insights, the study has limitations, including a small sample size, potential confounding factors such as nutrition and physical activity, and a lack of direct measures of neuroinflammation. Future longitudinal studies with larger samples are needed to clarify the role of inflammation and immune system changes in SLD and their potential impact on cognitive development.

Conclusion

In conclusion, elevated white blood cell (WBC) and lymphocyte levels, along with a low platelet-to-lymphocyte ratio (PLR), may reflect chronic inflammation and immune system dysregulation in individuals with learning disabilities. These findings offer valuable insights into the potential link between immune responses and neurodevelopmental processes in this population. To build on this foundation, future research should address

current limitations by utilizing larger sample sizes, adopting longitudinal study designs, and accounting for potential confounding factors. Furthermore, investigating direct markers of neuroinflammation and exploring the mechanisms behind these hematological changes in SLD is essential. A deeper understanding of these processes could pave the way for targeted interventions aimed at reducing the impact of inflammation on cognitive and learning outcomes in children with SLD.

Conflict of interest

The author declares no conflict of interest, financial or otherwise.

References

- [1] Jalenques I., Auclair C., Morand D., Legrand G., Marcheix M., Ramanoel C., Health-related quality of life, anxiety, and depression in parents of adolescents with Gilles de la Tourette syndrome: A controlled study, *Eur. Child Adolesc. Psychiatry*, 26 (5) (2017) 603-617.
- [2] Öncü T., An immunological perspective on neurodevelopmental disorders, *Curr. Perspect. Health Sci.*, 2 (2) (2021) 44-50.
- [3] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed., American Psychiatric Association, Washington, DC, 2013.
- [4] Leffa D.T., Torres I.L.S., Rohde L.A., A review on the role of inflammation in attention-deficit/hyperactivity disorder, *Neuroimmunomodulation*, 25 (5-6) (2018) 328-333.
- [5] Anand D., Colpo G.D., Zeni G., Zeni C.P., Teixeira A.L., Attention-deficit/hyperactivity disorder and inflammation: What does current knowledge tell us? A systematic review, *Front. Psychiatry*, 8 (2017) 228.
- [6] Dunn G.A., Nigg J.T., Sullivan E.L., Neuroinflammation as a risk factor for attention deficit hyperactivity disorder, *Pharmacol. Biochem. Behav.*, 182 (2019) 22-34.
- [7] Tural Hesapcioglu S., Kasak M., Citak Kurt A.N., Ceylan M.F., High monocyte level and low lymphocyte to monocyte ratio in autism spectrum disorders, *Int. J. Dev. Disabil.*, 65 (2) (2019) 73-81.
- [8] Karageorgiou V., Milas G.P., Michopoulos I., Neutrophil-to-lymphocyte ratio in schizophrenia: A systematic review and meta-analysis, *Schizophr. Res.*, 206 (2019) 4-12.
- [9] Bustan Y., Drapisz A., Ben Dor D.H., Avrahami M., Schwartz-Lifshitz M., Weizman A., Elevated neutrophil to lymphocyte ratio in non-affective psychotic adolescent inpatients: Evidence for an early association between inflammation and psychosis, *Psychiatry Res.*, 262 (2018) 149-153.
- [10] Örum M.H., Kara M.Z., Eğılmez O.B., Determination of vitamin B12, folate, and ferritin levels of inpatients in a psychiatry clinic: A one-year retrospective study, *Istanbul Bilim Univ. Florence Nightingale J. Med.*, 4 (2) (2018) 71-78.
- [11] Li C., Lai J., Huang T., Han Y., Du Y., Xu Y., Thyroid functions in patients with bipolar disorder and the impact of quetiapine monotherapy: A retrospective, naturalistic study, *Neuropsychiatr. Dis. Treat.*, 15 (2019) 2285-2290.
- [12] Esnafoğlu E., Serum folate, vitamin B12, homocysteine and vitamin D levels in children with specific learning disorder, *Bozok Tıp Derg.*, 8 (3) (2018) 59-64.
- [13] Taşkın B.D., Karalök Z.S., Aydoğmuş Ü., Gürkaş E., Yoldaş M., Güven A., Zinc and B12 vitamin levels in children with specific learning disorder, *J. Pediatr. Res.*, 3 (4) (2016) 187-190.
- [14] Altun H., Sahin N., Belge Kurutas E., Güngör O., Homocysteine, pyridoxine, folate and vitamin B12 levels in children with attention deficit hyperactivity disorder, *Psychiatr. Danub.*, 30 (3) (2018) 310-316.
- [15] Unal D., Çelebi F., Bildik H.N., Koyuncu A., Karahan S., Vitamin B12, and hemoglobin levels may be related with ADHD symptoms: A study in Turkish children with ADHD, *Psychiatry Clin. Psychopharmacol.*, 29 (4) (2019) 515-519.
- [16] Topal Z., Tufan A. E., Karadag M., Gokcen C., Akkaya C., Sarp A. S., Kilinc M., Evaluation of peripheral inflammatory markers, serum B12, folate, ferritin levels and clinical correlations in children with autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD), *Nordic journal of psychiatry*, 76(2) (2022) 150-157.
- [17] Chen, G., Gao, W., Xu, Y., Chen, H., & Cai, H., Serum TSH levels are Associated with hyperactivity behaviors in children with attention Deficit/Hyperactivity disorder. *Neuropsychiatric Disease and Treatment*, (2023) 557-564.
- [18] Chen X., Yao T., Cai J., Fu X., Li H., Wu J., Systemic inflammatory regulators and 7 major psychiatric disorders: A two-sample Mendelian randomization study, *Prog. Neuropsychopharmacol. Biol. Psychiatry*, 116 (2022) 110534.
- [19] Mitchell R.H., Goldstein B.I., Inflammation in children and adolescents with neuropsychiatric disorders: A systematic review, *J. Am. Acad. Child Adolesc. Psychiatry*, 53 (3) (2014) 274-296.
- [20] Voltas N., Arijia V., Hernández-Martínez C., Jiménez-Feijoo R., Ferré N., Canals J., Are there early inflammatory biomarkers that affect neurodevelopment in infancy?, *J. Neuroimmunol.*, 305 (2017) 42-50.
- [21] Sayit A.T., Gunbey P.H., Terzi Y., Is the mean platelet volume in patients with acute cholecystitis an inflammatory marker?, *J. Clin. Diagn. Res.*, 9 (6) (2015) TC05-TC07.
- [22] Bilaç, Ö., Tahillioğlu, H. A., Önder, A., & Kavrurma, C., Özgül öğrenme güçlüğünde enflamasyonun rolünün değerlendirilmesi, *Türk J Child Adolescent Mental Health*, 28 (1) (2021) 60-64.
- [23] Öz B., Ardiç B., The role of inflammation in children with specific learning disorders, *Ann. Indian Psychiatry*, 7 (2) (2023) 96.

Determination of Acidity Constant (pKa), Lipophilic Partition (LogP) and Distribution Coefficients (LogD) Values of Some 3-(2-Phenylethyl)-Tetrahydro-2H-1,3,5-Thiadiazine-2-Thione-5-Acetic Acid Derivatives

Gürkan Özen ^{1,a,*}, Erhan Palaska ^{2,b}, Emirhan Nemutlu ^{3,c}¹ Department of Analytical Chemistry, Faculty of Pharmacy, Baskent University, Ankara, Türkiye² Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Hacettepe University, Ankara, Türkiye³ Department of Analytical Chemistry, Faculty of Pharmacy, Hacettepe University, Ankara, Türkiye

*Corresponding author

Research Article

History

Received: 26/08/2024

Accepted: 11/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Invasive fungal infections pose a significant threat to human health, emphasizing the urgent need for new and effective antifungal agents. 3,5-Disubstituted-tetrahydro-2H-1,3,5-thiadiazine-2-thione derivatives have shown promising antifungal activity. This study focused on characterizing the physicochemical properties, specifically the acidity constant (pKa), lipophilic partition coefficient (logP), and distribution coefficient (logD), of four 3-(2-phenylethyl)-tetrahydro-2H-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives (EP1, EP2, EP3, and EP4) with varying α -carbon substituents. pKa values were determined using two independent methods: parallel factor analysis (PARAFAC) of UV spectroscopic data and reversed-phase high-performance liquid chromatography (RP-HPLC). Lipophilicity was assessed by measuring logD values using the shake-flask method with n-octanol, and logP values were calculated based on the determined pKa and logD values. The pKa values determined by both PARAFAC and RP-HPLC showed a good correlation. The α -carbon substituent significantly influenced pKa, with electron-withdrawing substituents resulting in lower pKa values, consistent with inductive effects. The logD-pH profiles exhibited typical bell-shaped curves, with logP values at pH 5.5 ranging from -0.38 to 3.00. EP1, EP2, and EP4 displayed lipophilic characteristics, while EP3 showed higher hydrophilicity. This study provides the first reported pKa, logP, and logD values for these novel thiadiazine derivatives. The findings emphasize the influence of structural modifications on physicochemical properties, which are critical for drug absorption, distribution, and target interaction. These data provide a valuable basis for elucidation of the structure-activity relationship and formulation optimization of these antifungal agents.

Keywords: Acidity constant, Lipophilic partition, Distribution coefficients, PARAFAC, HPLC.gurkanozen@baskent.edu.tr
enemutlu@hacettepe.edu.tr<https://orcid.org/0000-0003-0777-6402>
<https://orcid.org/0000-0002-7337-6215>epalaska@hacettepe.edu.tr<https://orcid.org/0000-0002-8914-8858>

Introduction

Invasive fungal infections present a serious and growing global health threat, causing an estimated 1.5 million deaths annually [1]. Despite medical progress in treating bacterial infections, the limited availability of effective antifungal drugs poses a significant challenge. Individuals with weakened immune systems, such as those undergoing chemotherapy, organ transplantation, or living with HIV/AIDS, are particularly susceptible to these infections, leading to severe consequences [2]. The emergence of antifungal resistance further complicates treatment efforts and highlights the urgent need for novel antifungal drug discovery [3].

The 3,5-disubstituted-tetrahydro-2H-1,3,5-thiadiazine-2-thione scaffold represents a promising source of novel therapeutic agents. Previous research has demonstrated the diverse biological activities of compounds containing this core structure, including antibacterial, antifungal, anti-inflammatory, antileishmanial, antitubercular, and anticancer properties [4-10]. Among these, 3-(2-phenylethyl)-tetrahydro-2H-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives, synthesized via a well-established

two-step process, have emerged as particularly promising candidates.

The journey of a drug molecule from the laboratory bench to a patient's bedside is a complex one, influenced by a delicate interplay of factors. Among these, the molecule's inherent physicochemical properties stand out as critical determinants of its fate and effectiveness within the body. These properties, including its acidity constant (pKa), lipophilic partition coefficient (logP), and distribution coefficient (logD), govern how a drug interacts with biological systems, ultimately influencing its absorption, distribution, metabolism, excretion, and overall therapeutic efficacy.

A drug molecule's degree of ionization, directly related to its pKa value, significantly impacts its behavior and efficacy. The pKa dictates the ratio of ionized to non-ionized forms at a specific pH, influencing crucial factors like solubility, cell membrane permeability, binding to plasma proteins, and interactions with target molecules [11]. While not the sole determinant of efficacy, understanding how a drug's ionization changes with pH is essential for optimizing its delivery to fungal cells,

ensuring effective target engagement, and predicting its overall pharmacokinetic profile. Therefore, pKa is a critical consideration throughout the design and development of new antifungal agents.

Understanding the logP and logD values provides further insights into a drug's behavior. While logP, a measure of the molecule's affinity for a lipophilic environment, offers valuable information, it only paints a partial picture. In contrast, logD, which takes into account the ionization state of the molecule across a range of pH values, provides a more comprehensive view of its lipophilic character in biological systems [12]. This becomes particularly crucial when considering that most

drugs exist as a mixture of ionized and non-ionized forms under physiological conditions.

PARAFAC has proven to be a valuable tool for pKa determination, with successful applications in analyzing UV [13-16], voltammetric [17, 18] and fluorescence [19] data.

To determine the pKa of these molecules, we employed two independent methods, including PARAFAC, which was applied to UV spectroscopic data and HPLC. PARAFAC, a powerful three-way analysis method (Figure 1), dissects the complex pH-dependent UV absorbance data, allowing us to discern the equilibrium between the acidic (HA) and basic (A⁻) forms of the molecule and precisely pinpoint its pKa value (Figure 2) [20, 21].

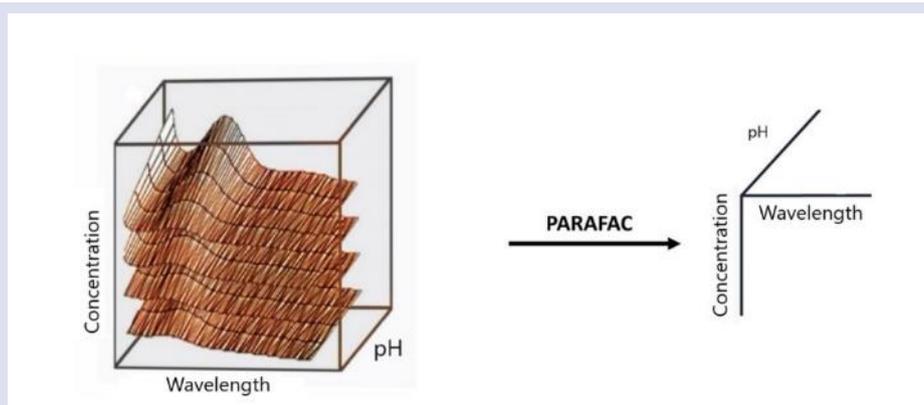


Figure 1. Datasets generated to determine pKa by PARAFAC three-way data analysis.

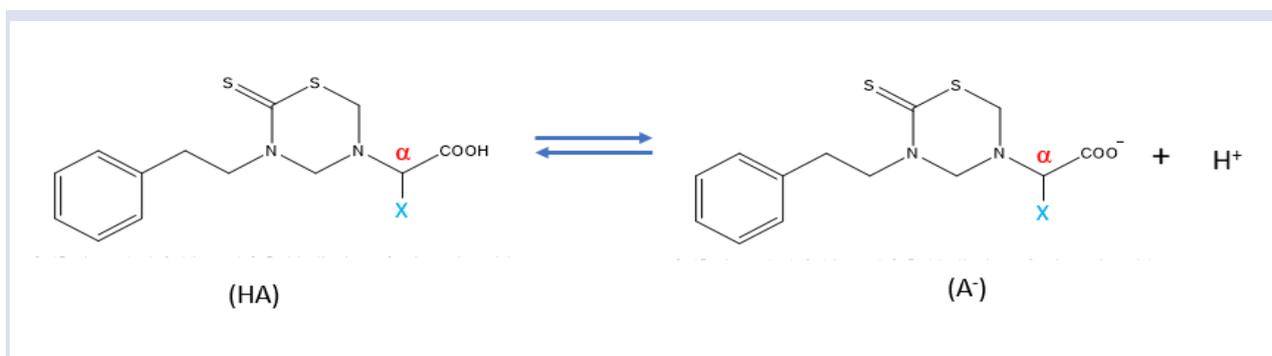


Figure 2. Acid dissociation equilibrium of EPO molecule.

According to Equation 1, when the acidic form concentration [HA] and the base form concentration [A⁻] are equal, the pH of the solution becomes equal to the pKa value of the relevant substance. Analytical signals can be decomposed with three-way analysis tools by arranging them as three-way data strings. This distinction helps us to observe the profiles of the three variants for all the components (HA and A⁻ concentrations) that contribute to the signals.

$$K_a = \frac{[H^+][A^-]}{[HA]} \quad (1)$$

In additionally, in this study, the pKa values of some 3-(2-phenylethyl)-tetrahydro-2h-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives were determined by the RP-HPLC method using acetonitrile (ACN)-water mobile phase. In

RP- HPLC method, acidic form (HA) and basic form (A⁻) of ionizable substances have different retention times. In the RP-HPLC method, the retention times of both species are expressed by the capacity factor (k') (Equation 2).

$$k' = \frac{t_r - t_0}{t_0} \quad (2)$$

In the equation, k' is the capacity factor, t_r is the retention time of the substance, and t₀ is the dead time.

Determining the pH value of the mobile phase is very important in pKa determination studies with HPLC. Different buffers and organic solvents (acetonitrile, methanol, etc.) or mixtures of these in mobile phases with certain ratios can be used for pKa determination in the HPLC method. In this case, the pH value of the aqueous phase and the pH value of the mobile phase with an

organic modifier may differ. For this reason, different pH scales in HPLC studies can be used. Although there are many different pH scales, three of them are widely used in pH scaling. The first of these is the aqueous pH scale (${}^w\text{pH}$); In this scale, the electrode system is calibrated with aqueous buffers and pH is measured in the aqueous buffers before adding organic modifiers. The second case is the pH scale relative to the solvent (${}^s\text{pH}$); In this scaling, the electrode system is calibrated with aqueous buffers, but the pH measurement is made in the mobile phase obtained by mixing the aqueous buffers with the organic modifier. In this case, the pH value will differ from the standard aqueous pH scale. The third is the pH scale in the mobile phase (${}^m\text{pH}$); The electrode system is calibrated with buffers prepared in the same mixed solvent used in the mobile phase and pH measurement is performed in the mobile phase. In this case, the pH value differs from the standard aqueous pH scale and also differs from the pH scale depending on the state of the solvent [22].

Determination of pKa using the k' and pH values of the mobile phase in liquid chromatography was first suggested by Horvat [23]. Later, further studies were performed on different compounds to measure the pKa values from the capacity factors [24, 25]. In this study, pKa values of EP1, EP2, EP3 and EP4 pharmaceuticals were determined with the help of k' -pH curves.

To unravel the lipophilic nature of these compounds, we employed UV-VIS spectroscopic methods in conjunction with the classic bottle-shaking method using *n*-octanol as the lipophilic phase, adhering to established OECD guidelines [26]. This approach allowed us to determine the partition coefficient (logP), a standard measure of lipophilicity [27], and the more understandable partition coefficient (logD), providing a comprehensive picture of the behavior of molecules in a two-phase system [12].

Lipophilicity is the affinity of a molecule for a lipophilic medium. It is measured by the dispersion behavior in a two-phase system, usually a liquid-liquid or solid-liquid system [27]. Passively absorbed active substances must cross biological membranes before entering the bloodstream. Due to the lipid structure of the membranes, the drug's passage rate directly depends on the lipophilic structure of the active substance. The logP observed in the water-*n*-octanol system of a molecule is adopted as the standard measure of lipophilicity (Equation 3). Most drugs in biological systems have an ionic nature. Therefore, evaluating the lipophilicity of the ionic forms of drugs is more important than the lipophilicity of the molecular form. The lipophilicity of the ionic structure of a drug is defined by its logD. The distribution coefficient is the ratio of the sum of the concentrations of all compound species in octanol to the sum of the concentrations of all compound species in water (Equation 4).

$$\text{LogP} = \text{Log} \frac{[\text{sum of the concentrations of neutral species}]_{\text{octanol}}}{[\text{sum of the concentrations of neutral species}]_{\text{water}}} \quad (3)$$

$$\text{LogD} = \text{Log} \frac{[\text{sum of the concentrations of all microspecies}]_{\text{octanol}}}{[\text{sum of the concentrations of all microspecies}]_{\text{water}}} \quad (4)$$

A certain balance of lipophilicity and hydrophilicity is required for a successful drug candidate substance. The solubility of a drug in water is directly proportional to its solubility in plasma and other aqueous biological fluids. Still, lipophilicity is the ability of the substance to penetrate biological membranes and is crucial for all drug compounds. The log D is a widely used coefficient to measure the lipophilicity of ionizable compounds. For non-ionizing compounds, lipophilicity is expressed by logP, and the logP value is equal to the logD value for these compounds' overall pH ranges. For ionizable compounds, logD considers both ionized and molecular forms. The logP values of acidic molecules can be calculated according to the Equation 5 depending on the pH, logD and pKa values [12].

$$\text{logP}_{(\text{pH})} = \text{logD} + \text{log}[1 + 10^{(\text{pH}-\text{pKa})}] \quad (5)$$

By meticulously analyzing the lipophilic profiles of these compounds across a wide pH range and quantifying their relationship using quadratic functions, we aim to gain a deeper understanding of their physicochemical properties and potential implications for drug development. This comprehensive analysis will provide crucial information for predicting the absorption, distribution, and ultimately, the antifungal efficacy of these promising drug candidates.

Our study marks the first comprehensive determination of pKa, logP, and logD values for these novels 3-(2-phenylethyl)-tetrahydro-2H-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives. The insights gained from our research will not only contribute to the growing body of knowledge surrounding this intriguing class of compounds but also pave the way for the development of potent and effective antifungal agents to combat the rising tide of fungal infections.

Material and Method

Chemicals and Reagents

Stock solutions (5000 ppm) of EP1, EP2, EP3, and EP4 were prepared in methanol and stored at 4°C. These stock solutions were used for all pKa, logD, and logP determination experiments. A 20 mM phosphate buffer was used throughout the study, with pH adjustments made using 0.1 M sodium hydroxide and 0.1 M hydrochloric acid solutions as needed. The calibrations of the pH electrode were made using aqueous buffers.

All reagents, including acetonitrile, methanol, octanol, sodium chloride, sodium dihydrogen phosphate, hydrochloric acid, and sodium hydroxide (Sigma-Aldrich), were used as received. All solutions were prepared using Milli-Q water (Barnstead Nanopure™ system).

Synthesis of EP1, EP2, EP3 and EP4 compounds

Following the synthesis procedure outlined by Ertan et al., four 3-(2-phenylethyl)-tetrahydro-2H-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives (EP1, EP2, EP3, and EP4, Figure 3) were prepared starting from phenylethylamine [4].

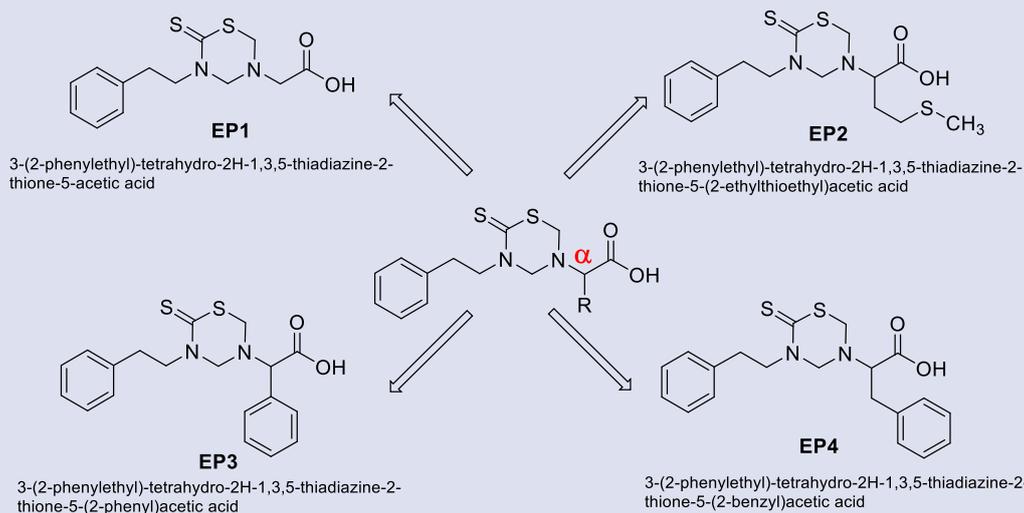


Figure 3. Structure and nomenclature of EP1, EP2, EP3 and EP4.

The synthesis involved reaction with carbon disulfide and potassium hydroxide to yield the potassium phenylethyl dithiocarbamate intermediate, followed by cyclocondensation with formaldehyde and the appropriate amino acids (Figure 4). This key step formed the thiadiazine ring and introduced the acetic acid moiety at position 5, a crucial structural element influencing the compounds' properties and potential antifungal activity. These compounds share a common core structure but differ in the substituent attached to the alpha (α) carbon of the acetic acid moiety, allowing for investigation into the impact of these subtle structural variations on their physicochemical properties and antifungal potential. This study aimed to comprehensively characterize these properties.

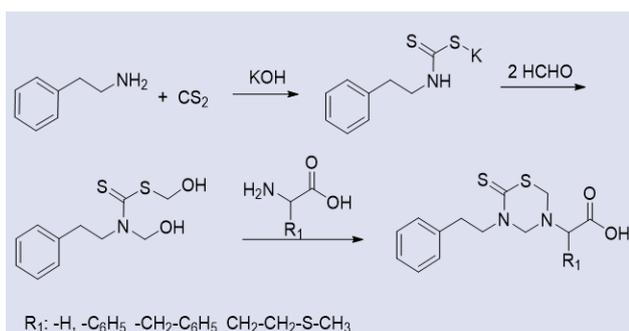


Figure 4. Synthesis reaction steps of 3-(2-phenylethyl)-tetrahydro-2H-1,3,5-thiadiazine-2-thione-5-acetic acid derivatives.

pKa Determination Using PARAFAC Analysis of UV Spectroscopic Data

UV spectra of the compounds were recorded for each compound at eleven pH levels (pH 2-12) using a Shimadzu UV-1280 UV/VIS spectrophotometer. Spectra were acquired in the range of 200-350 nm with a 0.1 nm interval and a scan speed of 100 nm/min. Buffer solutions were prepared and pH values were measured using a Thermo Fisher Orion Star 200 digital pH meter equipped with glass electrodes (uncertainty ± 0.01 pH units).

UV spectra were collected using Shimadzu UVProbe 2.43 software and exported to Microsoft Excel. Three-way data arrays, consisting of pH, wavelength, and absorbance data, were constructed and processed in MATLAB R2023a using the N-way Toolbox [28].

The application of PARAFAC for spectrophotometric pKa determination involves a straightforward procedure focused on constructing and analyzing three-dimensional (3D) datasets.

I. UV absorbance spectra of analytes (EP1, EP2, EP3, EP4) were recorded at a concentration (15 ppm) at a selected pH level ($n=3$).

II. Spectral recordings were repeated for each pH level within the relevant pH range (pH 2-12) to capture the changes in absorbance associated with ionization.

III. The UV absorbance data matrices obtained at each pH level were combined to create a three-way data array with dimensions of absorbance \times wavelength \times pH.

This three-way data array serves as the input for PARAFAC analysis, enabling the deconvolution of the spectral data and accurately determining pKa values.

pKa Determination Using RP-HPLC

Acidity constants were also determined using an RP-HPLC (Agilent 1100) system equipped with a UV detector. Chromatographic separations were performed on an ACE-121-2546 C18 column (4.6 \times 250 mm, 5 μ m particle size).

The mobile phase consisted of 20 mM phosphate buffer (pH 2-12) (A) and acetonitrile (ACN) (B) in isocratic elution mode. In pKa determination studies using the HPLC method, the pH value of the mobile phase was adjusted after mixing ACN with buffer solution in appropriate proportions, and the measured pH (s_pH) values were used for pKa calculation. To ensure accurate and reliable measurements, buffer solutions were prepared using ultra-pure water and stored in capped bottles to prevent carbon dioxide intrusion.

The column temperature was maintained at 25°C, and the injection volume was 10 μ L. UV detection was

performed at 283 nm. Data acquisition and processing were carried out using Agilent ChemStation for LC software (B.04.03-SP1). The column dead time (t_0) was 1.7 minutes, determined with methanol injection.

Capacity factors (k') were determined for each compound at various pH values ($n=6$). The relationship between k' and pH was analyzed using the Nonlinear Regression Shapes toolbox [29] in MATLAB. pK_a values were determined from the second derivative of the resulting k' -pH curves.

To assess the robustness of retention time estimations, experiments were also conducted using different ACN:water ratios in the mobile phase (65:35, 70:30, and 75:25, v/v). For each pH value, three independent replicates were performed to ensure accurate determination of k' values.

Determination of Lipophilicity ($\log P$ and $\log D$)

Lipophilicity was assessed by determining the $\log D$ using the classical shake-flask method with *n*-octanol, following OECD guidelines [26].

n-octanol and water were pre-saturated by shaking together for 24 hours. Aqueous phases were prepared using phosphate buffer, with pH adjustments (2.0-10.0) made using 0.1 M HCl or 0.1 M NaOH solutions. Ionic strength (μ) was maintained at 0.15 M using NaCl.

For each compound, 4 mL of the aqueous phase (at the desired pH) and 2 mL of *n*-octanol were combined and shaken for 1 hour. The samples were then centrifuged at 2000 g for 10 minutes to separate the phases.

Distribution coefficients (D) were determined using Equation 6:

$$D = \left(\frac{A_i - A_f}{A_f} \right) \times \frac{V_w}{V_o} \quad (6)$$

Where: A_i : UV absorbance of the aqueous phase before extraction, A_f : UV absorbance of the aqueous phase after extraction, V_w : Volume of the water-based layer, V_o : Volume of the *n*-octanol layer.

Each determination was performed in triplicate at each pH value. UV absorbance measurements were carried out using a Shimadzu UV-1280 spectrophotometer. All experiments were conducted at room temperature ($22.3 \pm 0.2^\circ\text{C}$). To account for spectral shifts due to pH, absorbance values were measured at 285 nm for pH 2-5 and at 290 nm for pH 5-10.

D values were calculated for each compound at pH 2-10 using Equation 6. $\log D$ values were then plotted as a function of pH, and quadratic equations describing the relationship were derived using polynomial second-order regression in MATLAB. $\log P$ values were calculated from the $\log D$ values using Equation 5.

Determination of pK_a values is crucial to understand the behavior of drug candidates. In this study, pK_a values of EP1, EP2, EP3 and EP4 with antifungal activity were determined using two independent methods (PARAFAC analysis of UV spectroscopic data and RP-HPLC).

Results and Discussion

Acidity Constant Determination via PARAFAC

The pH-dependent change in a three-way data set was analyzed to determine the pK_a values of EP1, EP2, EP3, and EP4. UV absorbance spectra were recorded for each compound using the PARAFAC method. Spectra were obtained at eleven pH values (pH 2–12, $\Delta\text{pH} = 1$) in the wavelength range of 200–340 nm and at a concentration of 15 ppm (Figure 5a–b). This process produced a series of 140×1 wavelength \times absorbance data matrices for each of the eleven pH levels.

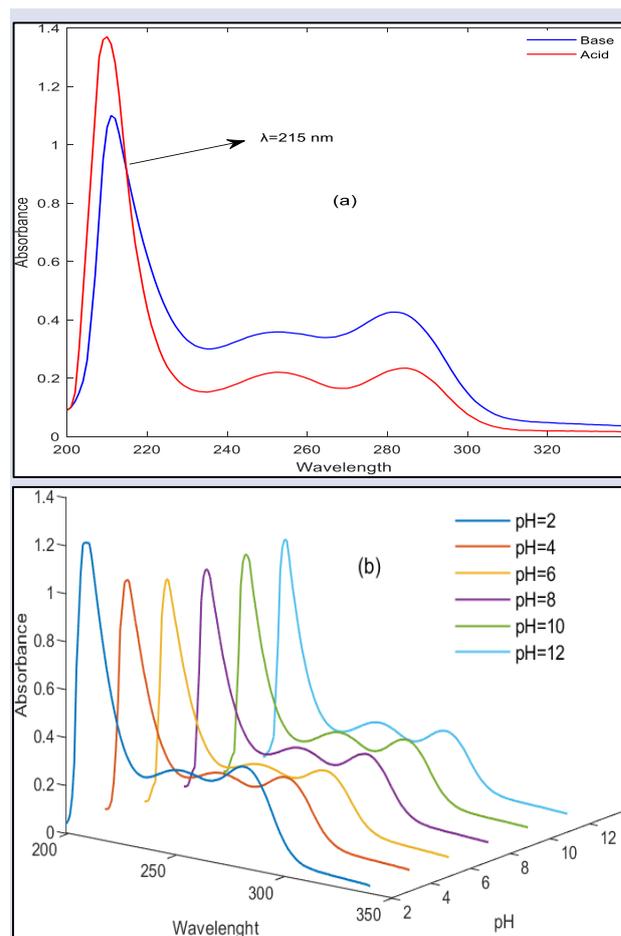


Figure 5. (a) UV spectrum (200-340 nm) of acidic and basic forms EP1 (15 ppm). (b) pH - absorbance - wavelength graph EP1 (15 ppm).

The eleven data matrices were then combined along the pH axis, resulting in a three-way data array (1 \times 140 \times 11 dimensions; absorbance \times wavelength \times pH) for analysis. This array was deconvoluted using the PARAFAC algorithm to differentiate the acidic and basic forms of the compounds. Both constrained (positive mode only) and unconstrained (positive and negative modes) PARAFAC models were evaluated. The unconstrained mode was selected for further analysis due to the negligible difference observed in the pH values at which the acidic and basic forms intersected between the two models.

Two-component PARAFAC modeling in the unconstrained mode, using CORCONDIA [30], yielded the optimal profiles for spectral, pH, and absorbance modes. The lowest CORCONDIA value obtained was 99.06% for all compounds using this two-component model. This result,

coupled with the observation that the sum of the acidic and basic forms approximated unity, strongly suggests that the ionizable proton in these molecules is associated solely with the carboxylic acid group.

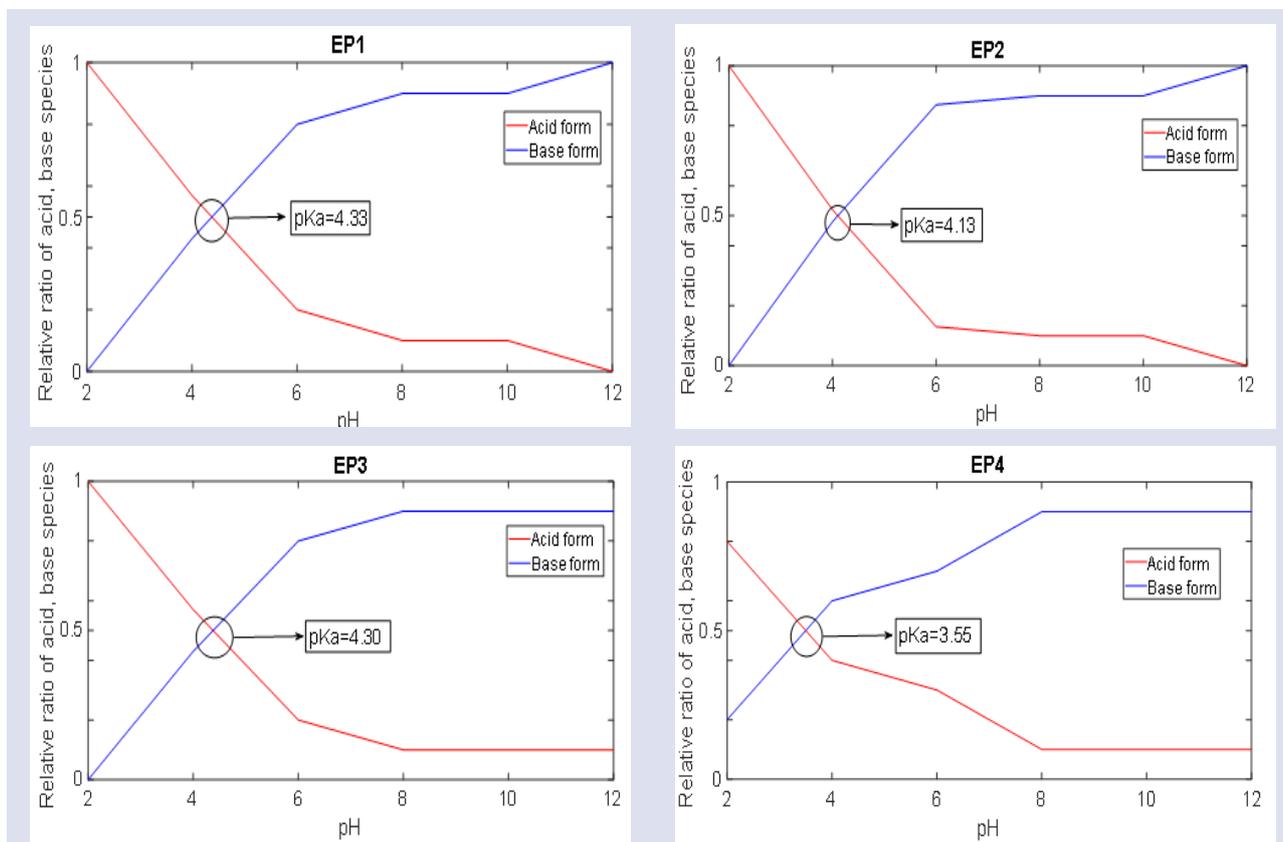


Figure 6. Estimated profiles for wavelength, pH, and concentration loadings of three-dimensional data in the PARAFAC deconvolution procedure.

Figure 6 displays the estimated profiles for absorbance, wavelength, and pH loadings resulting from the PARAFAC deconvolution. The red curves represent the relative amount of the compound in its acidic form, while the blue curves depict the relative amount in its basic form. The pKa value for each compound, representing the pH at which the acidic and basic forms are equal (intersection of the curves), is indicated on the graph (Equation 1). The pKa values determined using the PARAFAC method are presented in Table 1.

Isosbestic points in a UV absorption spectrum represent wavelengths where the absorbance of a molecule remains constant regardless of its ionization state (acid, base, or zwitterion). These points signify a direct relationship between the number of ionizable groups and the pKa values of the molecule. As the pKa value decreases, the corresponding isosbestic points generally shift to shorter wavelengths. The presence of a single isosbestic point at approximately 215 nm in the UV spectra of EP1, EP2, EP3, and EP4 suggests that each

molecule possesses a single ionizable group with a relatively low pKa value.

Acidity Constant Determination via HPLC

The pKa values of EP1, EP2, EP3, and EP4 were also determined using RP-HPLC under various elution conditions. Retention factors were measured for each compound at different pH values and mobile phase compositions (ACN: buffer ratios of 65:35, 70:30, and 75:25 v/v).

Capacity factor (k') values were plotted as a function of pH, and the resulting curves were fitted using second-order nonlinear regression (Figure 7). pKa values were then determined from the second derivative of these fitted equations (Table 1). The RP-HPLC method yielded pKa values of 4.270 ± 0.121 , 4.041 ± 0.280 , 4.124 ± 0.319 , and 3.332 ± 0.015 for EP1, EP2, EP3, and EP4, respectively. As shown in Figure 7, a strong correlation was observed between the experimental k' and the fitted curves across the entire pH range studied.

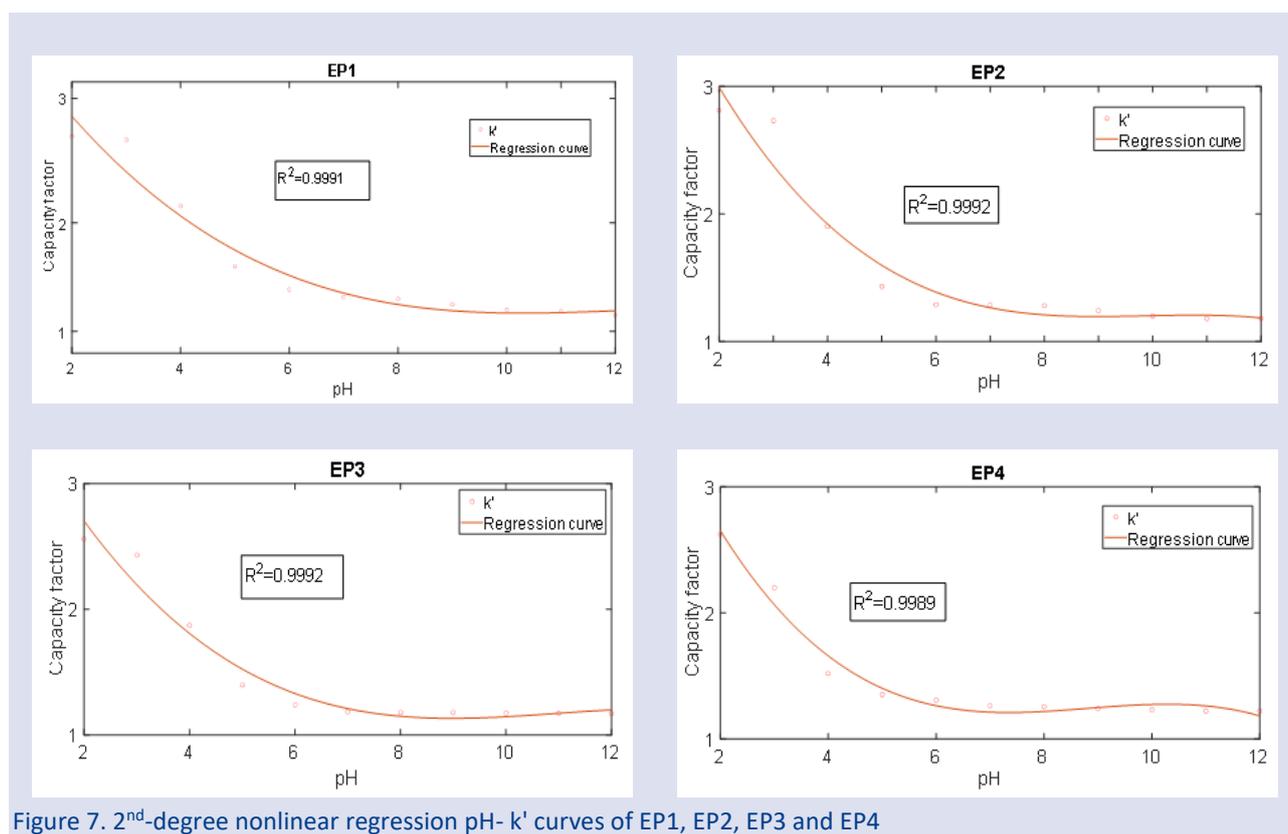


Figure 7. 2nd-degree nonlinear regression pH- k' curves of EP1, EP2, EP3 and EP4

Table 1. Experimental pKa Values for EP1, EP2, EP3 and EP4

| | PARAFAC ^a | RP-HPLC ^{a,b} |
|-----|----------------------|------------------------|
| EP1 | 4.332±0.016 | 4.270±0.121 |
| EP2 | 4.137±0.030 | 4.041±0.280 |
| EP3 | 4.305±0.029 | 4.124±0.319 |
| EP4 | 3.510±0.022 | 3.332±0.015 |

^a pKa ± standard error (SE)

^b The average of the pKa values obtained in three different mobile phase compositions (ACN: buffer 65:35, 70:30, 75:25 v/v).

A thorough understanding of ionization mechanisms is crucial in pharmaceutical science, as a drug's pKa value significantly influences key properties such as solubility, absorption, distribution, excretion, and stability.

The molecules investigated in this study contain two tertiary nitrogen atoms within the thiadiazine ring, theoretically allowing for up to three distinct pKa values. However, only a single pKa value was detected using both PARAFAC and RP-HPLC methods. This discrepancy might be attributed to the high reactivity of tertiary amines. Literature suggests that pKa values associated with cyclic tertiary nitrogen atoms typically fall within the range of 9-11 [33]. It is plausible that buffer components, interacting with various species formed within this pH range, could suppress the ionization of these nitrogen atoms during the experimental procedures.

The pKa values of carboxylic acids are influenced by the inductive effects of substituents attached to the α -carbon. Electron-withdrawing groups (-I effect) decrease

the pKa value, while electron-donating groups (+I effect) increase it. In this study, the experimentally determined pKa values were evaluated in light of the inductive effects exerted by the substituents on the α -carbon. Based on inductive effects alone, the following trends were expected:

The predicted pKa values for EP2, EP3, and EP4 were influenced by the electron-withdrawing nature of their respective α -carbon substituents. The strongly electron-withdrawing sulfur atom in the 2-methylthioethyl group of EP2 suggested a lower pKa than both EP1 and EP3. Similarly, the phenyl ring of EP3, being more electron-withdrawing than a hydrogen atom, was expected to result in a lower pKa than EP1. Finally, the benzyl group of EP4, exerting both inductive and resonance electron-withdrawing effects, was anticipated to yield the lowest pKa among all the compounds.

The experimentally observed pKa values were consistent with these predictions based on inductive effects, supporting the influence of substituent electronics on the acidity of these compounds.

In our previous study, the antifungal activity of EP1, EP2, EP3, and EP4 against *C. albicans*, *C. parapsilosis*, *C. stellatoidea*, and *C. pseudotropicalis*, minimum inhibitory concentrations (MICs) and minimum fungicidal concentrations (MFCs) were determined [4]. Comparing the pKa values determined in this study with the previously reported antifungal activities revealed a trend of lower average activity for compounds with pKa values between 4.0 and 4.3. Interestingly, EP2 and EP4, with pKa values outside of this range, exhibited approximately 30% higher average activity. While a statistically significant difference in activity was not observed, these preliminary

findings highlight the potential influence of pKa on antifungal efficacy and warrant further investigation. More comprehensive activity studies are needed to fully elucidate the relationship between pKa and antifungal activity for these compounds.

Distribution and Partition Coefficients

Following the OECD guidelines, as previously outlined, the distribution coefficients (D) of EP1, EP2, EP3, and EP4 were determined across a range of pH values (Table 2). The minimum logD value was observed at pH 6.0 for EP1 and EP2, and at pH 5.0 for EP3 and EP4.

The relationship between logD and pH was further analyzed by fitting the data to quadratic equations. The resulting equations and corresponding graphical representations are presented in Figure 8, demonstrating the pH-dependent lipophilicity profiles of each compound.

Table 2 logD experimental values for EP1, EP2, EP3 and EP4.

| pH | EP1 ^c | EP2 ^c | EP3 ^c | EP4 ^c |
|----|------------------|------------------|------------------|------------------|
| 2 | 2.250±0.010 | 0.758±0.782 | -0.792±0.185 | 0.348±0.151 |
| 3 | 2.081±0.814 | 0.405±0.103 | -1.213±0.936 | 0.247±0.253 |
| 4 | 1.913±0.665 | 0.057±0.574 | -1.646±0.689 | 0.135±0.835 |
| 5 | 1.717±0.129 | -0.292±0.823 | -2.119±0.743 | -0.001±0.929 |
| 6 | 1.706±0.050 | -0.496±0.828 | -2.094±0.247 | -0.033±0.142 |
| 7 | 1.882±0.998 | -0.261±0.470 | -1.000±0.117 | 0.036±0.871 |
| 8 | 2.252±0.137 | 0.107±0.518 | 0.584±0.303 | 0.208±0.521 |
| 9 | 2.631±0.781 | 0.794±0.015 | 1.641±0.593 | 0.476±0.883 |
| 10 | 3.034±0.545 | 1.452±0.313 | 2.695±0.846 | 0.752±0.366 |

^clogD±SE: Standard error

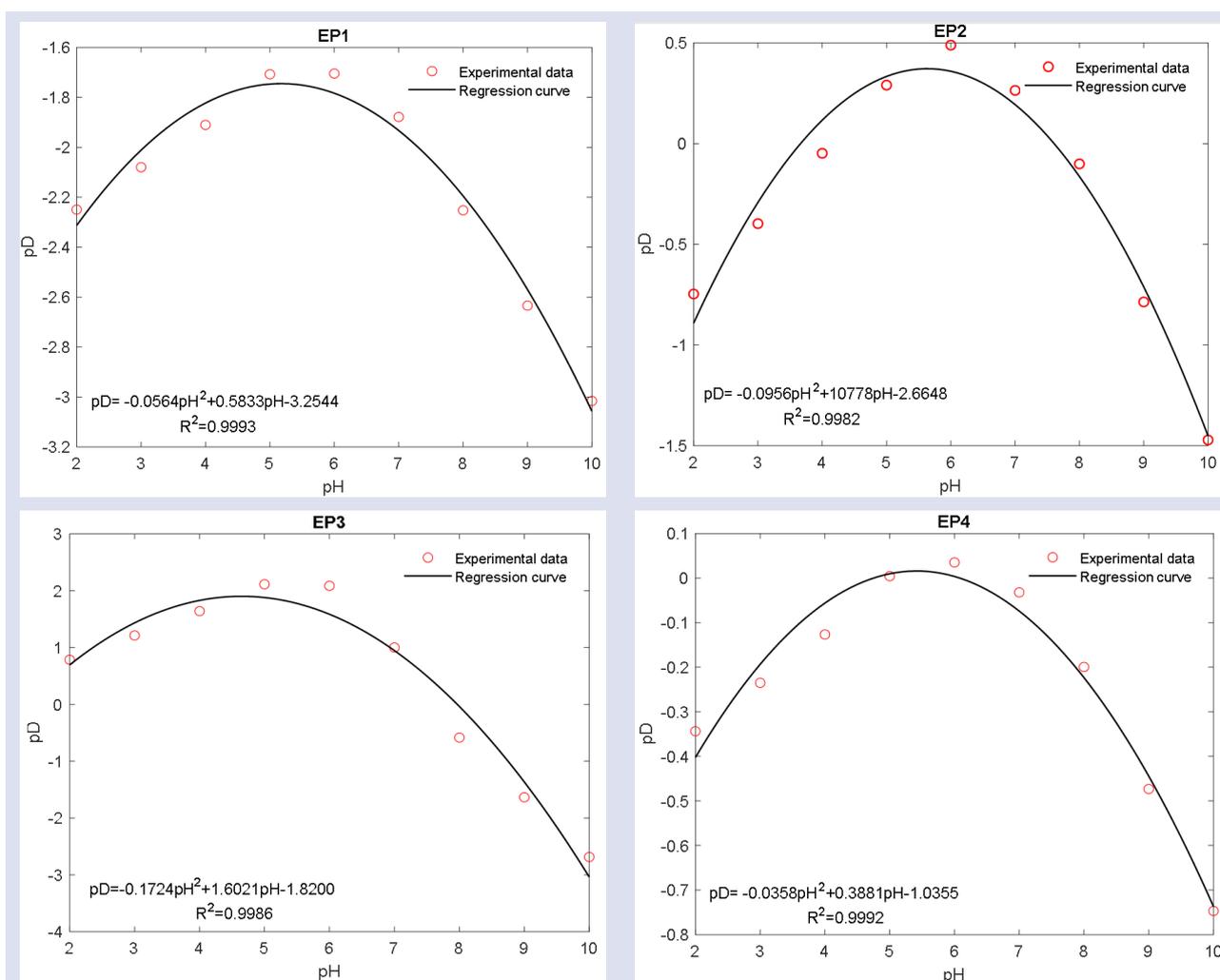


Figure 8. The logD - pH profiles and equations for EP1, EP2, EP3 and EP4.

To determine the logP values representing the neutral (molecular) forms of the compounds, the measured logD values were corrected for the relative proportions of ionized species at each pH using the pKa values [31, 32]. This correction is crucial because logD reflects the overall distribution of both ionized and non-ionized forms, while

logP specifically describes the partitioning of the neutral form. The calculated logP values were as follows: EP1 (3.00), EP2 (1.10), EP3 (-0.38), and EP4 (2.19).

The efficacy of antifungal drugs and their ability to reach target tissues are closely related to the logD partition coefficient, an important parameter reflecting

both water and lipid solubility of the drug molecule. The curves in the logD-pH profiles in Figure 8 showed a conical parabolic shape, which is typical for each analyzed compound in all four cases. However, the shape of the EP3's curve is comparatively less sharp than the others. This means that the logD of EP3 is less pH dependent than other substances. When highly acidic or basic conditions are evaluated, it can be said that logD values are higher in extremely acidic conditions.

LogP is a key parameter for predicting the in vivo absorption of a drug. Compounds with logP values between 1 and 3 typically exhibit good absorption. In this study, EP1, EP2, and EP4, with logP values above 1, are predicted to have favorable absorption profiles. A logP value below 1 generally indicates a hydrophilic compound, while a value above 1 suggests lipophilicity. Thus, EP1, EP2, and EP4 can be classified as lipophilic, while EP3, with a logP below 1, is hydrophilic.

Conclusion

This study determined the pKa values of several 2-(5-phenethyl-6-thioxo-1,3,5-thiadiazinan-3-yl) acetic acid derivatives for the first time, employing both UV spectroscopy and RP-HPLC. The results strongly suggest that proton removal occurs primarily from the carboxyl group, as illustrated by the proposed protonation reaction in Figure 4.

Accurate pKa determination relies on meticulous experimental techniques. For both UV spectroscopy and RP-HPLC, using freshly prepared buffer solutions tailored to the specific pH range is crucial. In PARAFAC analysis, careful selection of the appropriate number of components is critical, as even a single component deviation can result in a pKa error of approximately $\pm 25\%$. Furthermore, analyzing absorbance data within a wavelength range sensitive to pH-dependent changes enhances the sensitivity of PARAFAC analysis.

While both PARAFAC and RP-HPLC proved effective for pKa determination, the UV spectroscopy-assisted PARAFAC approach offered a significant time advantage over the longer elution times associated with RP-HPLC. These findings underscore the utility of both methods, particularly the PARAFAC approach, for accurate and efficient pKa determination.

The logD-pH profiles for all analyzed compounds exhibited a characteristic bell-shaped curve. The highest logD value for each compound was observed at a pH where approximately 50% of the molecules were in their ionized form, a point closely aligned with their respective pKa values. LogP values, calculated from the logD data, indicated that EP1, EP2, and EP4 possess favorable lipophilicity profiles for drug development. However, EP3, with a lower logP value, suggests higher hydrophilicity and potentially lower stability in acidic environments compared to the other compounds.

The accurate pKa values established in this work may be leveraged to systematically explore how modifications to the α -carbon substituent or other regions of the core

scaffold impact both the physicochemical properties and the antifungal efficacy of these derivatives in future studies. These findings can then guide in vivo studies using relevant animal models to evaluate the pharmacokinetic behavior of these compounds.

The pKa and logP/logD data generated in this study provide valuable parameters that will help develop optimal pharmaceutical dosage forms. These formulations should aim to enhance solubility, stability, and targeted delivery to sites of fungal infection. For example, pro-drug strategies that leverage pH-dependent ionization could be beneficial for EP3 because of its relatively low lipophilicity. Such strategies could potentially improve the absorption, distribution, and overall pharmacokinetic profile of this compound.

Employing computational modeling techniques, researchers can further investigate the relationship between the physicochemical properties, 3D structures, and potential interactions of these compounds with their molecular targets within fungal cells. This approach could provide valuable insights into their mechanism of action and guide further structural optimization efforts.

By pursuing these research avenues, a more comprehensive understanding of this promising class of antifungal agents can be achieved, potentially leading to the development of new and effective treatments for invasive fungal infections.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgment

I would like to express my gratitude to the Department of Analytical Chemistry, Faculty of Pharmacy, Hacettepe University, for providing the resources and facilities essential for the completion of this project.

Ethical Approval Statement

Ethical approval was not required for this study, as it involved only chemical synthesis, physicochemical characterization, and in silico analysis, with all procedures adhering to standard laboratory practices.

References

- [1] S. Campoy, J.L. Adrio, Antifungals, *Biochemical pharmacology*, 133 (2017) 86-96.
- [2] G. Wall, J.L. Lopez-Ribot, Current antimycotics, new prospects, and future approaches to antifungal therapy, *Antibiotics*, 9(8) (2020) 445.
- [3] J. Loeffler, D.A. Stevens, Antifungal Drug Resistance, *Clinical Infectious Diseases*, 36(Supplement_1) (2003) S31-S41.
- [4] M. Ertan, A. Bilgin, E. Palaska, N. Yulug, Syntheses and antifungal activities of some 3-(2-phenylethyl)-5-

- substituted-tetrahydro-2H-1, 3, 5-thiadiazine-2-thiones, *Arzneimittel-forschung*, 42(2) (1992) 160-163.
- [5] A.N. El-Shorbagi, New Tetrahydro-2H-1, 3, 5-thiadiazine-2-thione Derivatives as Potential Antimicrobial Agents, *Archiv der Pharmazie: An International Journal Pharmaceutical and Medicinal Chemistry*, 333(9) (2000) 281-286.
- [6] M. Ertan, S. Sarac, N. Yulug, Synthesis and antimicrobial activities of some new tetrahydro-2H-1, 3, 5-thiadiazine-2-thione derivatives of amoxicillin, *Arzneimittel-forschung*, 40(7) (1990) 790-795.
- [7] S. Raheem, R. Khan, X. Pan, R. Ullah, S.A. Halim, A. Khan, A. Al-Harrasi, One pot domino synthesis of new 3, 5-disubstituted-tetrahydro-2H-1, 3, 5-thiadiazine-2-thiones (THTTs) as anti-inflammatory and antinociceptive candidates: A proof from in-vivo to in-vitro and in-silico mechanistic studies, *Bioorganic Chemistry*, 127 (2022) 105974.
- [8] H. Ali, R. Khan, X. Pan, F. Shaheen, A. Jabeen, A. Rauf, M. Shah, U. Rashid, Y.S. Al-Awthan, O.S. Bahattab, Synthesis, characterization, anticancer, anti-inflammatory activities, and docking studies of 3, 5-disubstituted thiadiazine-2-thiones, *Green Processing and Synthesis*, 12(1) (2023) 20228136.
- [9] D. Katiyar, V. Tiwari, R. Tripathi, A. Srivastava, V. Chaturvedi, R. Srivastava, B. Srivastava, Synthesis and antimycobacterial activity of 3, 5-disubstituted thiadiazine thiones, *Bioorganic & medicinal chemistry*, 11(20) (2003) 4369-4375.
- [10] N. Arshad, J. Hashim, M.A. Minhas, J. Aslam, T. Ashraf, S.Z. Hamid, T. Iqbal, S. Javed, New series of 3, 5-disubstituted tetrahydro-2H-1, 3, 5-thiadiazine thione (THTT) derivatives: synthesis and potent antileishmanial activity, *Bioorganic & Medicinal Chemistry Letters*, 28(19) (2018) 3251-3254.
- [11] Y. Ishihama, M. Nakamura, T. Miwa, T. Kajima, N. Asakawa, A rapid method for pKa determination of drugs using pressure-assisted capillary electrophoresis with photodiode array detection in drug discovery, *Journal of pharmaceutical sciences*, 91(4) (2002) 933-942.
- [12] C. Hansch, A. Leo, Substituent constants for correlation analysis in chemistry and biology, *Journal of Quantum Information Science, Vol 1, No 2*, 1979.
- [13] E. Dinç, Z.C. Ertekin, N. Ünal, Three-way analysis of pH-UV absorbance dataset for the determination of paracetamol and its pKa value in presence of excipients, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 230 (2020) 118049.
- [14] E. Dinç, F. Selimoğlu, N. Ünal, Z.C. Ertekin, Simultaneous Determination of the Acid Dissociation Constants of Phenolics by Multivariate Analysis of PH and Ultraviolet-Visible Spectrophotometric Measurements, *Analytical Letters*, 54(16) (2021) 2624-2637.
- [15] E. Dinç, N. Ünal, Z.C. Ertekin, Three-way analysis-based pH-UV-Vis spectroscopy for quantifying allura red in an energy drink and determining colorant's pKa, *Journal of Food and Drug Analysis*, 29(1) (2021) 76.
- [16] F. Selimoğlu, N. Ünal, Z.C. Ertekin, E. Dinç, PARAFAC and MCR-ALS approaches to the pKa determination of benzoic acid and its derivatives, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 248 (2021) 119253.
- [17] Z. Yazan, S. Erden, E. Dinç, A comparative application of two-way and three-way analysis to three-dimensional voltammetric dataset for the pKa determination of vanillin, *Journal of Electroanalytical Chemistry*, 826 (2018) 133-141.
- [18] Z. Yazan, D. Eskiköy Bayraktepe, E. Dinç, Square wave voltammetric pKa determination of aspirin using multi-way data analysis models, *Chemical Papers*, 76(9) (2022) 5389-5397.
- [19] E. Dinç, Z. Yazan, Z.C. Ertekin, C. Yıldız, Two-and Three-Dimensional Strategies for Determining Consecutive Acidity Constants of Ciprofloxacin and Enrofloxacin Using Fluorescent Excitation-Emission Measurements, *ChemistrySelect*, 8(27) (2023) e202301809.
- [20] K.R. Murphy, C.A. Stedmon, D. Graeber, R. Bro, Fluorescence spectroscopy and multi-way techniques. PARAFAC, *Analytical methods*, 5(23) (2013) 6557-6566.
- [21] N.K.M. Faber, R. Bro, P.K. Hopke, Recent developments in CANDECOMP/PARAFAC algorithms: a critical review, *Chemometrics and Intelligent Laboratory Systems*, 65(1) (2003) 119-137.
- [22] S. Espinosa, E. Bosch, M. Rosés, Retention of ionizable compounds on HPLC. 12. The properties of liquid chromatography buffers in acetonitrile- water mobile phases that influence HPLC retention, *Analytical chemistry*, 74(15) (2002) 3809-3818.
- [23] C. Horváth, W. Melander, I. Molnar, Liquid chromatography of ionogenic substances with nonpolar stationary phases, *Analytical Chemistry*, 49(1) (1977) 142-154.
- [24] P.-C. Chiang, Y. Hu, Simultaneous determination of LogD, LogP, and pKa of drugs by using a reverse phase HPLC coupled with a 96-well plate auto injector, *Combinatorial chemistry & high throughput screening*, 12(3) (2009) 250-257.
- [25] T. Angelov, A. Vlasenko, W. Tashkov, HPLC determination of pKa of parabens and investigation on their lipophilicity parameters, *Journal of liquid chromatography & related technologies* 31(2) (2007) 188-197.
- [26] O.T. No, 107: Partition Coefficient (n-octanol/water): Shake flask method, OECD Guidelines for the Testing of Chemicals, Section 1 (1995) 1.
- [27] A.D. McNaught, A. Wilkinson, Compendium of chemical terminology, *Blackwell Science Oxford*, 1997.
- [28] C.A. Andersson, R. Bro, The N-way toolbox for MATLAB, *Chemometrics and intelligent laboratory systems*, 52(1) (2000) 1-4.
- [29] J. D'Errico, *Nonlinear Regression Shapes*, 2023.
- [30] E.E. Papalexakis, C. Faloutsos, Fast efficient and scalable core consistency diagnostic for the parafac decomposition for big sparse tensors, 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2015, pp. 5441-5445.
- [31] J. Vazquez, S. Merino, Ò. Doménech, M. Berlanga, M. Vinas, M. Montero, J. Hernandez-Borrell, Determination of the partition coefficients of a homologous series of ciprofloxacin: influence of the N-4 piperazinyl alkylation on the antimicrobial activity, *International journal of pharmaceutics*, 220(1-2) (2001) 53-62.
- [32] K. Takács-Novák, M. Józán, I. Hermeicz, G. Szász, Lipophilicity of antibacterial fluoroquinolones, *International Journal of Pharmaceutics*, 79(1-3) (1992) 89-96.
- [33] J. Kromann, F. Larsen, H. Moustafa, J. Jensen, Prediction of pKa values using the PM6 semiempirical method. *PeerJ*, 4: e2335, 2016.

Oxidative Stress Mediates Anti-proliferative Effects of Nifedipine on AGS Cells

Ziad Joha ^{1,a,*}, Sude Nur Göksu ^{2,b}

¹ Department of Pharmacology, Faculty of Pharmacy, Sivas Cumhuriyet University, Sivas, Türkiye

² Faculty of Medicine, Sivas Cumhuriyet University, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 18/08/2024

Accepted: 16/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Previous studies have demonstrated the anticancer properties of certain calcium channel blockers when administered as sole agents. This exploration aimed to explore the potential anti-proliferative activity of nifedipine on AGS gastric cancer cells and to determine the contribution of oxidative stress in mediating this response. To assess the anti-proliferative impact of nifedipine, AGS cell viability was calculated employing the XTT assay across a concentration range of 25, 50, 100, 200, and 500 µg/mL. Concurrently, TAS and TOS kits were employed to evaluate the drug's influence on oxidative stress levels. Nifedipine exhibited a concentration-dependent cytotoxic effect on AGS cells, with a statistically significant reduction in cell viability. The IC50 value for AGS cells after 24 hours was determined to be 98.49 µg/mL. At this concentration, a substantial augmentation in TOS and a concurrent diminution in TAS levels were observed relative to the control group. These observations imply that nifedipine-induced oxidative stress is one of the mechanisms underlying its cytotoxic action against gastric cancer cells. Our results highlight the potential therapeutic utility of nifedipine in this cancer type.

Keywords: Proliferative Activity, Nifedipine, Oxidative Stress, AGS Cells.

 zead-geha@hotmail.com

 <https://orcid.org/0000-0001-8520-3760>

 goksunursude@gmail.com

 <https://orcid.org/0009-0007-9021-2021>

Introduction

Stomach cancer persists as a substantial worldwide health crisis, occupying the fifth position among the most frequently identified cancers. Due to its often advanced detection, this disease carries a high mortality rate, claiming 784,000 lives worldwide in 2018, consequently positioning it as the third leading cause of cancer-related deaths [1]. A multitude of factors contribute to gastric cancer risk, including age, sex, genetics, smoking habits, ethnicity, poor diet, and infection with the *Helicobacter pylori* bacterium. Traditionally, treatment options for gastric cancer have encompassed chemotherapy, radiation therapy, surgical intervention, and immunotherapy [2]. Unfortunately, a significant number of patients struggle to tolerate the adverse effects of these conventional treatments [3]. Consequently, there is a critical need to identify novel cellular targets to enhance treatment effectiveness and mitigate the adverse consequences of current therapies. Accumulating evidence indicates that ion channels, specialized membrane proteins regulating ion flow, are implicated in the pathogenesis of various diseases, including cancer [4]. Calcium signaling modulation has shown promise in treating a diverse range of diseases, such as high blood pressure, coronary heart disease, and pain [5]. Additionally, calcium acts as a crucial intracellular messenger, governing cellular functions including cell growth, cell cycle progression, specialization, movement, and programmed cell death. Importantly, aberrant calcium channel activity has been implicated in tumorigenesis and cancer progression [6]. Calcium

signaling activation has been highlighted as a key driver of cancer progression. This is attributed to its role in fostering malignant cell behavior, interacting with multiple oncogenes and oncogenic signaling cascades, and suppressing immune responses [7]. Comparative analyses have revealed significantly elevated expression levels of CACNA1C and CACNA1D—genes encoding L-type calcium channels—in gastric, colorectal, breast, and prostate cancer tissues compared to their normal counterparts [8]. Research suggests that suppressing calcium signaling could be a promising approach for cancer treatment [7]. Furthermore, any ion channels or pumps that allow calcium ion passage represent potential therapeutic targets [9]. The calcium channel blocker nifedipine has demonstrated the ability to hinder the progression and immune evasion of colorectal cancer by inhibiting the translocation of NFATC2 into the cell nucleus [6]. Nifedipine has been shown to potentiate the antitumor efficacy of cisplatin against a range of human glioblastoma cells, including those exhibiting sensitivity, resistance, or even multidrug resistance to cisplatin [10]. Diltiazem has been shown to reverse the development of resistance to docetaxel and vincristine in human lung cancer cell populations [11]. Diltiazem has been demonstrated to suppress the growth and movement of hepatocellular cells in laboratory conditions [12]. The calcium channel inhibitors lercanidipine and amlodipine have exhibited the ability to suppress the transcription regulated by YY1, ERK, and TGF-β, thereby enhancing the responsiveness of gastric cancer cells to doxorubicin [13].

To date, the antiproliferative properties of nifedipine in gastric cancer have remained unexplored. Accordingly, this research sought to explore the inhibitory impact of nifedipine on the proliferation of AGS gastric cancer cells.

Materials and Methods

Cell culture

The CRL-1739 cell line was selected for this study due to its well-characterized nature as a human gastric adenocarcinoma cell line (AGS). AGS cells, obtained from the ATCC, were cultured in a 25 cm² flask under sterile conditions at 37°C and 5% CO₂ in DMEM (1:1) cell culture medium supplemented with 1% L-glutamine, 1% penicillin-streptomycin, and 10% fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific). Cells were passaged when they reached 80% confluency, and experiments were initiated after the third passage. Nifedipine (Tocris) was applied at different concentrations (25, 50, 100, 200, and 500 µg/mL). Stock solution of nifedipine was created by dissolving the compound in saline. Working solutions of various concentrations were subsequently generated by diluting appropriate volumes of the stock solution with DMEM [14].

Cell Viability Assay

To assess nifedipine's impact on AGS cell survival, an XTT assay was conducted. AGS cells were cultured in 96-well plates at a seeding density of 10,000 cells per well. Cells were exposed to various nifedipine concentrations for 24 hours. The XTT assay operates on the principle that metabolically active cells convert the tetrazolium salt XTT into a soluble orange formazan product. The formazan concentration, measured spectrophotometrically at 450 nm, correlates with the number of viable cells. Following the treatment period, XTT reagent was added to each well, and the plate was incubated for four hours under controlled conditions. The untreated cell group served as the viability control, set at 100%, against which treatment group results were compared [15,16].

Cell Homogenate Preparation

Cells from each experimental group were collected aseptically and subjected to centrifugation at 2000 rpm for 10 minutes. Following the removal of the supernatant, the cell pellet was resuspended in phosphate-buffered saline (PBS, pH 7.4) to achieve a cell density of approximately 1 million cells per milliliter. To release intracellular components, the cell suspension underwent multiple freeze-thaw cycles. Subsequently, the lysate was centrifuged at 4000 rpm for 10 minutes at 4°C, and the resulting supernatant was collected for biochemical analyses [17,18]. Total protein concentration within the samples was determined using a Bradford protein assay kit (Merck Millipore, Darmstadt, Germany) [17].

Measurement of Total Antioxidant Status (TAS) and Total Oxidant Status (TOS)

TAS levels in cell culture supernatants were quantified using an automated assay kit (Rel Assay Kit Diagnostics, Antep, Turkey) established by Erel (2004). This method is

based on the principle of colorimetric detection. Hydroxyl radicals, generated through the Fenton reaction, initiate a chain reaction producing colored dianisidyl radicals. Antioxidants in the sample inhibit this color development in proportion to their concentration [19,20]. Results were expressed as micromolar Trolox equivalents per milligram of tissue protein (µmol Trolox Eq/mg protein). TOS levels in cell culture supernatants were determined using an automated assay (Rel Assay Kit Diagnostics, Antep, Turkey) developed by Erel (2005). The assay measures TOS indirectly by quantifying ferric ions, formed from the oxidation of ferrous ions by oxidants in the sample. Xylenol orange is employed to detect ferric ions, and hydrogen peroxide serves as the assay standard [21,22]. Results were expressed as micromolar hydrogen peroxide equivalents per milligram tissue protein (µmol H₂O₂ Eq/mg protein).

Statistical Analysis

The statistical analysis of the obtained data was performed using the SPSS Version 25.0 program. Since the data showed a normal distribution, one-way ANOVA (Analysis of Variance) was applied. Values with $p < 0.05$ were considered statistically significant.

Results and Discussion

Nifedipine Inhibits AGS Cell Proliferation

Nifedipine was tested for its ability to inhibit AGS cell growth. Cells were exposed to concentrations ranging from 25 to 500 µg/mL for 24 hours. Cell viability was assessed using the XTT assay. No significant antiproliferative effect was observed at doses of 25 or 50 µg/mL. In contrast, a dose-dependent reduction in cell viability was evident at 100, 200, and 500 µg/mL ($p < 0.001$; Fig. 1). The half-maximal inhibitory concentration (IC₅₀) of nifedipine, calculated using GraphPad Prism, was 98.49 µg/mL.

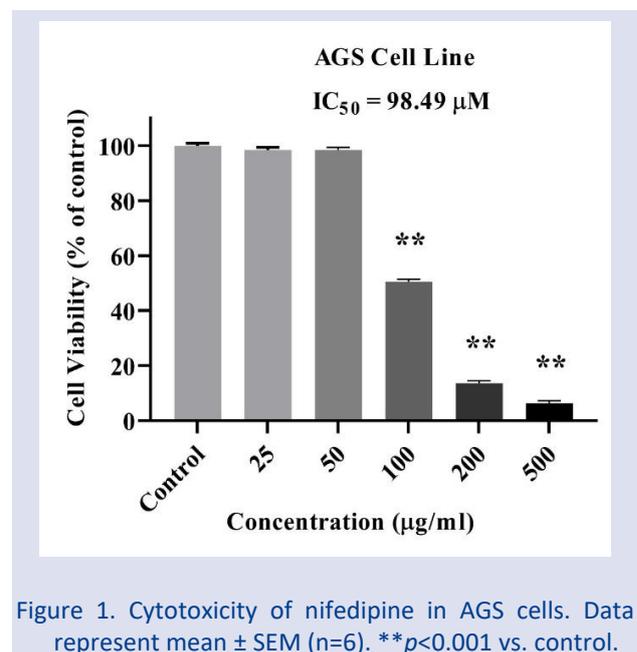
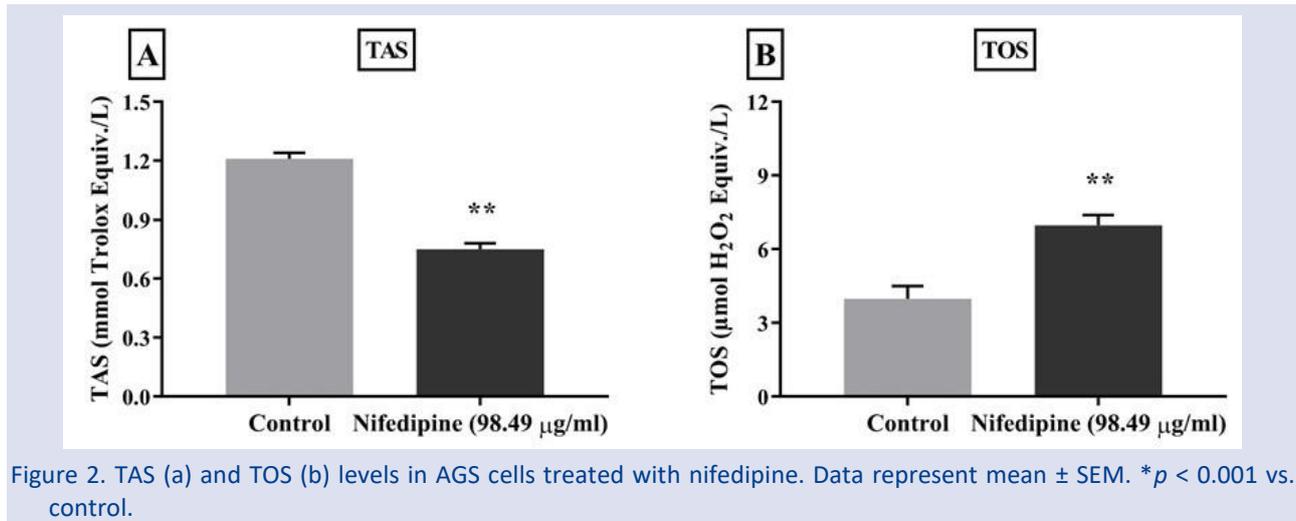


Figure 1. Cytotoxicity of nifedipine in AGS cells. Data represent mean ± SEM (n=6). ** $p < 0.001$ vs. control.

Nifedipine Modulates TAS and TOS in AGS Cells

Using commercial assay kits, Nifedipine's influence on TAS and TOS was evaluated in AGS cells. Treatment with

98.49 $\mu\text{g}/\text{mL}$ nifedipine resulted in a significant decrease in TAS and a concomitant increase in TOS levels compared to the control group ($p < 0.001$; Figure 2.a,b).



While multiple studies have documented Nifedipine's anticancer properties across various cancer cell types, to our knowledge, the literature lacks research specifically investigating nifedipine's cytotoxic impact on AGS cells. The present study revealed that the proliferation of gastric cancer cells was notably inhibited upon exposure to Nifedipine, a selective L-type calcium channel blocker, in a dose-dependent fashion. Mechanistic analysis indicated that this compound induced a marked elevation in TOS levels while concurrently diminishing TAS levels within AGS gastric cancer cells. These findings suggest that Nifedipine's inhibition of calcium channels may exert substantial anticancer effects by disrupting the oxidant-antioxidant balance in favor of oxidation. This oxidative shift is associated with the generation of free radicals, ultimately leading to cancer cell demise. Calcium ions (Ca^{2+}) and their associated intracellular calcium channels are primarily recognized for their roles in regulating cardiovascular and neural function. However, emerging evidence increasingly implicates these channels as key contributors to the initiation and progression of tumor development [6]. In connection with this, in cells, calcium channels play several roles in physiological and pathological processes, including cancer. Calcium channels are also frequently abnormally expressed in various cancer cells and are involved in the proliferation, cell cycle progression, and survival [23]. A growing body of research underscores the pivotal role of calcium homeostasis in established oncogenic pathways [7]. Numerous studies have demonstrated the capacity of calcium signaling to drive the progression of various cancers, including glioma [24], prostate cancer [25], and breast cancer [26], through the activation of STAT3, a key transcriptional regulator in cancer. Mechanistically, calcium signals stimulate the production of reactive

oxygen species (ROS) within mitochondria, which in turn triggers the phosphorylation and nuclear translocation of STAT3 [7]. Interfering with calcium signaling presents a potential strategy for enhancing antitumor immune responses. The MAPK pathway, another critical cellular signaling cascade, is interconnected with calcium signaling. A recent study demonstrated that downregulating KCNN4, a calcium-activated potassium channel, led to decreased expression of both MAPK and MMP-9 within cancer cells [7,27]. A significant finding is that numerous calcium channel blockers (CCBs), encompassing phenylalkylamine, dihydropyridine, and benzothiazepine subtypes, as well as other calcium antagonists, can effectively hinder drug efflux mediated by P-glycoprotein, thereby functioning as modulators of multidrug resistance (MDR) [28]. For instance, verapamil, a calcium channel blocker of the L-type, has been demonstrated to reduce P-gp expression in A704 (human kidney adenocarcinoma) and Caki-1 (human renal cancer) cell lines, consequently reversing multidrug resistance in renal cell carcinoma [29]. Furthermore, pretreatment with verapamil has demonstrated the ability to effectively counteract multidrug resistance to doxorubicin in ovarian cancer cells [30]. The calcium channel blocker nifedipine has been shown to inhibit the progression and immune evasion tactics of colorectal cancer by obstructing the nuclear translocation of NFAT2 [6]. Nifedipine has demonstrated the ability to amplify the antitumor efficacy of cisplatin across various human glioblastoma cell lines, including those exhibiting sensitivity, resistance, and even multidrug resistance to cisplatin [10]. Diltiazem has been shown to counteract the development of multidrug resistance induced by docetaxel and vincristine in human lung cancer cell lines [11]. Calcium channel blockers (CCBs) have also shown promise as standalone treatments. For

instance, diltiazem has been observed to suppress the proliferation and migratory behavior of hepatocellular cells in laboratory conditions [12,31]. Collectively, these findings suggest that disrupting calcium signaling could be a promising anticancer therapeutic approach. Carboxyamidotriazole (CAI), a compound that inhibits non-voltage-gated calcium channels and their associated signaling pathways, exemplifies this strategy [6]. Consistent with previous findings, our research demonstrated that nifedipine exhibits anti-proliferative properties against AGS gastric cancer cells. Oxidative stress, a common factor in numerous diseases, is significantly implicated in cancer progression. The body's reactive oxygen species (ROS) levels are pivotal in cancer development and metastasis. While moderate ROS can stimulate angiogenesis, metastasis, and cell survival by activating specific signaling pathways within the tumor microenvironment, excessive ROS can induce cancer cell apoptosis. This underscores the critical role of ROS concentration in determining cancer initiation or cell death [18]. The mechanism of action for many anticancer drugs involves inducing oxidative stress, believed to be the primary cause of extensive cellular macromolecular damage. These drugs target critical cellular components, including proteins, lipids within membranes, and DNA [32]. Total Oxidative Status (TOS) is a commonly used biomarker to evaluate the body's overall oxidative state [33]. Conversely, Total Antioxidant Status (TAS) is employed to assess the body's overall antioxidant capacity [19,34]. This study aimed to determine if nifedipine exerts a cytotoxic effect by altering TAS and TOS levels. Our results demonstrated that a 24-hour nifedipine treatment significantly increased TOS levels while simultaneously reducing TAS levels compared to control cells. These findings suggest that nifedipine induces oxidative stress in AGS cells, as evidenced by the elevated TOS and decreased TAS values. In conclusion, our study provides evidence that nifedipine, a commonly used calcium channel blocker, exhibits significant anti-proliferative effects on AGS gastric cancer cells. This effect is likely mediated through the induction of oxidative stress, as evidenced by the significant increase in TOS and decrease in TAS levels. These findings suggest that targeting calcium signaling pathways may represent a promising strategy for the development of novel anticancer therapies. While our study primarily focused on the antiproliferative effects of nifedipine and its association with oxidative stress in AGS cells, a more comprehensive understanding of the underlying mechanisms would benefit from further investigation. A cell cycle analysis, for instance, could provide additional valuable insights into the precise mode of action of nifedipine. Although such an analysis was beyond the scope of the current study due to resource limitations, our findings strongly suggest that nifedipine

exerts its anti-cancer effects by inducing oxidative stress and inhibiting cell proliferation. Future studies could delve deeper into these molecular mechanisms by incorporating cell cycle analysis and exploring the involvement of specific cell cycle regulatory proteins. Such investigations could provide further evidence supporting the therapeutic potential of nifedipine in the management of gastric cancer and guide the development of more targeted and effective treatment strategies.

Conflicts of interest

There are no conflicts of interest in this work.

Acknowledgments

The authors would like to acknowledge the indispensable support and infrastructure provided by the CUTFAM Research Center at the School of Medicine, Sivas Cumhuriyet University, Sivas, Turkey, which were crucial to the successful completion of this research. This study was funded by Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (2209-A Student Project, TÜBİTAK, NO. 1919B012221382, Turkey).

References

- [1] Ergül M., T-type Calcium Channel Blocker, NNC 55-0396, Suppress Cell Proliferation and Promote Apoptosis in SNU-1 Gastric Cancer Cells, *Cumhuriyet Science Journal*, 44 (2023) 254–7.
- [2] Joshi SS, Badgwell BD., Current treatment and recent progress in gastric cancer., *A Cancer Journal for Clinicians*, 71 (2021) 264–79.
- [3] Harada K, Sakamoto N, Ukai S, Yamamoto Y, Pham QT, Taniyama D, et al., Establishment of oxaliplatin-resistant gastric cancer organoids: importance of myoferlin in the acquisition of oxaliplatin resistance, *Gastric Cancer*, 24 (2021) 1264–77.
- [4] Gonçalves JCR, Coulidiati TH, Monteiro AL, de Carvalho-Gonçalves LCT, de Oliveira Valença W, de Oliveira RN, et al., Antitumoral activity of novel 1,4-naphthoquinone derivative involves L-type calcium channel activation in human colorectal cancer cell line. *Journal of Applied Biomedicine*, 14 (2016) 229–34.
- [5] Monteith GR, Davis FM, Roberts-Thomson SJ., Calcium channels and pumps in cancer: Changes and consequences, *Journal of Biological Chemistry*, 287 (2012) 31666–73.
- [6] Wu L, Lin W, Liao Q, Wang H, Lin C, Tang L, et al., Calcium Channel Blocker Nifedipine Suppresses Colorectal Cancer Progression and Immune Escape by Preventing NFAT2 Nuclear Translocation, *Cell Reports*, 33 (2020) 108327.
- [7] Wu L, Lian W, Zhao L., Calcium signaling in cancer progression and therapy, *The FEBS Journal*, 288 (2021) 6187–205.
- [8] Phan NN, Wang C-Y, Chen C-F, Sun Z, Lai M-D, Lin Y-C., Voltage-gated calcium channels: Novel targets for cancer therapy, *Oncology Letters* 14 (2) (2017) 2059–74.
- [9] Ferrari R, Pavasini R, Camici PG, Crea F, Danchin N, Pinto F, et al., Anti-anginal drugs—beliefs and evidence: systematic review covering 50 years of medical treatment, *European*

- Heart Journal*, 40 (2019) 190–4.
- [10] Kondo S, Yin D, Morimura T, Kubo H, Nakatsu S, Takeuchi J., Combination therapy with cisplatin and nifedipine induces apoptosis in cisplatin-sensitive and cisplatin-resistant human glioblastoma cells, *British Journal of Cancer*, 71 (1995) 282–9.
- [11] Chiu LY, Ko JL, Lee YJ, Yang TY, Tee YT, Sheu GT., L-type calcium channel blockers reverse docetaxel and vincristine-induced multidrug resistance independent of ABCB1 expression in human lung cancer cell lines, *Toxicology Letters*, 192 (2010) 408–18.
- [12] Guo R, Huang X, Jin X, Yang J., Diltiazem inhibits proliferation and motility of hepatocellular cells in vitro by downregulating calcium-activated chloride channel TMEM16A, *Journal of Southern Medical University*, 38 (2018) 818–23.
- [13] Panneerandian P, Rao DB, Ganesan K., Calcium channel blockers lercanidipine and amlodipine inhibit YY1/ERK/TGF- β mediated transcription and sensitize the gastric cancer cells to doxorubicin, *Toxicology in Vitro*, 74 (2021) 105152
- [14] Yang JL, Friedlander ML., Effect of nifedipine in metastatic colon cancer with DNA mismatch repair gene defect. *The Lancet*, 357 (9270) (2001) 1767-1768.
- [15] Taştemur Ş, Hacısüleyman L, Karataş Ö, Yulak F, Ataseven H. Anticancer activity of sinapic acid by inducing apoptosis in HT-29 human colon cancer cell line, *Canadian Journal of Physiology and Pharmacology* 101 (7) (2023) 361-368.
- [16] Ataseven D, Taştemur Ş, Yulak F, Karabulut S, Ergul M, GSK461364A suppresses proliferation of gastric cancer cells and induces apoptosis, *Toxicology in Vitro*, 90 (2023) 105610
- [17] Ciltas AC, Karabulut S, Sahin B, Filiz AK, Yulak F, Ozkaraca M, et al., FGF-18 alleviates memory impairments and neuropathological changes in a rat model of Alzheimer's disease, *Neuropeptides* 101 (2023) 102367.
- [18] Yulak F, Filiz AK, Joha Z, Ergul M., Mechanism of anticancer effect of ETP-45658, a PI3K/AKT/mTOR pathway inhibitor on HT-29 Cells, *Medical Oncology*, 40 (2023) 1–14.
- [19] Erel O., A novel automated direct measurement method for total antioxidant capacity using a new generation, more stable ABTS radical cation, *Clinical Biochemistry*, 37 (2004) 277–85.
- [20] Taskiran AS, Ergul M, Gunes H, Ozturk A, Sahin B, Ozdemir E., The Effects of Proton Pump Inhibitors (Pantoprazole) on Pentylentetrazole-Induced Epileptic Seizures in Rats and Neurotoxicity in the SH-SY5Y Human Neuroblastoma Cell Line, *Cellular and Molecular Neurobiology*, 41 (2021) 173–83.
- [21] Erel O., A new automated colorimetric method for measuring total oxidant status. *Clinical Biochemistry*, 38 (2005) 1103–11.
- [22] Taştemur Y, Gumus E, Ergul M, Ulu M, Akkaya R, Ozturk A, et al., Positive effects of angiotensin-converting enzyme (ACE) inhibitor, captopril, on pentylentetrazole-induced epileptic seizures in mice, *Tropical Journal of Pharmaceutical Research*, 19 (2020) 637–43.
- [23] Dziegielewska B, Gray LS, Dziegielewska J., T-type calcium channels blockers as new tools in cancer therapies, *Pflugers Archiv European Journal of Physiology*, 466 (2014) 801–10.
- [24] Li M, Ruan B, Wei J, Yang Q, Chen M, Ji M, et al., ACYP2 contributes to malignant progression of glioma through promoting Ca²⁺efflux and subsequently activating c-Myc and STAT3 signals, *Journal of Experimental and Clinical Cancer Research*, 39 (2020) 1–14.
- [25] Henderson VM, Hawsawi O, Burton LJ, Campbell T, Trice K, Dougan J, et al., Cancer-bone microenvironmental interactions promotes STAT3 signaling, *Molecular Carcinogenesis*, 58 (2019) 1349–61.
- [26] Li L, Zhang R, Liu Y, Zhang G., ANXA4 Activates JAK-STAT3 Signaling by Interacting with ANXA1 in Basal-Like Breast Cancer, *DNA and Cell Biology*, 39 (2020) 1649–56.
- [27] Li QT, Feng YM, Ke ZH, Qiu MJ, He XX, Wang MM, et al., KCNN4 Promotes Invasion and Metastasis through the MAPK/ERK Pathway in Hepatocellular Carcinoma, *J Investig Med.*, 68 (2020) 68–74.
- [28] Hamilton G, Theyer G, Baumgartner G., Calcium antagonists as modulators of multi-drug resistant tumor cells. *Wiener Medizinische Wochenschrift*, 143 (1993) 526–32.
- [29] Yu DS, Ma CP, Chang SY., Verapamil modulation of multidrug resistance in renal cell carcinoma, *Journal of the Formosan Medical Association*, 99 (2000) 311–6.
- [30] Zheng W, Li M, Lin Y, Zhan X., Encapsulation of verapamil and doxorubicin by MPEG-PLA to reverse drug resistance in ovarian cancer, *Biomedicine & Pharmacotherapy*, 108 (2018) 565–73..
- [31] Guo R, Jin X, Tian Y, Huang X, Li Z, Yang J., Calcium channel blocker diltiazem transiently inhibits migration and up-regulates metadherin expression in hepatocellular carcinoma cells in vitro, *Journal of Southern Medical University*, 39 (2019) 298–303.
- [32] Matos CP, Adiguzel Z, Yildizhan Y, Cevatemre B, Onder TB, Cevik O, et al., May iron(III) complexes containing phenanthroline derivatives as ligands be prospective anticancer agents?, *European Journal of Medicinal Chemistry*, 176 (2019) 492–512.
- [33] Yulak F, Üngür B., Investigation of the Effect of Indatraline on Oxidative Damage Induced by Hydrogen Peroxide in C6 Glioma Cell Line, *Cumhuriyet Science Journal*, 44 (2023) 645–9.
- [34] Çiltaş AÇ, Gündoğdu S, Yulak F., Levetiracetam Protects Against Glutamate-Induced Excitotoxicity in SH-SY5Y Cell Line, *International Journal of Nature and Life Sciences*, 6 (2022) 142–50.

Investigation of Cytotoxic Activity of *Anthriscus nemorosa* (M.Bieb.) Spreng. on Lung Cancer Cells

Kevser Taban ^{1,a,*}

¹ Department of Pharmacognosy, Faculty of Pharmacy, Sivas Cumhuriyet University, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 17/09/2024

Accepted: 25/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Medicinal plants are considered an important source of human health due to their therapeutic potential against various diseases including cancer. Cancer is a life-threatening disease characterized by uncontrolled cell growth and abnormal signaling processes. The incidence of cancer in society is increasing day by day. The search for active biological resources is important for the discovery of new anticancer drugs. *Anthriscus nemorosa* (M.Bieb.) Spreng. (Apiaceae) is a medicinal plant naturally distributed in Türkiye and traditionally used as food and against various diseases. This study investigated the methanol extracts of aerial parts and roots of *A. nemorosa* for cytotoxicity on lung cancer (A549) and non-cancerous (L929) cell lines. According to the results, both extracts showed significant dose-dependent cytotoxic effects on lung cancer cells. IC₅₀ value was recorded as 8.29 µg/mL in the aerial parts extract. On the other hand, it was recorded as 3.57 µg/mL in the roots extract. Furthermore, the selectivity indexes were calculated as 5.93 and 3.38 for aerial parts and roots extracts, respectively. In light of the findings, it has been concluded that *A. nemorosa* deserves further anticancer research.

Keywords: *Anthriscus nemorosa*, Apiaceae, Lung cancer, Antiproliferative activity.

^a kevsertaban@cumhuriyet.edu.tr  <https://orcid.org/0000-0001-8620-6402>

Introduction

Plants have had an important place for humans as food and medicine since ancient times. Especially, medicinal plants are considered an important source of human health due to their therapeutic capabilities against various diseases [1]. Identification of cytotoxic natural compounds and medicinal plants will be useful for cancer treatment. The genus *Anthriscus* Pers. (Apiaceae) has 16 species with traditional use worldwide [2]. The genus is represented by 8 species in Türkiye: *A. caucalis* M. Bieb., *A. cerefolium* (L.) Hoffm., *A. kotschyi* Fenzl ex Boiss., *A. lamprocarpa* Boiss., *A. nemorosa* (M.Bieb.) Spreng., *A. ruprechtii* Boiss., *A. sylvestris* (L.) Hoffm., and *A. tenerrima* Boiss. & Spruner [3]. *A. nemorosa* is known as "gimigimi, peçek" and fruits of the plant are used in the treatment of inflammation, gastrointestinal disorders, and rheumatism [4, 5]. *A. nemorosa* is also used as food in the eastern and southeastern parts of Anatolia [6].

There are limited biological activity studies on *A. nemorosa* in the literature. The essential oil of *A. nemorosa* was reported to exhibit antimicrobial activity against *Bacillus subtilis* and *Candida albicans* [7]. In another study, *A. nemorosa* essential oil inhalation was found to improve memory formation and showed anxiolytic and antidepressant effects in treatment groups [2]. Karakaya et al. reported that the essential oil and ethyl acetate fraction of *A. nemorosa* root had high total phenolic content and DPPH radical scavenging activity. The essential oil of the plant also exhibited high butyryl cholinesterase enzyme (BuChE) inhibition in the same study [8]. The essential oil of *A. nemorosa* and farnesene

in the chemical composition of the essential oil was reported to be effective against *Trypanosoma brucei* [9]. Forouhandeh et al. investigated the cytotoxic activity of *n*-hexane, dichloromethane, and methanol extracts of *A. nemorosa* on breast cancer cells. *n*-Hexane and 60%, 80%, and 100% fractions of *n*-hexane extract were found to inhibit the proliferation of MCF-7 cells in a dose-dependent manner [10]. *A. sylvestris* and its main compound, anthricin, were found to have strong cytotoxic effects on various cancer cell lines [11-14]. Therefore, *A. nemorosa* may be a promising source against lung cancer due to its phylogenetic relationship.

In this study, the methanol extracts of the roots and aerial parts of *A. nemorosa* were examined on lung cancer and healthy cells using the MTT method. In this way, it was aimed to introduce new natural strategies with cytotoxic effects against lung cancer to the literature.

Materials and Methods

Plant Material

Anthriscus nemorosa (M.Bieb.) Spreng was collected from Karayün, Sivas in May 2023. The plant was identified by Assoc. Prof. Mehmet Ufuk Özbek from Gazi University, Faculty of Science, Department of Biology. A voucher specimen was deposited at the Herbarium of Faculty of Pharmacy, Gazi University, Ankara (GUEF No: 3846).

Preparation of the Extracts

The air-dried roots and aerial parts of the plant were separated and ground to powder using a laboratory-type mill. The plant materials were macerated with methanol at room temperature for 24 h. Then, the extracts were filtered and fresh methanol was added to the residue. These extraction processes were repeated three times (24 h x 3). The filtrates were combined and the solvent was removed by rotary evaporation at 40°C. Extracts obtained from the process were held at +4°C until they were used.

Cell Lines and Cell Culture

Adenocarcinomas of human alveolar basal epithelial cell line (A549, ATCC) and healthy mouse fibroblast cell line (L929, ATCC) were used for the cytotoxicity test of extracts. The cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM, Capricorn Scientific) supplemented with 10% fetal bovine serum (FBS, Serana Europe), 1% L-glutamine, and 1% penicillin-streptomycin (Serana Europe). The cells were incubated at 37 °C under 5% CO₂ humidified atmosphere.

Cell Viability Assay

The cytotoxic activity of the extracts and cisplatin were determined by 3-(4,5-Dimethyl thiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. MTT assay called the "gold standard" of cell viability tests is based on measuring metabolic activity [15, 16]. The cells reaching 80% confluence were seeded in 96-well plates (6 × 10³ cells/well). After 24 h, cells were treated with five concentrations of the extracts (1-100 µg/mL) in DMSO. Cisplatin was used as a positive control and tested cells were also subjected to cisplatin (1-100 µM). The plates were incubated with test materials for 48 h. After treatment, the supernatant was removed and fresh colorless DMEM and 50 µL MTT solution were added to each well. The cells were maintained at 37 °C for 3 h. DMSO was used to dissolve the resulting formazan crystals. The absorbance was measured at 570 nm using a spectrophotometer (SpectraMax i3x; Molecular Devices, San Jose, CA, USA). Each experiment was conducted in triplicates, and cell viability was indicated as a percentage relative to the control (100% of viability). Data were analyzed using GraphPad Software Prism 8.0 (San Diego, CA, USA; demo version). Nonlinear regression analysis (dose-response) was used to determine the IC₅₀ values. IC₅₀ values were calculated as µg/mL for extracts and as µM for cisplatin.

Selectivity Index

The selectivity index (SI) indicates the selectivity of test materials between cancer and normal cells. SI was calculated by using the formula [(IC₅₀ values of non-cancerous cells)/(IC₅₀ values of cancerous cells)]. The

compounds with SI greater than 1 are more likely to inhibit cancer cells than non-cancerous cells [17]. When the SI value of a compound is greater than 10, it is considered a selective anticancer agent [18].

Results and Discussion

Extract Yields

The extract yields of *A. nemorosa* were given in Table 1.

Table 1. The extract yields of *A. nemorosa*

| Extract | Amount (g) | Yield (% w/w) |
|---------------------------------|------------|---------------|
| <i>A. nemorosa</i> aerial parts | 7.621 | 15.24 |
| <i>A. nemorosa</i> roots | 6.328 | 12.66 |

Activity Results

The roots and aerial parts of methanol extracts of *A. nemorosa* were studied to find new approaches against lung cancer. According to the National Cancer Institute (NCI), the extracts with an IC₅₀ value of less than 20 µg/mL could be considered promising anticancer substances [19]. In this study, *A. nemorosa* exhibited strong anticancer activity on A549 cell lines in a dose-dependent manner. Cell viability results of the extracts and cisplatin on the tested cells are given in Figure 1.

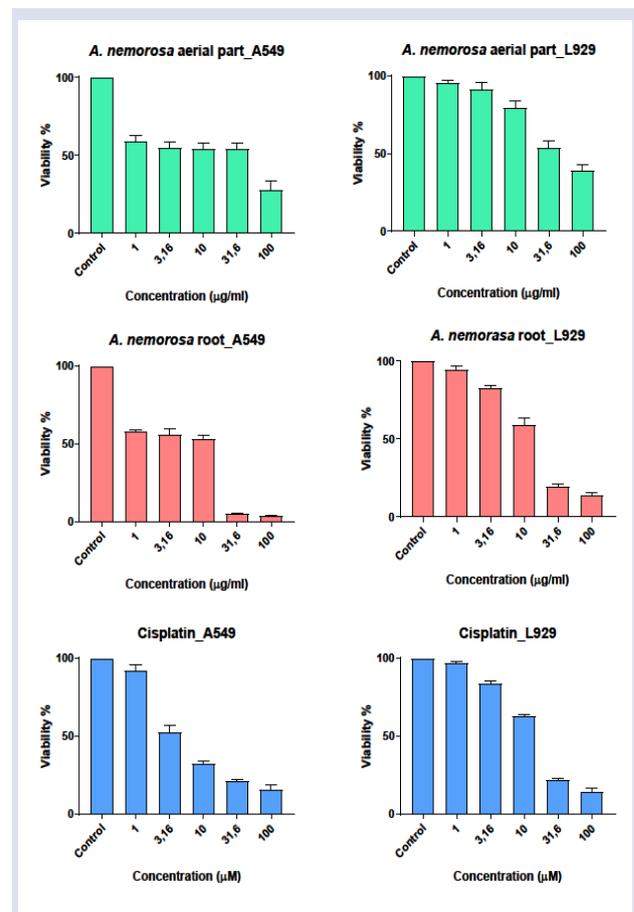


Figure 1. Cell viability inhibition of *A. nemorosa* extracts, and cisplatin against A549, and L929 cell lines after 48 h exposure

Both extracts had an IC₅₀ value of ≤ 20 µg/mL against lung cancer cells. IC₅₀ values of the aerial parts and roots extracts of *A. nemorosa* on A549 were calculated as 8.29 and 3.57 µg/mL, respectively. It was determined as 49.13 µg/mL for aerial parts extract and 12.07 µg/mL for roots extract on L929. The selectivity of both extracts was higher than cisplatin against normal cells. The cytotoxic effect of *A. nemorosa* extracts and cisplatin against A549, and L929 cell lines were shown in Table 2 and Figure 2.

Table 2. Cytotoxic effect of *A. nemorosa* extracts, and cisplatin against A549, and L929 cell lines after 48 h exposure

| | IC ₅₀ ± SEM (µg/mL) | | SI |
|--------------------------------|--------------------------------|--------------|------|
| | A549 | L929 | |
| <i>A. nemorosa</i> aerial part | 8.29 ± 2.57 | 49.13 ± 5.58 | 5.93 |
| <i>A. nemorosa</i> root | 3.57 ± 0.57 | 12.07 ± 0.78 | 3.38 |
| Cisplatin | 4.89 ± 0.61 | 13.80 ± 0.90 | 2.82 |

The results of cisplatin were given as IC₅₀ ± SEM (µM). SEM: Standard error of the mean; SI: Selectivity index

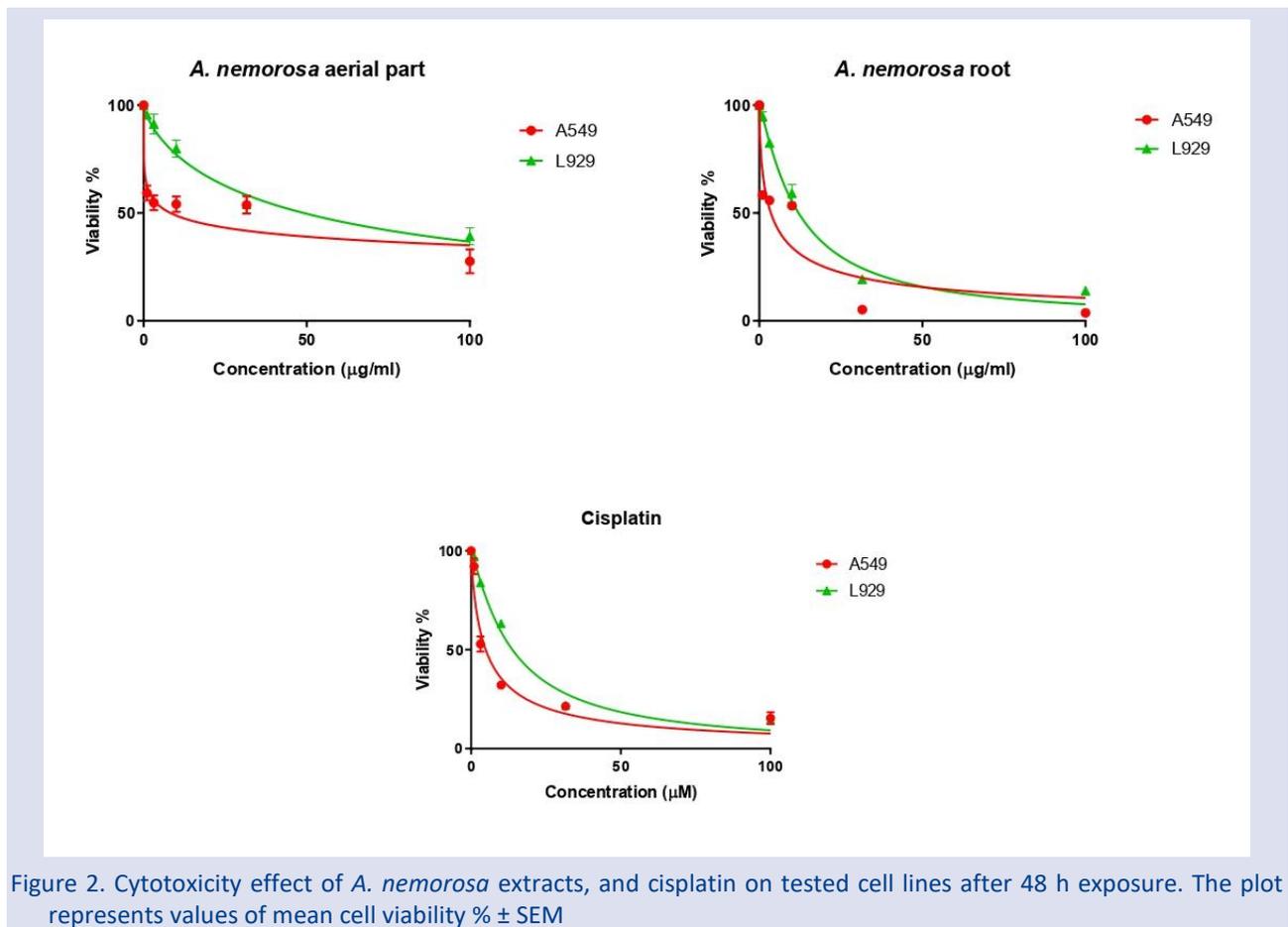


Figure 2. Cytotoxicity effect of *A. nemorosa* extracts, and cisplatin on tested cell lines after 48 h exposure. The plot represents values of mean cell viability % ± SEM

There are a limited number of phytochemical and biological activity studies on *A. nemorosa*. Bagci et al. investigated the effect of *A. nemorosa* essential oil on memory, anxiety, and depression-like behaviors in scopolamine-treated rats. *A. nemorosa* essential oil inhalation improved memory formation and had anxiolytic and antidepressant effects in treatment groups [2]. The root ethyl acetate fraction and essential oil of plant were found to have high total phenolic content with values of 677.31 ve 509.39 GAE g⁻¹. The essential oil of roots showed high butyrylcholinesterase inhibition (88.51%). Major monoterpene of roots and aerial parts of *A. nemorosa* were α-pinene (25.5%), myristicin (10.4%), p-cymene (8.2%), limonene (6.0%), and fatty alcohol 1-heptadecanol (7.5%). In another study, the main components of *A. nemorosa* roots essential oil were identified as n-nonane (12.1%), n-hexadecanol (6.9%), δ-cadinene (6.4%), β-pinene (6%) and germacrene D (5.4%).

In the same study, *A. nemorosa* essential oil was found to be sensitive against *Bacillus subtilis* (MIC= 6.25 µg/mL), and *Candida albicans* (MIC= 50 µg/mL) [7].

There is only one study in the literature on the cytotoxicity evaluation and anticancer potential of *A. nemorosa*. In this study, the cytotoxic effect of n-hexane, dichloromethane, and methanol extracts of aerial parts of *A. nemorosa* was examined on breast cancer cells at the concentration range of 50-800 µg/mL. According to the results, n-hexane and 60%, 80% and 100% fractions of n-hexane extract inhibited the proliferation of MCF-7 cells with IC₅₀ values of 75.63, 22.6, 26.82, 14.71, respectively. The predominant escape composition of extracts was identified as non-terpenoid. The most common ingredient was palmitic acid and non-terpenoids were associated with activity against breast cancer cells [10].

Another species of genus *Anthriscus*, *A. sylvestris*, had cytotoxicity against various cancer cells. The methanol

extract of *A. sylvestris* was reported to have cytotoxicity on human chronic myeloid leukemia K562 cell [14]. Ikeda et al. found that the root and aerial parts of *A. sylvestris* had strong inhibition on the proliferation of gastric adenocarcinoma (MK-1), cervical adenocarcinoma (HeLa), and melanoma (B16F10) cells [13]. The petroleum ether fraction of *A. sylvestris* showed a strong cytotoxic effect on HeLa and human hepatocellular carcinoma (HepG2) cell lines with IC₅₀ values of 18.25 and 36.53 µg/mL, respectively [11]. *n*-Hexane and dichloromethane fractions of *A. sylvestris* inhibited the growth of human gastric adenocarcinoma (AGS) cells [12]. Anthricin (deoxypodophyllotoxin) was reported as one of the main lignans in *A. sylvestris* [20]. This compound had various biological activities, including anticancer, and was considered the most important compound of the plant [21-23]. Anthricin had a five-ring cyclolignan structure and similar to podophyllotoxin. Podophyllotoxin served as a key starting compound for the synthesis of etoposide and teniposide, anticancer drugs used in the treatment of various leukemia and solid tumors. Since podophyllotoxin was synthesized directly from deoxypodophyllotoxin in the biosynthetic pathway, anthricin also gained importance as the starting compound in the synthesis of etoposide and teniposide [22, 24]. Anthricin was reported to have a strong antiproliferative effect on HepG2, osteosarcoma (MG63), melanoma (B16), HeLa, and breast cancer cells (MCF7, MDA-MB-231) [11, 12, 21]. It was reported that anthricin inhibited polymerization via binding directly to microtubules, causing cell cycle arrest at the G2/M phase and apoptosis [23].

Essential oil analysis stands out in the literature on *A. nemorosa*. However, in a study, two lignan lactones namely savinin and nemerosin were isolated from methanol extract of *A. nemorosa* root [25]. Lignan-type compounds such as savinin and nemerosin might be responsible for anticancer activity. Detailed and further phytochemical studies are needed.

Conclusion

The methanol extracts of the roots and aerial parts of *A. nemorosa* showed significant dose-dependent cytotoxic effects on lung cancer cells. IC₅₀ value was recorded as 8.29 µg/mL in the aerial parts extract and it was recorded as 3.57 µg/mL in the roots extract. Furthermore, the selectivity indexes were calculated as 5.93 and 3.38 for aerial parts and roots extracts, respectively. According to the NCI guidelines, the methanol extract of *A. nemorosa* exhibited promising antiproliferative activity on lung cancer cells. Moreover, the extracts had lower cytotoxicity against normal cells. The selective and potent activity results of *A. nemorosa* indicate that this compound deserves further anticancer research.

Considering our results, *A. nemorosa* could be determined as a promising candidate for cancer treatment. In further studies, it is recommended to

discover cytotoxic compounds against lung cancer by activity-guided isolation studies on *A. nemorosa*.

Conflicts of interest

The author declares no conflict of interest.

Acknowledgement

This study was supported by the Scientific Research Projects Coordination Unit of Sivas Cumhuriyet University (ECZ-2023-091).

Many thanks to Cumhuriyet University Faculty of Medicine Research Center (CÜTFAM) for laboratory equipment support.

References

- [1] Al Qaisi Y., Alfarrayeh I., Alsarayreh A., Khleifat K., Abu-Nwas N., Assessment of antioxidant potential, cytotoxicity, and anticancer activity of methanolic extracts from selected wild medicinal plants, *Phytomed. Plus.*, 4 (2) (2024) 100534.
- [2] Bağcı E., Aydın E., Ungureanu E., Hritcu L. *Anthriscus nemorosa* essential oil inhalation prevents memory impairment, anxiety and depression in scopolamine-treated rats, *Biomed. Pharmacother.*, 84 (2016) 1313-1320.
- [3] Menemen Y., *Anthriscus* Pers. In: Güner A., Aslan S., Ekim T., Vural M., Babaç M.T., (Eds). Türkiye Bitkileri Listesi (Damarlı Bitkiler). İstanbul: Nezahat Gökyiğit Botanic Garden Publications, (2012) 62-64.
- [4] Altundag E., Ozturk M., Ethnomedicinal studies on the plant resources of East Anatolia, Turkey, *Procedia Soc. Behav. Sci.*, 19 (2011) 756-777.
- [5] Gairola S., Sharma J., Bedi Y.S., A cross-cultural analysis of Jammu, Kashmir and Ladakh (India) medicinal plant use, *J. Ethnopharmacol.*, 155 (2) (2014) 925-986.
- [6] Tarakçı Z., Coskun H., Tunçtürk Y., Some properties of fresh and ripened herby cheese, a traditional variety produced in Turkey, *Food Technol. Biotechnol.*, 42 (1) (2004) 47-50.
- [7] Pavlović M., Petrović S., Milenković M., Couladis M., Tzakou O., Niketić M., Chemical composition and antimicrobial activity of *Anthriscus nemorosa* root essential oil, *Nat. Prod. Commun.*, 6 (2) (2011).
- [8] Karakaya S., Yılmaz S.V., Koca M., Demirci B., Sytar O., Screening of non-alkaloid acetylcholinesterase inhibitors from extracts and essential oils of *Anthriscus nemorosa* (M. Bieb.) Spreng. (Apiaceae), *S. Afr. J. Bot.*, 125 (2019) 261-269.
- [9] Baldassarri C., Falappa G., Mazzara E., Acquaticci L., Ossoli E., Perinelli D.R., Antitrypanosomal activity of *Anthriscus nemorosa* essential oils and combinations of their main constituents, *Antibiotics*, 10 (11) (2021) 1413.
- [10] Forouhandeh H., Zarchini M., Safarzadeh E., Molavi O., Asgharian P., Tarhriz V., Evaluation of the cytotoxic activity of *Anthriscus nemorosa* on breast cancer cells, *J. Contemp. Med. Sci.*, 7 (4) (2021) 217-226.
- [11] Chen H., Jiang H.Z., Li Y.C., Wei G.Q., Geng Y., Ma C.Y., Antitumor constituents from *Anthriscus sylvestris* (L.) Hoffm., *Asian Pac. J. Cancer Prev.*, 15 (6) (2014) 2803-2807.
- [12] Cho E.J., Choi J.M., Kim H.M., Choi K., Ku J., Park K.W., Antibacterial activity and protective effect against gastric cancer by *Anthriscus sylvestris* fractions, *Hortic. Environ. Biotechnol.*, 54 (2013) 326-330.

- [13] Ikeda R., Nagao T., Okabe H., Nakano Y., Matsunaga H., Katano M., Mori M., Antiproliferative Constituents In Umbelliferae Plants. III. Constituents in the Root and the Ground Part of *Anthriscus sylvestris* Hoffm., *Chem. Pharm. Bull.*, 46 (5) (1998) 871-874.
- [14] Lim Y.H., Leem M.J., Shin D.H., Chang H.B., Hong S.W., Moon E.Y., Cytotoxic constituents from the roots of *Anthriscus sylvestris*, *Arch. Pharm. Res.*, 22 (1999) 208-212.
- [15] Emerce E., Taban Akça K., Cytotoxic Activity Methods. In: *Methods for Preclinical Evaluation of Bioactive Natural Products*, (2023) 149-176.
- [16] van Tonder A., Joubert A.M., Cromarty A.D., Limitations Of The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay when compared to three commonly used cell enumeration assays, *BMC Res. Notes*, 8 (1) (2015) 1-10.
- [17] Badisa R.B., Darling-Reed S.F., Joseph P., Cooperwood J.S., Latinwo L.M., Goodman C.B., Selective cytotoxic activities of two novel synthetic drugs on human breast carcinoma MCF-7 cells, *Anticancer Res.*, 29(8) (2009) 2993-2996.
- [18] Acar Çevik U., Celik I., Görgülü Ş., Şahin Inan Z.D., Bostancı H.E., Özkay Y., Kaplacıklı Z.A., New benzimidazole-oxadiazole derivatives as potent VEGFR-2 inhibitors: synthesis, anticancer evaluation, and docking study, *Drug Dev. Res.*, 85 (4) (2024) e22218.
- [19] Suffness M., Assays related to cancer drug discovery. In: Hostettmann K. (Ed) *Methods plant biochem: assays bioactivity*. London: Academic Press, (1990) 71-133
- [20] Noguchi T., Kawanami M., Studies on the constituent of *Anthriscus sylvestris*. Hoffm., *Yakugaku Zasshi*, 60 (1940) 2809-2814.
- [21] Jung C.H., Kim H., Ahn J., Jung S.K., Um M.Y., Son K.H., Anthricin isolated from *Anthriscus sylvestris* (L.) Hoffm. inhibits the growth of breast cancer cells by inhibiting Akt/mTOR Signaling, and its apoptotic effects are enhanced by autophagy inhibition, *Evid. Based. Complement. Alternat. Med.*, (2013).
- [22] Khaled M., Jiang Z.Z., Zhang L.Y., Deoxypodophyllotoxin: a promising therapeutic agent from herbal medicine, *J. Ethnopharmacol.*, 149 (1) (2013) 24-34.
- [23] Olaru O.T., Nițulescu G.M., Orțan A., Băbeanu N., Popa O., Ionescu D., Dinu-Pirvu C.E., Polyphenolic content and toxicity assessment of *Anthriscus sylvestris* Hoffm., *Rom. Biotechnol. Lett.*, 22 (6) (2016) 12054.
- [24] Berežni S., Mimica-Dukić N., Domina G., Raimondo F.M., Orčić D., *Anthriscus sylvestris*-noxious weed or sustainable source of bioactive lignans, *Plants*, 13(8) (2024) 1087.
- [25] Turabelidze D.G., Mikaya G.A., Kemertelidze E.P., Vul'fson N.S., Lignan lactones from the roots of wild chervil *Anthriscus nemerosa* (Bieb.) Spreng., *Bioorg. Khim.*, 8(5) (1982).

Chemical Composition of Taraxacum Officinale Extracts and Anti-Proliferative Effect on A549 and HT-29 Cell Lines Through Regulating the Expression of Apoptosis-Related Genes

Serkan Kapancik^{1,a,*}¹ Department of Biochemistry, Cumhuriyet University School of Medicine, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 22/06/2024

Accepted: 04/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

The use of medicinal plants is still of great importance as an alternative treatment method in neurodegenerative diseases, cardiovascular patients and cancer. Taraxacum officinale plant is one of the medical plants used in the treatment of many diseases as diuretic, hypolipidemic and antidiabetic thanks to the active components it contains. In order to examine the effect of Taraxacum officinal on proliferation and apoptosis in cancer, in our study, we determined the chemical composition of the Taraxacum officinal plant by gas chromatography-mass spectrometry (GC-MS) and its antioxidant status using a commercial kit. Additionally, we determined the anti-cancer effect of water, methanol and chloroform extracts of the Taraxacum officinal plant on A549 and HT-29 cancer cell lines with the help of the MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide) method. Finally, we aimed to determine the relationship between the chloroform extract of the Taraxacum officinal plant, which has the highest anti-cancer properties, and apoptosis by determining the expressions of BAX, BCL-2 and Caspase-3, which are among the genes associated with apoptosis, in A549 and HT-29 cancer cell lines, with the help of Real-time Polymerase Chain Reaction (RT-PCR) analysis. We determined that the Taraxacum officinal plant has an anti-cancer effect by increasing caspase-3 expressions in A549 lung cancer cell lines and by increasing caspase-3 and BAX expressions and decreasing BCL-2 expression in HT-29 colon cancer cell lines. In light of the data we obtained, we can say that the Taraxacum officinal plant has anti-cancer properties that cause a decrease in cell proliferation by regulating the expression of genes related to apoptosis.

Keywords: Cancer, GC-MS, Taraxacum officinal, Apoptosis, Gene expression.^a serkankapancik@gmail.com^{ib} <https://orcid.org/0000-0003-3019-4275>

Introduction

Medicinal plants have been used by humanity since ancient times to treat diseases through the active ingredients they contain [1]. Today, as a result of developments in the field of pharmacology, although non-plant-based synthetic drugs are produced for the treatment of diseases, active substances obtained from medicinal plants are still widely used [2]. In particular, the use of medicinal plants remains important as an alternative treatment method in neurodegenerative diseases, cardiovascular patients and cancer [3-5].

Taraxacum officinale plant is one of the medical plants used in the treatment of many diseases due to the active components it contains. Taraxacum officinale plant is a perennial 40 cm tall plant and grows in temperate and hot climates. It is a wild herbaceous plant that blooms in yellow and orange colors [6].

Considering the studies on the medical benefits of the Taraxacum officinale plant, it has been determined that the hydroethanolic extract of Taraxacum officinale has a diuretic effect on humans [7]. In another study, it was shown that water and methanol extracts of the plant inhibit the activities of α -amylase and α -glucosidase enzymes under in vitro conditions, thus having anti-diabetic effects [8]. When the root and leaf of Taraxacum

officinale were included in the diet of rabbits fed with high cholesterol, it was determined that the activities of enzymes that produce antioxidant effects increased and lipid levels decreased in plasma samples taken from rabbits. Based on these results, it has been reported that Taraxacum officinale has antioxidant and hypolipidemic effects and may contribute to the treatment of cardiovascular diseases by providing protection against atherosclerosis [9]. However, it was determined that hepatotoxicity in mice in which liver damage was caused by the administration of acetaminophen and in mice given extracts obtained from Taraxacum officinale leaves was significantly reduced through the antioxidant properties of Taraxacum officinale [10]. In studies on the role of Taraxacum officinal in cancer, it has been determined that it reduces cell proliferation in breast and prostate cancer [11], has an antitumor effect in pediatric cancers [12], and its hydroethanolic extract induces apoptosis in breast cancer cells [13].

The emergence of cancer is associated with a shift in the balance of proliferation and apoptosis, which is maintained in healthy cells, towards proliferation. Apoptosis is a programmed form of cell death in cells that has genetic basis [14]. There are many genes associated with apoptosis and the proteins these genes encode.

Caspase-3, BCL-2 and BAX are among the proteins that have the most important role in the regulation of apoptosis. While caspase-3 and BAX mediate the induction of apoptosis, BCL-2 plays a role in inhibiting it [15].

In order to examine the effect of *Taraxacum officinale* on proliferation and apoptosis in cancer, in our study, we aimed to determine the anticancer effect of water, methanol and chloroform extracts of the *Taraxacum officinale* plant, whose chemical composition and antioxidant status we analyzed, on A549 and HT-29 cancer cell lines, and to determine the anti-cancer effect of the *Taraxacum officinale* plant exposed to chloroform extract. We aimed to determine their relationship with apoptosis by determining the expressions of BAX, BCL-2 and Caspase-3 in A549 and HT-29 cancer cell lines.

Materials and Methods

Collection of Plant Material

Taraxacum officinale plant was collected in Sivas province of Turkey in 2023 and stored in the dark compartment under room temperature conditions.

Extract Preparation

After the *Taraxacum officinale* plant was collected, it was divided into 10 gram portions. Then, 10 grams of plant samples were ground. Water, methanol and chloroform extracts were prepared from the ground plant samples. The methanol extract of the plant was prepared in the Soxhlet extractor under 60°C and 2-hour extraction conditions. The extract was concentrated using vacuum below 45°C. A portion of the extract, from which methanol was removed, was dissolved in chloroform and water. The resulting fractions were evaporated and the remaining part was stored in the dark at 4°C [16].

Gas Chromatography-mass Spectroscopic Analysis

Shimadzu GCMS-QP2010 Ultra Gas Chromatograph Mass Spectrometer (Shimadzu, Kyoto, Japan) device was used in GC-MS analyzes to determine the chemical content of the plant. The system setup includes a Flex 2 autosampler (EST Analytical, Ohio, United States) for automatic HS-SPME sampling. Data were analyzed with GC-MS Postrun Analysis software (ver. 4.53) [16].

Cell Culture Maintenance

Preparation of cell culture

A549 and HT-29 cells were grown in medium containing Dulbecco's Modified Eagle Medium (DMEM). The medium also contained 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin of the total volume. Cells were grown under 37°C and 5% CO₂ culture conditions [17].

MTT method

Water, methanol and chloroform extracts of the *Taraxacum officinale* plant were applied to A549 and HT-29 cells for 24, 48 and 72 hours and IC₅₀ values were determined via the MTT method. For this, cells were

seeded in 96-well plates at 1 x 10⁵/ml and incubated overnight. Then, the % viability values in each well were determined in accordance with the MTT kit protocol. % viability graphs were drawn with GraphPad Prism and statistical analysis was performed [16]. Images of A549 and HT-29 cells exposed to *Taraxacum officinale* plant extracts for 24 hours were recorded through a microscope.

Expression analysis of BAX, BCL-2 and Caspase-3, Genes Associated with Apoptosis

cDNA was synthesized using the kit (Thermo Scientific, USA) from total RNA isolated using the kit (GeneAll Hybrid-R, South Korea) from A549 and HT-29 cells exposed to the IC₅₀ dose of chloroform extract of the *Taraxacum officinale* plant for 24 hours. Using these cDNAs, expression analysis of Caspase-3, BAX and BCL-2 was performed on the RT-PCR device with the help of the SYBR Green qPCR Mastermix kit [18]. Statistical analysis of the data was done with the program (<https://geneglobe.qiagen.com/us/analyze>).

Total Antioxidant Status (TAS) Measurement

The method developed by Erel was used to measure TAS levels in extracts via spectrophotometer. The basis of this method is based on the measurement of the inhibition of the amounts of colored dianisidyl radicals formed by free radical reactions in the Fenton reaction by antioxidants. The higher the active antioxidant capacity in the extracts, the more color formation is suppressed. In this method, the results are expressed as millimolar Trolox equivalents per L extract (mmol Trolox Equiv./L). In our study, TAS was determined by spectrophotometer using a commercial kit in accordance with the manufacturer's instructions (Rel Assay Diagnostic, Türkiye) [19,20].

Results

Chemical Composition of the Plant

The chemical composition was determined by GC-MS analysis of the chloroform extract of the *Taraxacum officinale* plant. It was determined that there are 38 active components in the chloroform extract of the *Taraxacum officinale* plant. In the chloroform extract of *Taraxacum officinale* plant.; Hexadecanoic acid, Phytol, 9-Octadecanamide, Neophytadiene, 7,9-Di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione were determined to be the most abundant active compounds at the concentration of 16.25%, 14.39%, 11.02%, 9.42% and 5.07%, respectively.

MTT Assay

A549 and HT-29 cell lines were exposed to 1000µg/ml, 500µg/ml, 250µg/ml, 100µg/ml, 50µg/ml and 10µg/ml chloroform, methanol, and water extracts of *Taraxacum officinale* plant for 24, 48, and 72h to determine IC₅₀ effective doses (Figures 1-3).

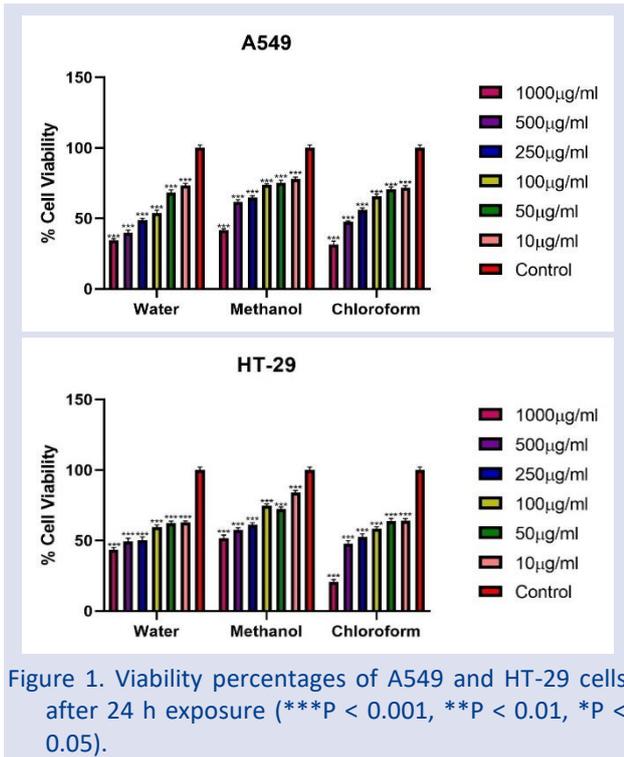


Figure 1. Viability percentages of A549 and HT-29 cells after 24 h exposure (***P* < 0.001, ***P* < 0.01, **P* < 0.05).

In A549 cell lines, the IC50 effective dose of the water extract was determined as 195.30 µg/ml, the IC50 effective dose of the methanol extract was 977.90 µg/ml, and the IC50 effective dose of the chloroform extract was 322.20 µg/ml for the 24th hour. In HT-29 cell lines, the IC50 effective dose of water extract at the 24th hour was determined as 379.20 µg/ml, the IC50 effective dose of methanol extract was 1208.00 µg/ml, and the IC50 effective dose of chloroform extract was 168.60 µg/ml (Figure 1.).

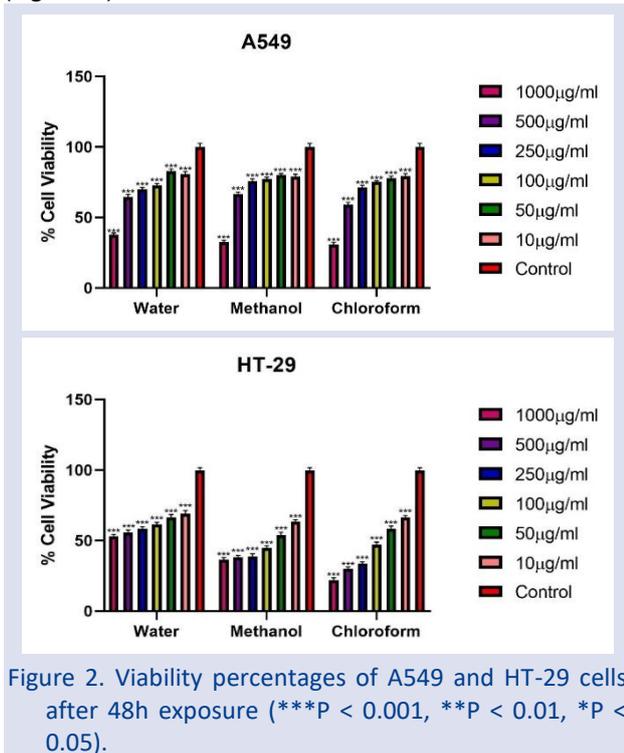


Figure 2. Viability percentages of A549 and HT-29 cells after 48h exposure (***P* < 0.001, ***P* < 0.01, **P* < 0.05).

In A549 cell lines, the IC50 effective dose of water extract was determined as 853.00 µg/ml, the IC50 effective dose of methanol extract was 813.60 µg/ml, and the IC50 effective dose of chloroform extract was 604.50 µg/ml for the 48th hour. The IC50 effective dose of the water extract was 853.00 µg/ml, the IC50 effective dose of the methanol extract was 813.60 µg/ml, and the IC50 effective dose of the chloroform extract was 604.50 µg/ml for the 48th hour in HT-29 cell lines, and the 48th hour in A549 cell lines. determined. In HT-29 cell lines, the IC50 effective dose of water extract was determined as 2239.00 µg/ml, the IC50 effective dose of methanol extract was 69.81 µg/ml, and the IC50 effective dose of chloroform extract was 69.46 µg/ml for the 48th hour (Figure 2.).

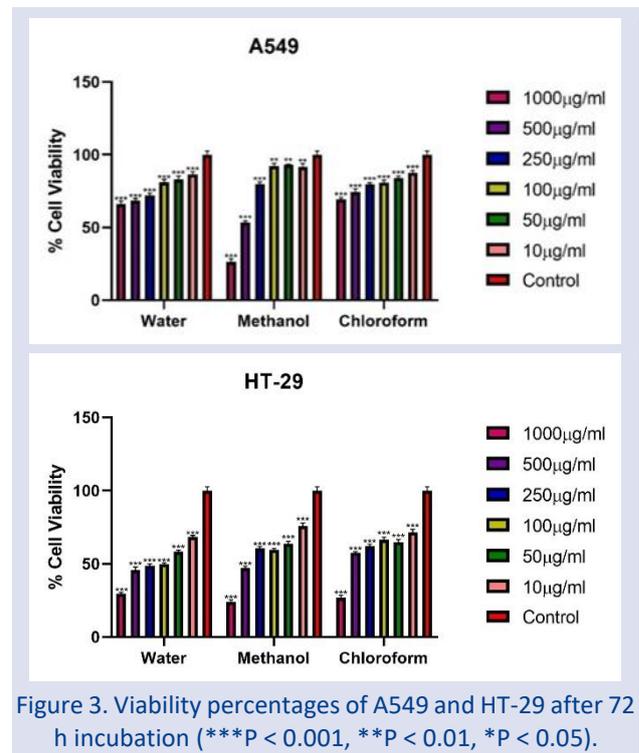
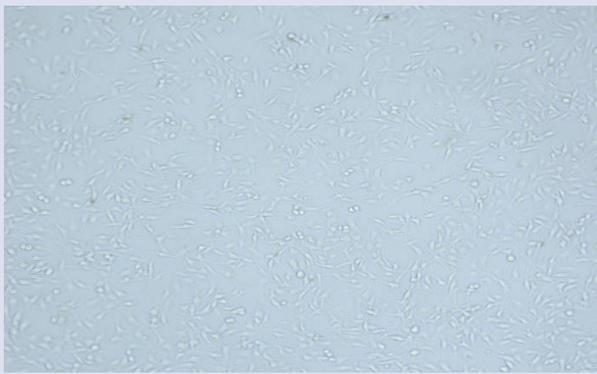


Figure 3. Viability percentages of A549 and HT-29 after 72 h incubation (***P* < 0.001, ***P* < 0.01, **P* < 0.05).

In A549 cell lines, the IC50 effective dose of the water extract was determined as 8169.00 µg/ml, the IC50 effective dose of the methanol extract was 544.10 µg/ml, and the IC50 effective dose of the chloroform extract was 29258.97 µg/ml for the 72nd hour. In HT-29 cell lines, the IC50 effective dose of water extract was determined as 139.00 µg/ml, the IC50 effective dose of methanol extract was 244.40 µg/ml, and the IC50 effective dose of chloroform extract was 399.00 µg/ml for the 72nd hour (Figure 3.).

Microscope images

Microscope images of A549 and HT-29 cell lines exposed to 50 µg/ml concentration of water, methanol and chloroform extract of Taraxacum officinal plant at 24h were presented (Figures 4-7.).

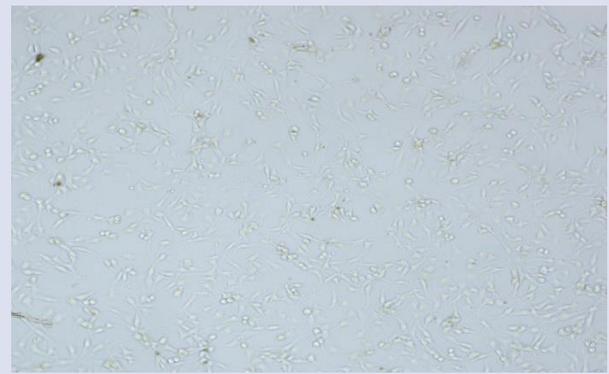


(a)

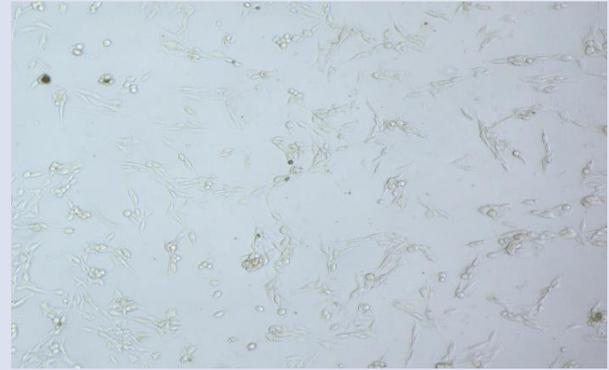


(b)

Figure 4. Microscope images of (a) A549 and (b) HT-29 cell lines exposed to water extract of *Taraxacum officinal* plant for 24h.

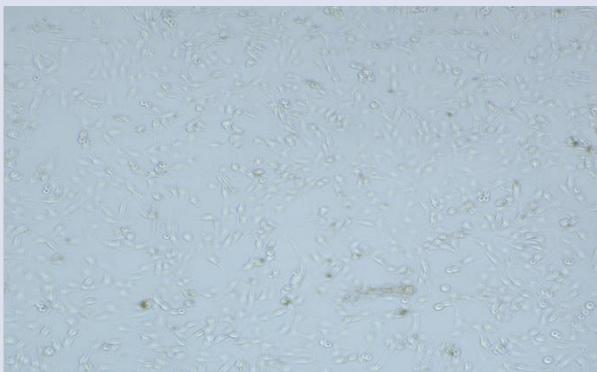


(a)

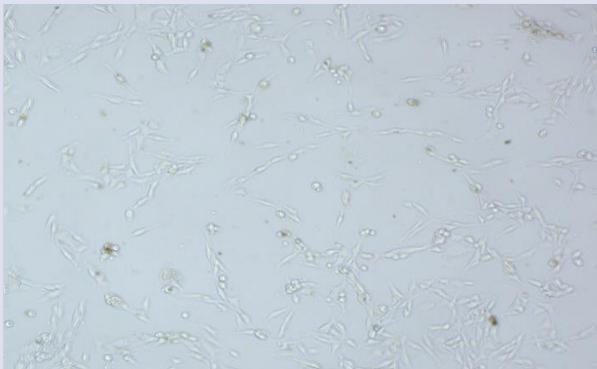


(b)

Figure 6. Microscope images of (a) A549 and (b) HT-29 cell lines exposed to chloroform extract of *Taraxacum officinal* plant for 24h.



(a)

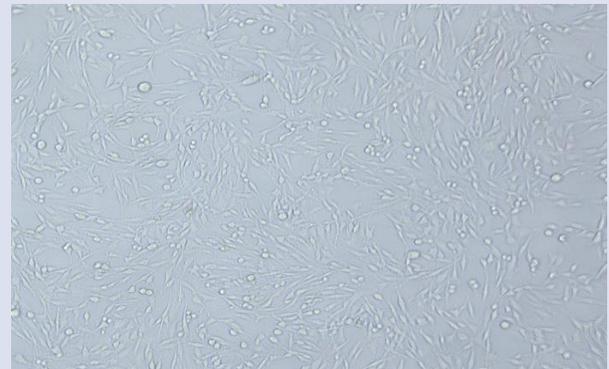


(b)

Figure 5. Microscope images of (a) A549 and (b) HT-29 cell lines exposed to methanol extract of *Taraxacum officinal* plant for 24h.



(a)



(b)

Figure 7. Microscope images of control group (a) A549 and (b) HT-29 cell lines at 24h.

BAX, BCL-2 and Caspase-3 Expression Levels

Caspase 3, BAX and BCL-2 expression levels were determined in cDNA samples obtained from A549 and HT-29 cell lines and control group cell lines, to which the effective dose of chloroform extract of Taraxacum officinal plant was applied for 24 hours (Figure 8.).

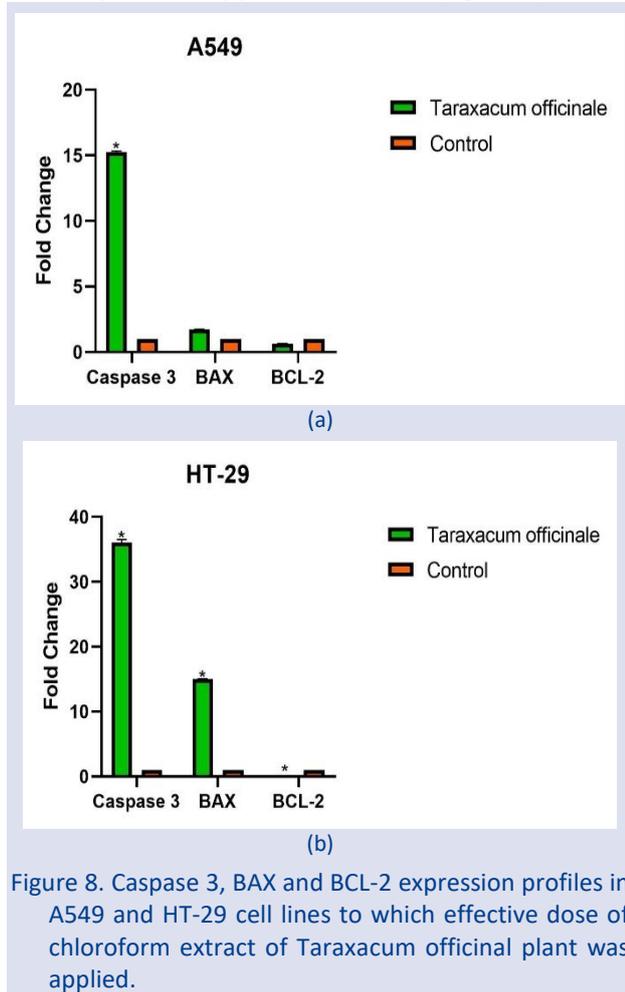


Figure 8. Caspase 3, BAX and BCL-2 expression profiles in A549 and HT-29 cell lines to which effective dose of chloroform extract of Taraxacum officinal plant was applied.

It was determined that Caspase 3 increased significantly by 15.24 times in A549 cell lines exposed to the effective dose of chloroform extract of the Taraxacum officinal plant at the 24th hour ($p < 0.05$). In HT-29 cell lines exposed to the effective dose of chloroform extract of the Taraxacum officinal plant at the 24th hour, BAX increased significantly by 15.03 times, Caspase 3 by 36.02 times, and BCL-2 increased significantly by 10.87 times. It was determined that it decreased ($p < 0.05$).

Total Antioxidant Status

TAS levels were determined in four different doses of water, methanol and chloroform extract of Taraxacum officinal plant concentrations between 10 μ g/ml and 1000 μ g/ml (Table 1.).

It was determined that the TAS levels of water, methanol and chloroform extracts of the Taraxacum officinal plant increased with increasing concentration, and the highest TAS level was at 1000 μ g/ml concentrations.

Table 1. TAS levels of water, methanol and chloroform extract of Taraxacum officinal

| | Water | Methanol | Chloroform |
|-----------------|-----------------------------|-----------------------------|-----------------------------|
| | TAS (mmol Trolox Equiv. /L) | TAS (mmol Trolox Equiv. /L) | TAS (mmol Trolox Equiv. /L) |
| 10 μ g/ml | 0,219 \pm 0,012 | 0,491 \pm 0,023 | 0,359 \pm 0,019 |
| 100 μ g/ml | 0,256 \pm 0,011 | 0,609 \pm 0,026 | 0,541 \pm 0,027 |
| 500 μ g/ml | 0,689 \pm 0,031 | 0,762 \pm 0,035 | 0,677 \pm 0,025 |
| 1000 μ g/ml | 1,270 \pm 0,037 | 1,111 \pm 0,045 | 0,889 \pm 0,041 |

The TAS level of the water extract of the Taraxacum officinal plant at a concentration of 1000 μ g/ml was 1,270 mmol Trolox Equiv. /L, TAS level at 1000 μ g/ml concentration of methanol extract is 1.111 mmol Trolox Equiv. /L, TAS level at 1000 μ g/ml concentration of chloroform extract is 0.889 mmol Trolox Equiv. /L.

Discussion

There are many studies investigating the effects of Taraxacum officinal on cancer. If we touch upon these studies, it has been shown that Taraxacum officinal extracts strongly inhibit breast cancer stem cell proliferation [21]. It has been determined that ethanol extracts of Taraxacum officinal have a proliferation reducing effect on MCF-7 breast cancer cell lines and WRL-68 liver cancer cell lines [22]. It has also been reported that ethanol extract of Taraxacum officinal flowers has an anti-cancer effect by inducing apoptosis in SK-OV-3 ovarian cancer cell lines [23]. However, in a study examining the anti-cancer activity of DMSO (dimethyl sulfoxide) and ethanol extracts of Taraxacum officinal leaves on glioblastoma, an aggressive type of cancer, alone and in combination with chicoric, chlorogenic, and caftaric acids, it was reported that DMSO extracts showed the best anti-tumor activity and that the anti-cancer activity of the plant was affected by the type of extract and the extract concentration [24]. In another study on glioblastoma regarding Taraxacum officinal, it was determined that DMSO, ethanol and water extracts of Taraxacum officinal leaves and roots all suppressed cell viability in U-138 cancer cell lines. It was reported that ethanol extracts inhibited proliferation at the highest level compared to other extracts [25]. In cancer cell culture studies using the nanomaterial developed by green synthesis method and consisting of Taraxacum officinal leaf extracts and ZnO, it was reported that this nanomaterial had cytotoxic effects on cancer cells. Therefore, it was emphasized that this nanomaterial could be used for cancer treatment with an environmentally friendly approach [26]. In a different study conducted on mouse fibrosarcoma and hepatoma cancer cell lines, it was determined that methanol extracts of Taraxacum officinal had dose-dependent anticancer effects on fibrosarcoma cell lines. In addition, caffeic acid, chlorogenic acid, and ursolic acid isolated from Taraxacum officinal methanol extract were shown to significantly suppress cell proliferation in these two cancer cell lines

[27]. In a study investigating the antioxidant and anti-cancer effects of the essential oil obtained from *Taraxacum officinale* in vivo and in vitro, it was shown that the essential oil obtained from the plant mediated an increase in antioxidant enzyme levels against oxidative stress induced by paracetamol in mice and protected the liver and kidneys of mice against oxidative damage. In addition, it was determined that *Taraxacum officinale* essential oil had proliferation inhibitory effects in HeLa cancer cells in relation to the dose used [28]. It was found that hydroalcoholic extracts of *Taraxacum officinale* had anti-tumor effects on triple-negative breast cancer cells 4T1 and increased apoptosis (by increasing expressions of p53, BAX, Beclin-1 and Atg-7 and BAX/BCL-2 ratio, and also by suppressing BCL-2 expression) and thus decreased proliferation in a dose-dependent manner [13]. It was shown that water and ethanol extracts of *Taraxacum officinale* prevented cell migration in hepatocyte carcinoma Huh7 cells, which frequently metastasize, and thus had metastasis-inhibiting properties in addition to its anti-cancer properties [29]. In another study, the anti-cancer activity of water, methanol, ethanol and chloroform extracts of the plant was investigated in prostate cancer cell lines PC-3, colorectal cancer cell lines Caco-2, breast cancer cell lines MCF-7 and chronic myeloid leukemia cell lines K562. It was determined that the chloroform extract, which was determined as the most active form of *Taraxacum officinale*, selectively showed an anti-proliferation effect in prostate and chronic myeloid leukemia cell lines [30]. Finally, it was shown that ethanol extracts of *Taraxacum officinale* plant roots applied to cervical cancer cell lines caused the induction of apoptosis in cancer cells in a manner closely related to the dose, thus inhibiting proliferation [31]. According to the data of our study, we can say that water, methanol and chloroform extracts of *Taraxacum officinale* plant have anti-cancer effects in A549 and HT-29 cancer cell lines. If we look at the results of studies in the literature emphasizing the anti-cancer properties of the *Taraxacum officinale* plant, they seem to be closely related to the results we obtained.

We determined that the chloroform extract of *Taraxacum officinale* activated cell apoptosis by inducing Caspase 3 expression in A549 cell lines. We found that it induced apoptosis by mediating an increase in Caspase-3 and BAX expression and a decrease in BCL-2 expression in HT-29 cell lines. In a study related to this subject, it was determined that cells were directed to apoptosis by increasing Bax and p53 expression and decreasing Bcl-2 expression in breast cancer 4T1 cells exposed to hydroalcoholic extract of *Taraxacum officinale* [13]. In a different study, it was found that *Taraxacum officinale* also had an apoptosis-inducing effect in leukemia cell lines [32]. Based on the results of our study, we can say that the inducing effects of *Taraxacum officinale* on apoptosis in A549 and HT-29 cell lines are possible through the regulation of different apoptosis-related genes.

In the study investigating the antioxidant capacity in water and ethanol-water extracts obtained from the leaves of *Taraxacum officinale* plant, it was determined

that the antioxidant capacity was quite high. In addition, it was determined that the extracts had a rich polyphenol content. In this respect, it was reported that *Taraxacum officinale* plant may be effective in the treatment of liver and gall diseases [33]. It was determined that the antioxidant capacity was high in all water-ethanol extracts obtained from the roots, stems, leaves and flowers of *Taraxacum officinale* by different methods. However, it was determined that the extraction method showing the best antioxidant activity was the Soxhlet method [34]. It was shown that the methanol extract was the extract type with the highest phenolic content among the methanol, acetone and n-hexane extracts of *Taraxacum officinale*. It was reported that the highest antioxidant activity was again found in the methanol extract [35]. In the study where methanol and ethyl acetate extracts of *Taraxacum officinale* leaves were evaluated in terms of phytochemicals, antioxidant capacity and vitamins, it was reported that the extracts contained high concentrations of phenolic content, flavonoids, alkaloids, vitamins A and C. When methanol and ethyl acetate extracts were compared in terms of antioxidant capacity, it was determined that the highest antioxidant activity was in the methanol extracts of the plant [36]. In a different study where the antioxidant capacity of the fruit extracts of the plant was investigated, it was revealed that the antioxidant content of the fruit extracts of *Taraxacum officinale* was also high [37]. Similar to these studies in the literature, we determined that the antioxidant status of water, methanol and chloroform extracts of *Taraxacum officinale* plant was high, that the antioxidant capacity increased dose-dependently with increasing extract concentrations and that the methanol extract of the plant had the highest antioxidant activity.

Conclusions

In conclusion, in this study where we examined the anti-cancer effects of *Taraxacum officinale* plant in lung and colon cancer, we determined that the extracts of this plant prevented proliferation in A549 lung and HT-29 colon cancer cells. We showed that *Taraxacum officinale* achieves this by regulating gene expressions in the apoptosis pathway in a way that induces apoptosis in cancer cells. Using these data obtained from our study, we can say that *Taraxacum officinale* plant may have therapeutic properties in lung and colon cancer. However, in order to express this prediction more strongly, more advanced and in-vivo studies related to *Taraxacum officinale* and cancer are needed.

Conflicts of interest

There are no conflicts of interest in this work.

Acknowledgement

There is no acknowledgement.

References

- [1] Petrovska B.B., Historical review of medicinal plants' usage, *Pharmacognosy Reviews*, 6 (11) (2012) 1.
- [2] Dutra R.C., Campos M.M., Santos A.R., Calixto J.B., Medicinal plants in Brazil: Pharmacological studies, drug discovery, challenges and perspectives, *Pharmacological Research*, 112 (2016) 4-29.
- [3] Sarkar B., Rana N., Singh C., Singh A., Medicinal herbal remedies in neurodegenerative diseases: an update on antioxidant potential, *Naunyn-Schmiedeberg's Archives of Pharmacology*, (2024) 1-29.
- [4] Tokhirova, A.A., Maksudjanovna M.G., Phytopharmacokinetics of Medicinal Plants and Drug Preparations Affecting Cardiovascular Diseases, *Web of Semantics: Journal of Interdisciplinary Science*, 2 (2) (2024) 10-12.
- [5] Gielecińska A., Kciuk M., Mujwar S., Celik I., Kotat D., Kałuzińska-Kotat Ż., Kontek R., Substances of natural origin in medicine: Plants vs. cancer, *Cells*, 12 (7) (2023) 986.
- [6] Di Napoli A., Zucchetti P., A comprehensive review of the benefits of Taraxacum officinale on human health, *Bulletin of the National Research Centre*, 45 (1) (2021) 110.
- [7] Clare B.A., Conroy R.S., Spelman K., The diuretic effect in human subjects of an extract of Taraxacum officinale folium over a single day, *The Journal of Alternative and Complementary Medicine*, 15 (8) (2009) 929-934.
- [8] Mir M.A., Sawhney S.S., Jassal M.M.S., In-vitro antidiabetic studies of various extracts of Taraxacum officinale, *The Pharma Innovation*, 4 (1) (2015) 61.
- [9] Choi U.K., Lee O.H., Yim J.H., Cho C.W., Rhee Y.K., Lim S.I., Kim Y.C., Hypolipidemic and antioxidant effects of dandelion (Taraxacum officinale) root and leaf on cholesterol-fed rabbits, *International journal of molecular sciences*, 11 (1) (2010) 67-78.
- [10] Colle D., Arantes L.P., Gubert P., da Luz S.C.A., Athayde M.L., Teixeira Rocha J.B., Soares F.A.A., Antioxidant properties of Taraxacum officinale leaf extract are involved in the protective effect against hepatotoxicity induced by acetaminophen in mice, *Journal of medicinal food*, 15 (6) (2012) 549-556.
- [11] Sigstedt S.C., Hooten C.J., Callewaert M.C., Jenkins A.R., Romero A.E., Pullin M.J., ... & Steelant W.F., Evaluation of aqueous extracts of Taraxacum officinale on growth and invasion of breast and prostate cancer cells, *International journal of oncology*, 32 (5) (2008) 1085-1090.
- [12] Menke K., Schwermer M., Felenda J., Beckmann C., Stintzing F., Schramm A., Zuzak T.J., Taraxacum officinale extract shows antitumor effects on pediatric cancer cells and enhance mistletoe therapy, *Complementary therapies in medicine*, 40 (2018) 158-164.
- [13] Ahmadi S., Saberivand A., Jalili C., Asadpour R., Khordadmehr M., Saberivand M., Hydroalcoholic extract of Taraxacum officinale induces apoptosis and autophagy in 4T1 breast cancer cells, *In Veterinary Research Forum*, 14 (9) (2023) 507.
- [14] Letai A., Apoptosis and cancer, *Annual Review of Cancer Biology*, 1 (1) (2017) 275-294.
- [15] Kiraz Y., Adan A., Kartal Yandim M., Baran Y., Major apoptotic mechanisms and genes involved in apoptosis, *Tumor Biology*, 37 (2016) 8471-8486.
- [16] Kapancık S., Çelik M.S., Demiralp M., Ünal K., Çetinkaya S., Tüzün B., Chemical composition, cytotoxicity, and molecular docking analyses of Thuja orientalis extracts. *Journal of Molecular Structure*, 1318 (2024) 139279.
- [17] Motafeghi F., Shahsavari R., Mortazavi P., Shokrzadeh M., Anticancer effect of paroxetine and amitriptyline on HT29 and A549 cell lines, *Toxicology in Vitro*, 87 (2023) 105532.
- [18] Liu G., Wang T., Wang T., Song J., Zhou Z., Effects of apoptosis-related proteins caspase-3, Bax and Bcl-2 on cerebral ischemia rats, *Biomedical reports*, 1 (6) (2013) 861-867.
- [19] Erel O., A novel automated method to measure total antioxidant response against potent free radical reactions, *Clinical biochemistry*, 37 (2) (2004) 112-119.
- [20] Kapancık S., Çelik V., Kılıçkap S., Kacan T., Kapancık S., The relationship of agmatine deficiency with the lung cancer, *Uhod-Uluslararası Hematoloji-Onkoloji Dergisi*, 26 (2016) 103-109.
- [21] Trinh N.V., Doan-Phuong Dang N., Hong Tran D., Van Pham P., Taraxacum officinale dandelion extracts efficiently inhibited the breast cancer stem cell proliferation, *Biomedical Research and Therapy*, 3 (2016) 1-13.
- [22] Rawa'a A.M., Cytotoxic activity of taraxacum officinale ethanolic plant extract against human breast cancer (MCF-7) cells and human hepatic (WRL-68) cells, *Iraqi Journal of Cancer and Medical Genetics*, 11 (1) (2018).
- [23] Choi E.J., Kim G.H., Dandelion (Taraxacum officinale) flower ethanol extract inhibits cell proliferation and induces apoptosis in human ovarian cancer SK-OV-3 cells, *Food science and biotechnology*, 18 (2) (2009) 552-555.
- [24] Fulga A., Casian A., Casian I., Protopop S., Gudumac V., Tagadiuc O., Effects of Taraxacum officinale on glioblastoma cell culture and their correlation with hydroxycinnamic acids content, *In Medical Sciences Forum*, 21 (1) (2023) 18.
- [25] Fulga A., Todiras M., Gudumac V., Tagadiuc O., Taraxacum officinale leaves and roots suppress glioma cell viability, *Journal of Biosciences and Medicines*, 10 (3) (2022) 175-189.
- [26] Kasthuri K., Rajkumar P., Ranchani A.A.J., Jayaprakash K., Bio-inspired synthesis of ZnO nanoparticles using Taraxacum officinale for antibacterial, antifungal, and anticancer applications, *Inorganic Chemistry Communications*, 170 (3) (2024) 113409.
- [27] Korbášová M., Tomenendálová J., Chloupek J., Antitumor effect of combinations of three acids isolated from Taraxacum officinale, *Acta Veterinaria Brno*, 91 (1) (2022) 77-85.
- [28] Kamal F.Z., Lefter R., Mihai C.T., Farah H., Ciobica A., Ali A., ... & Ech-Chahad A., Chemical composition, antioxidant and antiproliferative activities of Taraxacum officinale essential oil, *Molecules*, 27 (19) (2022) 6477.
- [29] Yoon H.S., An H., Park C.M., Anti-metastatic Effect of Taraxacum Officinale Water and Ethanol Extracts Through the Regulation of Epithelial-Mesenchymal Transition in Huh7 Cells, *Journal of The Korean Society of Integrative Medicine*, 11 (3) (2023) 59-67.
- [30] Al-Eisawi Z., Abderrahman S.M., Al-Khalaf I.F., Al-Abbassi R., Bustanji Y.K., Taraxacum officinale extracts exhibit safe and selective anticancer activity, *The Natural Products Journal*, 12 (1) (2022) 69-77.
- [31] Venezuela R.F., Mosmann J.P., Mugas M.L., Kiguen A.X., Montoya S.C.N., Konigheim B.S., Cuffini C.G., Dandelion Root Extract Affects the Proliferation, Survival and Migration of Cervical Cancer Cell Lines, *Research Square*, (2021) 1-15.

- [32] Appiah-Opong R., Tuffour I., Annor G.K., Doris A., Blankson-Darku P.C., Kissi-Twum A.A., ... & Ocloo A., Antiproliferative, antioxidant activities and apoptosis induction by *Morinda lucida* and *Taraxacum officinale* in human HL-60 leukemia cells, *Journal of Global Biosciences*, 5 (7) (2016) 4281-4291.
- [33] Erdem M., Özaslan M., Investigation of the antioxidant capacity of *Taraxacum officinale* l. leaf extract, *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 28, (2024) 575-580.
- [34] Dedić S., Džaferović A., Jukić H., Chemical Composition and Antioxidant Activity of Water-Ethanol Extracts of Dandelion (*Taraxacum officinale*), *Hrana u zdravlju i bolesti: znanstveno-stručni časopis za nutricionizam i dijetetiku*, 11 (1) (2022) 8-14.
- [35] Akhtar W., Ali G., Ashraf N., Fatima I., Kayani W.K., Shaheen H., ... & Khames A., Efficiency of Multiple Extraction Solvents on Antioxidant, Cytotoxic, and Phytotoxic Potential of *Taraxacum officinale* (L.) Weber ex FH Wigg. from Poonch Valley, Azad Kashmir, Pakistan, *Evidence-Based Complementary and Alternative Medicine*, 1 (2022) 5118553.
- [36] Daniel I.E., Mathew K.N., John P.L., Evaluation of vitamin contents, antioxidant and antimicrobial activities of different leaf extracts of *Taraxacum officinale* (dandelion), *Journal of Complementary and Alternative Medical Research*, 13 (1) (2021) 13-26.
- [37] Lis B., Jedrejek D., Rywaniak J., Soluch A., Stochmal A., Olas B., Flavonoid preparations from *Taraxacum officinale* L. fruits—A phytochemical, antioxidant and hemostasis studies, *Molecules*, 25 (22) (2020) 5402.



Classification of Grapevine Leaf Types with Vision Transformer Architecture

Esra Kavalcı Yılmaz^{1,a,*}, Hatice Aktaş^{1,b}, Kemal Adem^{2,c}

¹ Department of Computer Engineering, 'Sivas University of Science and Technology, Sivas, Türkiye.

² Department of Computer Engineering, Sivas Cumhuriyet University, Sivas, Türkiye.

*Corresponding author

Research Article

History

Received: 11/09/2024

Accepted: 13/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Viticulture plays an important role in agriculture. Farmers prefer grapevine cultivation because not only its fruit but also its leaves are used in various fields. Both the use and trade of grapevine leaves within the country is an important source of income. Grapevine leaves, which are grown in almost all countries and used as edible, vary in terms of species. Determining and cultivating the species according to their suitability in terms of productivity is important. In this study, artificial intelligence methods were used to classify grapevine leaf species. The dataset consisting of five different classes, including 100 grapevine leaf images for each class, totalling 500 images, was classified using ViT, VGG19 and MobileNet methods. When the methods used in this study to help increase productivity in production are evaluated, ViT method has the best accuracy rate with 94%.

Keywords: Grapevine leaf classification, Deep learning, ViT, Artificial intelligence.

^a esra.kavalcı@sivas.edu.tr

^b <https://orcid.org/0000-0003-1314-4495>

^c haticeaktas@sivas.edu.tr

^d <https://orcid.org/0000-0002-1104-9307>

^e kemaladem@cumhuriyet.edu.tr

^f <https://orcid.org/0000-0002-3752-7354>

Introduction

Agriculture is one of the most important sources of livelihood in our country. Viticulture has a great place in agriculture, which has many varieties and branches. It is often preferred by farmers both for the livelihood it provides and for its diversity of use. Grapevine is one of the wide range of products provided by viticulture. In our country, the fruit obtained from the grapevine is used in different areas. In addition, grapevine leaves are in demand as much as the fruit [1]. Grapevine leaves, which are frequently consumed in every region of our country, is a product with economic return. The trade of grapevine leaves, which are made durable by using different storage methods, has also increased in recent years. It is important to cultivate, collect, and store different types of grapevine leaves suitable for edible use in different regions. The grapevine leaves, which have different characteristics according to their thickness, shape and hairiness, are carefully selected during cultivation and collection.

The high diversity of grapevine leaves grown in almost every region and the manual determination of the species can negatively affect the classification process and efficiency of the appropriate grapevine leaves to be selected for consumption. The use of systems established with artificial intelligence in the field of agriculture, as in every field, is quite common today. [2]. In the literature, there are many studies involving artificial intelligence for the detection of diseases, identification and classification of species in the field of agriculture [3, 4]. Artificial intelligence has been used in some previous studies on grapevine and grapevine leaves. In these studies, diseases such as mould, grapevine yellowness, leaf blight, black measles, and esca seen in grapevine and grapevine leaves

were tried to be detected by artificial intelligence [5-10]. In the presented studies, in addition to methods such as XGBoost, SVM, DNN, CNN, AlexNet, GoogleNet, and ResNet-50, artificial intelligence methods were also used in the image processing phase. In another study [11], nitrogen concentration estimation of grapevine leaves was also performed using artificial intelligence. Some artificial intelligence methods used to classify the diversity of grapevine leaves were also presented in [12, 13]. In the study [12], where data augmentation was first performed using a dataset containing images of grapevine leaves, Support Vector Machines were used for classification and an accuracy rate of 96.14% was obtained. In the classification study [13], in which both CNN model and Support Vector Machines were used together, grapevine leaf image data set containing 5 different classes was used and 97.6% success rate was obtained in classification.

In addition to agriculture, studies in the literature show that the ViT model yields successful results in different fields. The best example of this can be given as studies in the field of health. Zang and Wen presented a transformer-based framework for automatic diagnosis of COVID-19 from chest CT scans. In this framework, lung segmentation was performed with UNet and 3D volume level features were extracted using the Swin transformer. The study was performed on the MIA-COV19D dataset and they obtained an F1 score of 0.935 and an accuracy of 94.3% with the Swin transformer [14]. Tyagi et al. used CNN, VGG16 and Vision Transformer (ViT) models to automatically detect pneumonia from chest X-rays. The ViT model showed the best performance in pneumonia detection with an accuracy of 96.45% and it was stated that it can be used to recognize other lung-related

diseases [15]. Dai et al. proposed the TransMed model, which combines the advantages of CNN and transducers for multimodal medical image classification. The model aims to establish long-range dependencies between modalities while extracting low-level features of images. It was tested on two datasets, classifying parotid gland tumors and knee injuries. For parotid gland tumors, 88.9% success was achieved with the proposed TransMed model [16]. In another study, Kamran et al. proposed an approach that converts fundus images into Fluorescence Angiography (FA) images to diagnose retinal abnormalities. Fundus images are synthesized into FA images using a Generative Adversarial Network (GAN) and these images are given to a transform model. The proposed model achieved 85.7% accuracy, 83.3% sensitivity and 90% specificity in classifying normal and abnormal retina, respectively [17]. Zeid et al. investigated the effectiveness of Vision Transformers in the classification of Colorectal Cancer (CRC) histological images. In their study, they used a publicly available CRC histology dataset and performed experiments with a total of 5000 images with eight different tissue categories. By training Vision Transformer and Compact Convolutional Transformer models, they achieved 93.3% and 95% accuracy rates, respectively [18]. In another study, Xu et al. proposed a model that combines transfer learning with attention mechanism for glaucoma detection. This method strengthened the distinction between general and specific features by identifying the regions containing information in images, and also allowed visualization of these regions thanks to the attention mechanism. The model was evaluated on two different datasets and achieved impressive results with 85.77% accuracy, 84.9% sensitivity, 86.9% specificity and 0.929 AUC values [19].



Figure 1. Sample images of classes

Deep Learning Methods

Deep learning methods are a subset of machine learning involving neural networks with multiple layers to model complex patterns in data. These methods are especially effectively used in tasks such as image and speech recognition, natural language processing, and autonomous driving. Deep learning algorithms can learn from large amounts of data by simulating the structure of the human brain, making them powerful tools for predictive analytics and decision making [21,22]. It is widely used for disease diagnosis in healthcare, fraud detection in finance, plant species and disease diagnosis in agriculture, and content recommendations in

Vision Transformer has been used in different fields with its minimization of data loss, self-attention mechanism and highlighting of important features and it has been seen that good results have been obtained. In this study, considering the advantages of the ViT method, grapevine leaf classification was performed according to their types. The ViT model is trained with image dataset containing grapevine varieties for multiclass classification. Grapevine leaf images were processed with artificial intelligence and classified according to 5 different previously determined species classes. With this study, it is aimed to determine the species of grapevine leaves, which are of great interest both in the domestic and foreign markets, more effectively and to contribute to increasing the cultivation efficiency.

Materials and Methods

Data Set

The 'Grapevine Leaf Image Dataset' used in the study was created by Köklü et al. [13] and shared on the Kaggle website. The dataset consists of images of 5 different types of grapevine leaves, namely Ak, Ala İdris, Büzgülü, Dimnit, and Nazlı. In the dataset, there are 100 images of each class in 512x512 size, totaling 500 images in total [20]. Sample images of each class are presented in Figure 1. In this study, the data set was tripled by using data augmentation methods such as random rotation, horizontal and vertical translation, color changes and random cropping. The dataset, which initially consisted of 400 training and 100 test images, reached 1200 training and 300 test images as a result of these processes.

entertainment [23-29]. The deep learning methods used in this study are briefly mentioned below.

- Vision Transformer (ViT): Vision Transformer (ViT), a deep learning model commonly used in computer vision, was introduced by Dosovitskiy et al. in 2020 [30]. Unlike traditional convolutional neural networks (CNNs), ViT uses a transformer architecture, which has become a popular alternative for image processing tasks. ViT utilizes a set of attention mechanisms instead of convolutional layers during the training of data. It helps to reduce image size by splitting images into smaller patches, allowing higher-dimensional images to be processed compared to previous models. By converting images into patches and applying an attention mechanism, ViT can more efficiently

capture complex patterns in the data. Thanks to this approach, higher success can be achieved with less data in studies using ViT. This model has proven its effectiveness in tasks such as image classification and object detection, showing significant improvements over CNNs in various benchmarks. As a result, ViT has become a powerful tool for advancing the state of the art in computer vision [31]. The main reason for choosing the ViT (Vision Transformer) model in this study is to avoid the loss of information caused by the pooling layers used in traditional CNN models. While CNN models are successful in capturing local features, they are limited in modeling long-range relationships and can lead to loss of details during pooling operations. The ViT model divides the image into fixed-size patches, enriches the features of each patch with position information, and can globally model the relationships between all patches in the image with a self-attention mechanism. In this way, it provides superior performance, especially in applications that require detail precision and work on limited data sets [32].

VGG19: VGG19, proposed by Simonyan and Zisserman [33] in 2015, is a convolutional neural network model designed for image classification tasks. VGG19 consists of 19 layers, 16 convolutional layers and 3 fully connected layers, making it a deep and powerful architecture. VGG19 provides a robust framework for complex image recognition tasks. The architecture of the network is characterised by its simplicity; it consistently uses small 3x3 convolutional filters throughout the layers, which allows it to effectively capture complex features in images. Despite its depth, VGG19 is relatively easy to implement, making it a popular choice for researchers and developers alike. One of the greatest strengths of VGG19 is its ability to generalise well across a wide range of visual tasks, making it widely used in computer vision studies [34].

MobileNet: MobileNet, developed by Google, is a convolutional neural network model specifically designed for efficient performance on mobile and embedded devices. One of its key features is the fast convergence of the model during training. In addition, MobileNet is ideal for environments with limited resources due to its low memory consumption [35]. It also reduces the overall computational load by requiring fewer computational operations compared to traditional models. This efficiency is achieved through the use of deeply separable convolution layers that significantly reduce the number of required parameters and computations. Its lightweight architecture ensures high accuracy while minimising latency, even when running on hardware with limited capabilities. This makes MobileNet a popular choice for applications such as image recognition, object detection, and other machine learning tasks that need to be

performed on the go. Its flexibility also allows developers to easily scale the model to meet the specific needs of their applications, ensuring a balance between accuracy and efficiency [36]

Results and Discussion

In this study, three different deep learning models, namely ViT, VGG19, and MobileNet, were used to successfully classify grapevine leaves according to their species. The studies were carried out in Python language using the Tensorflow library. In the study, image sizes were set to 224x224x3, 'batch' size = 16 and 'epoch' = 100. The precision, recall, and f1 score values obtained as a result of the classification processes of ViT, VGG19, and MobileNet models for each class are presented in Table 1, Table 2, and Table 3, respectively.

Table 1. ViT results

| Class | Precision | Recall | f1-Score |
|-----------|-----------|--------|----------|
| Ak | 0.98 | 1.00 | 0.99 |
| Ala Idris | 0.94 | 0.98 | 0.96 |
| Büzgülü | 0.97 | 0.93 | 0.95 |
| Dimnit | 0.95 | 0.93 | 0.94 |
| Nazli | 0.97 | 0.95 | 0.96 |

Table 2. VGG19 results

| Class | Precision | Recall | f1- Score |
|-----------|-----------|--------|-----------|
| Ak | 0.84 | 0.91 | 0.88 |
| Ala Idris | 0.78 | 0.79 | 0.81 |
| Büzgülü | 0.80 | 0.67 | 0.70 |
| Dimnit | 0.79 | 0.78 | 0.75 |
| Nazli | 0.83 | 0.85 | 0.86 |

Table 3. MobileNet results

| Class | Precision | Recall | f1- Score |
|-----------|-----------|--------|-----------|
| Ak | 0.87 | 0.98 | 0.92 |
| Ala Idris | 0.86 | 0.80 | 0.83 |
| Büzgülü | 0.81 | 0.72 | 0.76 |
| Dimnit | 0.84 | 0.80 | 0.82 |
| Nazli | 0.83 | 0.92 | 0.87 |

When the results of the ViT Model (Table 1) are analyzed, it is seen that the most successful classification process among the classes is performed in Dimnit and Ala Idris classes. The most unsuccessful results were obtained in the images belonging to the Büzgülü class.

Table 4 shows the best results obtained by using the hyper parameters "activation_function='gelu', learning_rate =1e-4, optimization_algorithm=RectifiedAdam, loss_function =CategoricalCrossentropy". Accuracy values, average precision- recall-F1-score values, kappa values and loss values are listed for the models used

Table 4. ViT, VGG19, MobileNet doğruluk değerleri

| Model | Accuracy | Prec Avg | Recall Avg | F1-Score Avg | Kappa | Loss |
|-----------|----------|----------|------------|--------------|-------|------|
| ViT | 0.96 | 0.96 | 0.95 | 0.96 | 0.95 | 0.87 |
| VGG19 | 0.81 | 0.81 | 0.80 | 0.80 | 0.74 | 0.98 |
| MobileNet | 0.84 | 0.84 | 0.84 | 0.84 | 0.80 | 0.94 |

When Table 4 is analyzed, it is seen that the most successful classification process is provided by the ViT Model (0.96). Also, when the loss values are analyzed, it can be said that the ViT model provides a more stable and efficient learning compared to the others with a lower loss value (0.87). The ViT model provides superiority in capturing the global context thanks to its attention mechanism, enabling effective learning of long-range dependencies in images. In addition, it allows the image to be processed by dividing it into patches. This patch-based input method makes the application easier by reducing the need for extensive preprocessing. Thanks to these advantages, it can be said that the ViT model offers better generalization and more successful results after the training process compared to other models. The box plot of ViT, VGG19, and MobileNet models organized according to the classification results is presented in Figure 2.

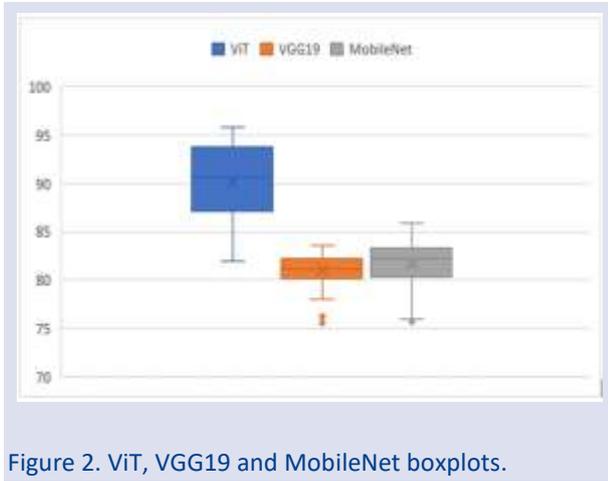


Figure 2. ViT, VGG19 and MobileNet boxplots.

Figure 2 shows that ViT offers the widest range of performance with an accuracy value ranging between 86 and 94 and a median value of around 91. According to these results, ViT indicates high but variable performance. The VGG19 model, on the other hand, has a closer range, the accuracy is mostly between 80 and 83 and the median value is about 82. Therefore, it can be said that it is more consistent than the ViT model but indicates a slightly lower performance. Finally, the MobileNet model shows the closest range with an accuracy between 80 and 84 and a median of around 83. According to these results, it can be said that it shows the lowest and most consistent performance among the three models. ViT has potentially the best performance but shows higher variability, while VGG19 offers a stable performance. MobileNet lags behind in both range and average accuracy. Overall, ViT stands out for its high but fluctuating accuracy, VGG19 for its consistency and MobileNet for its low but stable performance.

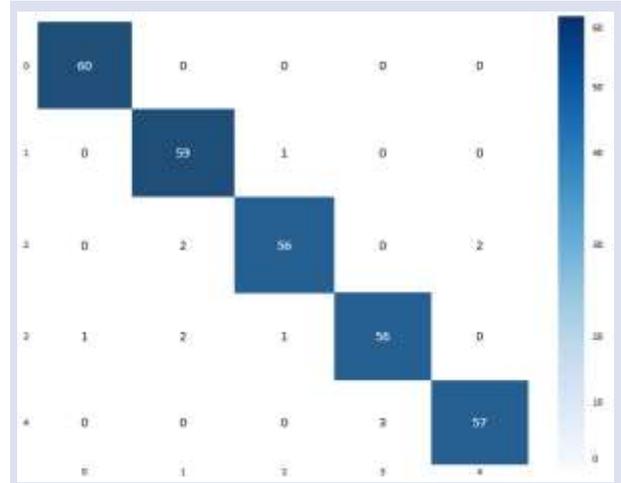


Figure 3. Confusion matrix of the ViT model.

Conclusion and Recommendations

Correct classification of agricultural products increases productivity in production. When the classification process is carried out by means of artificial intelligence methods, the process will be independent of human beings, the expert performing the classification process will not be affected by external factors such as physical, mental conditions and experience level, and the classification process will be concluded more successfully. In addition, the experts performing the classification process will use their time more efficiently by completing more complex tasks instead of routine tasks during this period. In line with these reasons, in this study, classification processes were performed using deep learning methods in order to increase the classification success of the grapevine leaf. Three different deep learning methods, namely ViT, VGG19, and MobileNet, were used in the study and the results were evaluated comparatively. Considering the results obtained, it is seen that the most successful result is obtained with 94% accuracy rate with ViT, a current deep learning model. In future studies, it is thought that more accurate results can be obtained by using various data augmentation methods, increasing and diversifying the data set. In addition, the classification success can be improved by hyperparameter optimization.

Conflict of interest

There are no conflicts of interest in this work.

References

- [1] Gülcü M., Torçuk A. İ., *Yemelik Asma Yaprağı Üretimi ve Pazarlamasında Kalite Parametreleri, Meyve Bilimi*, c. 1, ss. 75-79, (2016)
- [2] Adem K., Yılmaz E. K., Ölmez F., Çelik K., Bakır H., *A Comparative Analysis of Deep Learning Parameters for Enhanced Detection of Yellow Rust in Wheat, UMAG*, 16(2) (2024)

- [3] Yilmaz E. K., Oğuz T., Adem K., *A CNN-Based Hybrid Approach to Classification of Raisin Grains*, 1st International Conference on Frontiers in Academic Research, (2023).
- [4] Yilmaz E. K., Adem K., Kılıçarslan S., Aydın H. A., *Classification of lemon quality using hybrid model based on Stacked AutoEncoder and convolutional neural network*, *Eur Food Res Technol*, 249(6) (2023) 1655-1667.
- [5] Hernández I., Gutiérrez S., Ceballos S., Iñíguez R., Barrio I., Tardaguila J., *Artificial Intelligence and Novel Sensing Technologies for Assessing Downy Mildew in Grapevine*, *Horticulturae*, 7(5) (2021)
- [6] Cruz A. vd., *Detection of grapevine yellows symptoms in Vitis vinifera L. with artificial intelligence*, *Computers and Electronics in Agriculture*, 157 (2019) 63-76.
- [7] Poblete-Echeverría C., Hernández I., Gutiérrez S., Iñíguez R., Barrio I., Tardaguila J., *Using artificial intelligence (AI) for grapevine disease detection based on images*, *BIO Web Conf.*, 68 (2023) 01021.
- [8] Nagi R. Tripathy S. S., *Deep convolutional neural network based disease identification in grapevine leaf images*, *Multimed Tools Appl*, 81 (18) (2022) 24995-25006.
- [9] Alessandrini M., Calero Fuentes Rivera R., Falaschetti L., Pau D., Tomaselli V., ve Turchetti C., *A grapevine leaves dataset for early detection and classification of esca disease in vineyards through machine learning*, *Data in Brief*, 35 (2021) 106809.
- [10] Jaisakthi S. M., Mirunalini P., Thenmozhi D., Vatsala, *Grape Leaf Disease Identification using Machine Learning Techniques*, 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), (2019) 1-6.
- [11] Moghimi A., Pourreza A., Zuniga-Ramirez G., Williams L. E., Fidelibus M. W., *A Novel Machine Learning Approach to Estimate Grapevine Leaf Nitrogen Concentration Using Aerial Multispectral Imagery*, *Remote Sensing*, 12(21) (2020) 3515.
- [12] İmak A., Doğan G., Şengür A., ve Ergen B., *Asma Yaprağı Türünün Sınıflandırılması için Doğal ve Sentetik Verilerden Derin Öznitelikler Çıkarma, Birleştirme ve Seçmeye Dayalı Yeni Bir Yöntem*, *International Journal of Pure and Applied Sciences*, 9(1) (2023) 46-55.
- [13] Koklu M., Unlarsen M. F., Ozkan I. A., Aslan M. F., Sabanci K., *A CNN-SVM study based on selected deep features for grapevine leaves classification*, *Measurement*, 188 (2022) 110425.
- [14] Zhang, L., Wen, Y. A transformer-based framework for automatic COVID19 diagnosis in chest CTs. In Proceedings of the IEEE/CVF international conference on computer vision, (2021) 513-518.
- [15] Tyagi, K., Pathak, G., Nijhawan, R., & Mittal, A. Detecting pneumonia using vision transformer and comparing with other techniques. In 2021 5th international conference on electronics, communication and aerospace technology (ICECA), (2021) 12-16.
- [16] Dai, Y., Gao, Y., Liu, F., Transmed: Transformers advance multi-modal medical image classification, *Diagnostics*, 11(8) (2021) 1384.
- [17] Kamran, S. A., Hossain, K. F., Tavakkoli, A., Zuckerbrod, S. L., Baker, S. A., Vtgan: Semi-supervised retinal image synthesis and disease prediction using vision transformers. In Proceedings of the IEEE/CVF international conference on computer vision, (2021)3235-3245.
- [18] Zeid, M. A. E., El-Bahnasy, K., Abo-Youssef, S. E., Multiclass colorectal cancer histology images classification using vision transformers. In 2021 tenth international conference on intelligent computing and information systems (ICICIS) (2021) 224-230.
- [19] Xu, X., Guan, Y., Li, J., Ma, Z., Zhang, L., Li, L., Automatic glaucoma detection based on transfer induced attention network, *BioMedical Engineering OnLine*, 20(1) (2021) 39.
- [20] *Grapevine Leaves Image Dataset*, <https://www.kaggle.com/datasets/muratkokludataset/grapevine-leaves-image-dataset>
- [21] Alaca Y., Emin B., Akgul A., *A comparative study of deep learning models and classification algorithms for chemical compound identification and Tox21 prediction*, *Computers & Chemical Engineering*, 189 (2024) 108805,
- [22] Assim O. M., Mahmood A. F., *A novel Universal Deep Learning Approach for Accurate Detection of Epilepsy*, *Medical Engineering & Physics*, (2024) 104219.
- [23] Közkurt C., Diker A., Elen A., Kılıçarslan S., Dönmez E., ve Demir F. B., *Trish: an efficient activation function for CNN models and analysis of its effectiveness with optimizers in diagnosing glaucoma*, *J Supercomput*, 80 (11) (2024) 15485-15516.
- [24] Kılıçarslan S., Diker A., Közkurt C., Dönmez E., Demir F. B., ve Elen A., *Identification of multiclass tympanic membranes by using deep feature transfer learning and hyperparameter optimization*, *Measurement*, c. 229, s. 114488, (2024)
- [25] Daza A., González Rueda N. D., Aguilar Sánchez M. S., Robles Espiritu W. F., Chauca Quiñones M. E., *Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works*, *International Journal of Information Management Data Insights*, 4 (2) (2024) 100267.
- [26] Altaş Z., Ozguven M., Adem K., *Bazı Bağ Hastalıklarının Faster R-CNN Modeli ile Otomatik Tespit Edilmesi ve Sınıflandırılması*, *Turkish Journal of Agriculture - Food Science and Technology*, 11(97) 97-103.
- [27] Adem K., Ozguven M. M., ve Altaş Z., *A sugar beet leaf disease classification method based on image processing and deep learning*, *Multimed Tools Appl*, 82(8) (2023) 12577-12594,
- [28] Alnasyan B., Basherri M., Alassafi M., *The power of Deep Learning techniques for predicting student performance in Virtual Learning Environments: A systematic literature review*, *Computers and Education: Artificial Intelligence*, 6 (2024) 100231.
- [29] Dönmez E., Kılıçarslan S., ve Diker A., *Classification of hazelnut varieties based on bigtransfer deep learning model*, *Eur Food Res Technol*, c. 250, sy 5, ss. 1433-1442, (2024)
- [30] Dosovitskiy A. vd., *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*, International Conference on Learning Representations, (2020)
- [31] Dümen S., Yilmaz E. K., Adem K., ve Avaroglu E., *Performance of vision transformer and swin transformer models for lemon quality classification in fruit juice factories*, *European Food Research and Technology*, ss. 1-12, (2024)
- [32] Heo, J., Seo, S., Kang, P., Exploring the differences in adversarial robustness between ViT-and CNN-based models using novel metrics, *Computer Vision and Image Understanding*, 235 (2023) 103800.
- [33] Simonyan K. Zisserman A., *Very Deep Convolutional Networks for Large-Scale Image Recognition*, (2015) arXiv: arXiv:1409.1556.

- [34] Bansal M., Kumar M., Sachdeva M., Mittal A., *Transfer learning for image classification using VGG19: Caltech-101 image data set*, *J Ambient Intell Human Comput*, 14(4) (2023) 3609-3620.
- [35] Shome N., Kashyap R., Laskar R. H., *Detection of tuberculosis using customized MobileNet and transfer learning from chest X-ray image*, *Image and Vision Computing*, 147 (2024) 105063.
- [36] Kılıçarslan S., Aydın H. A., Adem K., Yılmaz E. K., *Impact of optimizers functions on detection of Melanoma using transfer learning architectures*, *Multimed Tools Appl* (2024).

Characterization of Local Okra Seed (*Abelmoschus Esculentus*) Extract and Evaluation of Bioactivity

Bilgenur Celebi ^{1,a}, Sevil Ozer ^{1,b,*}, Sezen Canim Ates ^{1,c}

¹ Biomedical Engineering Department, Istanbul Yeniuzyil University, İstanbul, Türkiye

*Corresponding author

Research Article

History

Received: 20/07/2024

Accepted: 16/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

A. esculentus (okra) plant is an economically important medicinal plant grown in tropical and subtropical parts of the world. The fruit and seeds of the *A. esculentus* (okra) plant; which are rich in oil and protein. The *in vitro* cytotoxic, bactericidal, and antioxidant activities of extracts obtained from different parts of okra (leaves, fruits, and seeds) were studied widely in the medical industry. This study evaluated the biocompatible and antioxidant potential of an optimized ethanolic extract of *A. esculentus* seeds. Commercially obtained *A. esculentus* seeds were first extracted and characterized by UV-Vis and FT-IR spectrophotometry. *In vitro*, antioxidant activity of okra seed extracts was evaluated using 1,1 -diphenyl -2picrylhydrazyl (DPPH) with different concentrations (0.5-5 mg/mL). Results show that the antioxidant activity increased in direct proportion to the increase in the concentration of the extract. *In vitro*, the cytotoxic effect of the extracts on C6 Glioma cancer cell line and L929 mouse fibroblast cell line was studied using 3-(4, 5-dimethylthiazolyl -2,5-diphenyl-tetrazolium bromide (MTT) assay with the concentrations between 50-1000 µg/mL. MTT results showed effective cytotoxicity on cancer cell lines with increasing extract concentration (IC50 value for C6 was 273.4 µg/ml, while the IC50 value for L929 was 431.45 µg/ml). The study indicates that the extract isolated from *A. esculentus* seeds shows that all concentrations have a substantial amount of anticancer and antioxidant activity. In conclusion, *A. esculentus* seeds could be considered innovative products and be proposed for alternative end-uses in pharmaceutical industries with antimicrobial and bioactive properties.

Keywords: *Abelmoschus esculentus*, Cancer, C6 Glioma, L929, Cytotoxicity, Antioxidants.

^a 23242401003@ogr.yeniuzyil.edu.tr ^b <https://orcid.org/0009-0006-5737-1840>

^c sezen.canimates@yeniuzyil.edu.tr ^d <https://orcid.org/0000-0003-2196-7053>

^e sevil.ozero@yeniuzyil.edu.tr

^f <https://orcid.org/0000-0002-0186-763X>

Introduction

Different civilizations have widely used functional plants for thousands of years for traditional treatment purposes. They are exploited for their bioactive compounds like polyphenols, carotene, folic acid, thiamine, vitamin C, and amino acids, which are highly demanded in cosmetic, pharmaceutical, and nutraceutical industries [1,2]. It is reported that plants have antimicrobial, anticarcinogenic, anti-glycemic, antioxidant, anti-cholesterol, antiparasitic, and anti-inflammatory properties, thanks to the phenolic compounds found at different levels in almost all parts of plants [3,4].

A. esculentus is a member of the mallow (Malvaceae) family and is often used as an edible immature fruit. This herbaceous annual plant is native to Africa (Ethiopia) and is now almost cosmopolitan, cultivated in tropical, subtropical, and warm temperate climates in different countries from Africa to Asia, southern Europe, and America [5,6]. In traditional medicine, okra (*A. esculentus*) is thought to exhibit diuretic, antispasmodic, relaxing, laxative, and anti-cancer effects. It is also known for its beneficial mucus content. The public uses it to treat diseases such as colds, gonorrhoea, dental disease, and bronchitis. *A. esculentus* attracts the attention of the public and researchers. In addition to the fruit of the plant, its leaves, flowers, stems, and seeds are rich in oil and

protein. *In vivo* and *in vitro* antioxidant, antidiabetic, antihyperlipidemic, antimicrobial, antifatigue, hepatoprotective, and neuroprotective effects of okra (*A. esculentus*) have been reported [7-10]. A substance frequently mentioned in the okra seed extract is isoquercitrin, with higher bioavailability than quercetin, displaying a few chemoprotective effects against cancer [11]. Isoquercitrin has shown inhibition of urinary bladder and pancreatic cancer progress [12,13], as well as colon cancer suppression [14]. The profile of the bioactive components in different parts of *A. esculentus* (okra) is well documented: for its pod polyphenolic compounds, carotene, folic acid, thiamine, riboflavin, niacin, vitamin C, oxalic acid, and amino acids; for its seed polyphenolic compounds, mainly oligomeric catechins and flavonol derivatives, protein (i.e., high lysine levels), and oil fraction (in particular, its derived oil is rich in palmitic, oleic, and linoleic acids); for root carbohydrates and flavonol glycosides, and mainly minerals, tannins, and flavonol glycosides for leaves.

Despite the numerous studies mentioned above, there is a paucity of reports regarding the anticancer effects of *A. esculentus* (okra) seeds. Extractive techniques and conditions used recently are almost obsolete and must be optimized for higher efficiency. Leonardo G. Monte and his colleagues researched the anti-tumor properties of a

newly discovered lectin from *A. esculentus* (AEL) on human breast cancer (MCF7) and skin fibroblast (CCD-1089SK) cells. Their results show that AEL in its natural form could potentially create an anti-tumor effect on breast cancer cells and could be a therapeutic support for the treatment of human breast cancer [15]. Watcharaphong et al. [16] investigated the effects of okra seed extract (OSE) on the proliferation and viability of three human cancer cell lines (MCF-7, HeLa, and HepG2). It was found that both the naturally applied extract and the extract loaded into polymeric micelles had dose-dependent cytotoxic effects on all three cell lines. OSE-loaded micelles were more cytotoxic than natural OSE, especially in the HepG2 cell line. Additionally, it was found that OSE inhibited the migration of MCF-7 and HeLa cells. Hayaza and coworkers [17] conducted a study to examine raw polysaccharides extract of okra (*A. esculentus* L.) as an anticancer agent against Huh7it human liver cancer cells. Their extract could reduce the metabolism of cancer cells. It also has potential as an anti-cancer agent by inhibiting up to 15 % cell growth at the highest dose.

Glioblastoma multiforme (GBM) stands as one of the most formidable challenges in oncology, characterized by its aggressive infiltration into surrounding brain tissue, resistance to conventional therapies, and high rates of recurrence [18]. Despite decades of research and therapeutic advancements, the prognosis for GBM patients remains bleak, with a median survival of only 12-15 months post-diagnosis. This pressing clinical need underscores the imperative for novel therapeutic approaches that can effectively target GBM with improved efficacy and reduced toxicity [19]. C6 is a glial cell line isolated from the brain of a rat with Glioma. C6 glioma cells are frequently used for glioblastoma research since they have high mitotic activity and were determined to be the cell line most similar to the mechanism of human brain tumors. L929 cells play a pivotal role in biological science with versatile applications. They serve as a crucial resource for toxicity assessment, aiding in evaluating the potential adverse effects of various compounds. The cytotoxic and anticarcinogenic potentials of *A. Esculentus* seed extracts were determined for C6 rat glial cell and L929 rat fibroblastoma cell line using MTT at different concentrations.

A lot of research has been done on the antioxidants, anti-inflammatory, and anticancer activity of medical plants that are rich in polyphenols and flavonoids and are still being done [20,21]. While the interest of the biological activities of okra seed extract has not yet been explored for comparatively studied for cancer and normal cell lines. In addition, the importance of okra as a low-cost functional feed still deserves to be further highlighted in the potential of cancer treatment. Considering this, our study comprehensively evaluated the antioxidant properties and cytotoxic effects of *A. esculentus* seeds extract against C6 glioblastoma and L929 fibroblastoma cell lines using an in vitro model. The extractive techniques and conditions used in the present study are also optimized for higher efficiency.

Material and Method

Preparation of *A. Esculentus* Seed Extract

A. esculentus seeds were obtained from traditional markets in Turkiye. Then, they were washed thoroughly with distilled water to remove foreign material. To get the extract, briefly, the dried 6 g. of fresh *A. esculentus* seeds were rinsed with 30 mL of 70% ethanol in a shaking water bath (N-Biotek NB-30330) for 48 hours. The extracts were prepared using ethanol (70%) (ISOLAB), known as a good solvent for polyphenol extraction, safe for human consumption, low cost, and environmentally friendly in nature [22]. After 72 hours, the ethanol-soluble fraction was filtered through a 0.45 µm microporous membrane and concentrated in a vacuum oven (Nuve CO2 Incubator) at +50°C. It was then freeze-dried with a lyophilizer and stored in the refrigerator at +4°C until use.

Sigma-Aldrich provided suitable media and analytical-grade chemicals. All the chemicals were analytical grades and used without further modification or purification. Double-distilled and deionized water was used in the extraction and for rinsing the clusters.

Characterization of *A. esculentus* Seed Extract

A. esculentus extracts were prepared at a concentration of 2 mg/mL. In UV-Vis Spectrophotometry (Shimadzu UV-1700), the spectrums of each sample were used in the range of 200-800 nm by using a blank instead of pure water. At the same time, to analyze the chemical structure of the extract and the functional groups it contains, it was examined on an FT-IR spectrophotometer (Shimadzu Prestige 2100) in ATR mode, and FT-IR spectra ranging from 600 to 4000 cm⁻¹ were used.

Antioxidant Activity Analysis of *A. esculentus* Seed Extract

The antioxidant activity of *A. esculentus* seed extracts were tested using a 1,1-diphenyl-2-picryl hydroxyl (DPPH) technique. To make the stock solution, 0.63 milligrams of DPPH were dissolved in 15.98 mL of ethanol. Filtration of the DPPH stock solution using methanol yielded a usable 0.1 mM concentration DPPH mixture with an absorbance of around 0.973 at 517 nm. To ensure the reliability of the test, the study was conducted with 3 repetitions.

In a 96-well microplate, 100 µL of DPPH stock solutions were combined with the concentration of *A. esculentus* seed extracts between 0.5 -5 mg/mL. As for negative control, 100 µL of DPPH solution and 100 µL of pure water were added to the control wells. After that, the plate was kept in complete darkness for 30 min. The absorbance was therefore determined at 517 nm. The following formula was used to compute the percentage of antioxidants.

$$\% \text{ of antioxidant activity} = [(A_c - A_s) \div A_c] \times 100$$

where: A_c —Control reaction absorbance; A_s —Testing specimen absorbance.

The absorbance and % antioxidant activity values of seed extracts were obtained with SkanIt™ Software for Microplate Readers using the Elisa Plate Reader (Thermo

Scientific Multiskan Go). Antioxidant activity values of the extracts were calculated using the Microsoft Excel program.

Cell Culture

The ethanolic plant extracts were screened for cytotoxicity using C6 glioma (glioblastoma multiforme - GBM) and L929 fibroblastoma (mouse fibroblast) cell lines. The C6 and L929 cell lines were seeded into 25 cm² plastic tissue culture flasks in DMEM/F12 medium (Sigma), supplemented with 10% FCS, 1% penicillin/streptomycin into 500 ml of DMEM/F12 medium containing L-glutamine. Cells were grown at 25 °C in a humidified atmosphere (85% humidity). After achieving exponential growth (approximately 90% confluent), cells were detached from culture flasks by brief exposure to trypsin (0.25% in PBS, pH 7.2–7.4), according to the standard trypsinization methods. The detached cells were collected by centrifugation (Nuve NF 800, 1000 rpm, 5 min, 25 °C) and counted on Thoma slides.

MTT Assay

The cytotoxicity analyses of *A. esculentus* seed powders at different concentrations on C6 Glioma and L929 fibroblastoma cell lines were performed comparatively by 3-(4,5-dimethyl thiazolyl- 2-yl)- 2, 5-diphenyltetrazolium bromide (MTT) assay, (ISO B. 10993-5: biological evaluation of medical devices). Briefly, intact C6 and L929 cells were seeded into 96-well plates at a 1x10⁴ cells/ml ratio in DMEM containing 5% fetal bovine serum (FBS) and 1% penicillin-streptomycin. Cells were incubated at 37°C in a 5% CO₂ for 24 h. After the overnight incubation, the medium was replaced with a fresh medium containing a series of dilutions of increasing the *A. esculentus* seed powder concentrations applied to C6 and L929 cells from 50 µg/mL to 1000 µg/mL. C6 and L929 cells were seeded on wells without extracts as a control group. After 24 incubations, the medium from each well was replaced with an MTT reagent then plates were incubated at 37 °C for 3 h. Following incubation, formazan crystals formed in the cell were dissolved in DMSO, and absorbance was recorded at 570 nm with a 96-well plate reader. Results were obtained using SkanIt™ Software for Microplate Readers using the Elisa Plate Reader (Thermo Scientific Multiskan Go). Values are displayed as mean ± SD of a minimum of three times. Statistical analyses and IC50 values of the extracts were evaluated in GraphPad Prism 8 software to find the percent cell viability [23].

Results and Discussion

UV-VIS Spectrophotometer Results

The flavonoid isolate compounds were identified using a UV-Vis spectrophotometer to determine the absorbance value of the compound at the maximum wavelength. The UV-Vis spectrum of the *A. esculentus* extracts can be seen in Figure 1. The extracts caused a peak at a wavelength of 272 nm. It is known that tannins and phenolic compounds

peak in the 270-280 nm range [24]. The spectrum results obtained were the maximum wavelengths that may occur due to the absorption of the benzoyl ring on the flavonoid structure, mainly occurring at a wavelength of 256 nm [25].

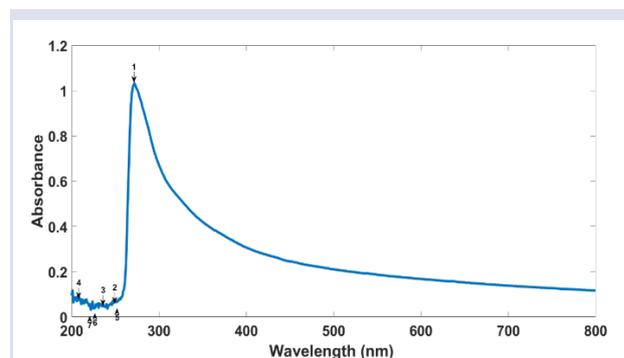


Figure 1. UV-Vis spectrophotometry of *A. esculentus* seed extract

FT-IR Spectrophotometer Results

The FTIR results shown in Figure 2 indicate the presence of various functional groups such as alkyl, ketone, aldehyde, carboxylic acids, esters, and amide in the aqueous and methanolic extracts of *A. esculentus*, respectively [26]. The FTIR confirmed that the so-called fingerprint region of the spectrum consists of characteristic bands between 1500 cm⁻¹ and 1000 cm⁻¹, which can be attributed to polysaccharides. The band at 925 cm⁻¹ belongs to the C-O stretching vibrations of alcohols/phenols [5]. The band located at 1041 cm⁻¹ might refer to the C=O stretching vibrations of carbonyl compounds. Absorption at 1590 cm⁻¹ due to ester carbonyl is detected and agrees with the okra spectrum of Archana et al. [27]. The band at 1402 cm⁻¹ belongs to the stretching vibrations of aromatic C=C bonds. The band at 3285 cm⁻¹ belongs to the hydrogen-bonded O-H stretching vibration due to water and carbohydrates [28]. The bands at 2851 cm⁻¹ belong to aliphatic compounds' asymmetric C-H stretching vibrations [29]. The bands at 552 cm⁻¹, 563 cm⁻¹ and 591 cm⁻¹ may be attributed to the C-O-H bending vibrations of aliphatic compounds or phenolic ring torsions [30,31]. Results agreed with those reported in the literature [32].

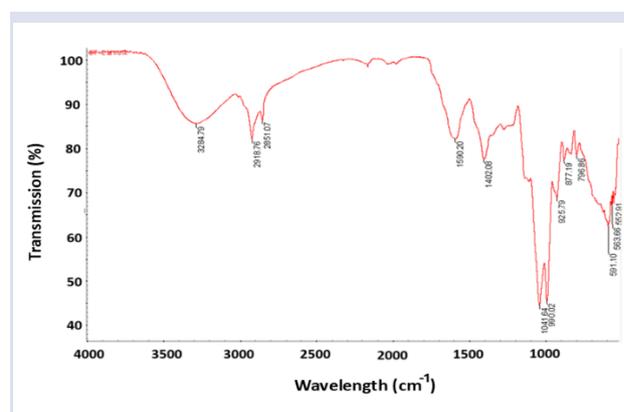


Figure 2. FT-IR spectrophotometry of *A. esculentus* seed extracts

Antioxidant Activity

The total antioxidant activity, measured by the DPPH method, of *A. esculentus* seed extracts was determined in a range of concentrations from 0.5 to 5 mg/mL (with an increase of 0.5 mg/mL). Antioxidants' ability to donate hydrogen or their radical scavenging impact on DPPH radical's scavenging activity. When a DPPH solution is combined with a material that can donate a hydrogen atom, diphenyl picryl hydrazine is produced in its reduced form, losing its violet color [33].

Table 1 and Figure 3 give the absorbance and % antioxidant activity values of seed extracts. The results showed that ethanolic *A. esculentus* extracts at different concentrations have remarkable antioxidant activities with concentration (dose) dependency.

Table 1. Absorbance and % antioxidant activity values of *A. esculentus* extract

| Concentration (mg/mL) | Extract+DPPH | Extract + Ethanol | Difference | Antioxidant Activity (%) |
|-----------------------|--------------|-------------------|------------|--------------------------|
| Control | 0,188 | - | - | - |
| 0,5 | 0,156 | 0,047 | 0,108 | 42,38 |
| 1 | 0,127 | 0,056 | 0,070 | 62,59 |
| 2 | 0,1185 | 0,058 | 0,0605 | 67,81 |
| 2,5 | 0,125 | 0,067 | 0,058 | 69,14 |
| 3 | 0,125 | 0,073 | 0,052 | 72,34 |
| 3,5 | 0,136 | 0,089 | 0,047 | 75,00 |
| 4 | 0,124 | 0,088 | 0,035 | 81,21 |
| 4,5 | 0,121 | 0,102 | 0,018 | 90,25 |
| 5 | 0,141 | 0,129 | 0,012 | 93,62 |

The results indicated a significant increase in the antioxidant activity at the highest concentration tested (93.62% at 5 mg/mL) compared to the lowest concentration (42.38% at 0.5 mg/mL). The antioxidant capacity of the extract should be related to its high phenolic content since polyphenols have been reported to possess antioxidant capacities. The content of flavonoids in a plant extract could also result in higher antioxidant activity.

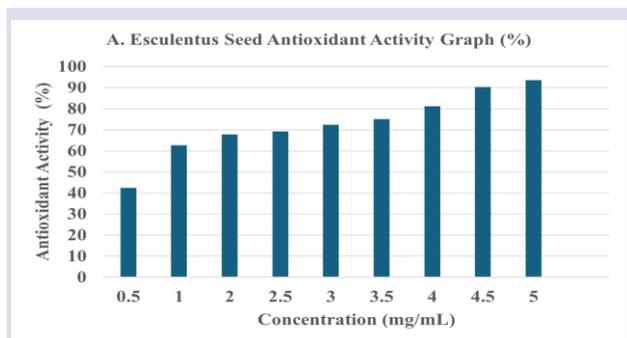


Figure 3. Antioxidant activity of *A. Esculentus* seed extracts

Cytotoxicity Results

The possible cytotoxic effects of ethanolic extracts were evaluated on L929 and C6 cells after 24 h of

incubation with different concentrations of extracts ranging from 50 µg/mL to 1000 µg/mL (50, 100, 250, 500, 1000 µg/mL), comparing the results with the control group, which consisted of cells incubated only with the medium.

The results showed that the cytotoxic effect of okra extracts depended on the extract dose and the cell line in question. As it is seen in Figure 4, *Abelmoschus Esculentus* seed extracts produced a significant reduction in the viability of the L929 cell line compared to the control samples at all concentrations. It was observed that cell vitality decreased from 86.1% to 10.2% by increasing the *A. esculentus* concentrations applied to L929 cells from 50 µg/mL to 1000 µg/mL. The IC₅₀ value for L929 cells was 431.45 µg/ml. Indeed, the determination of IC₅₀ was not achieved concerning the incubation times tested. Thus, the extract seemed to be toxic, dependent on the concentration range.

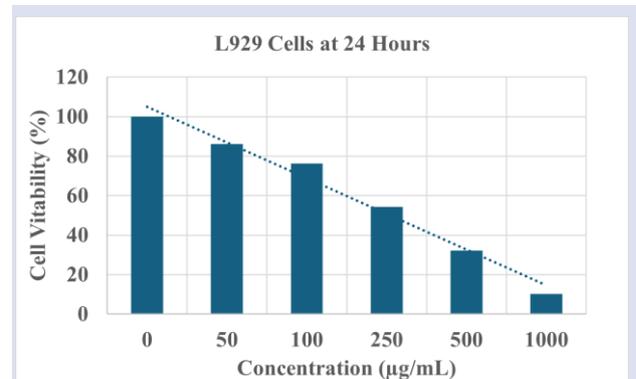


Figure 4. Cytotoxic activity (expressed as the percentage of viable cells) on L929 cell lines after 24 h of incubation with of *A. esculentus* seed extracts

When *A. esculentus* extracts were tested on the C6 Glioma cancer cell line, we observed a potentiated cytotoxic effect for the extracts, resulting in 72.19% C6 cell viability even at a lower concentration (50 µg/mL) than the control group. Figure 5 shows the cytotoxicity of *A. esculentus* extracts on C6 Glioma cells at the 24th hour. It was observed that by increasing the *A. esculentus* concentrations applied to C6 Glioma cells from 50 µg/mL to 1000 µg/mL, the vitality rate in the cells decreased from 72.2% to 12.9%. The IC₅₀ value for C6 glioma cells was 273.4 µg/ml. The sample concentrations of 1000µg/ml, 500 µg/ml, 250µg/ml, 100µg/ml, and 50µg/ml show %12.9, %15.4, %28.2, %52.5, and %72.2 cell viability values against the glioma cancer C6 cell line respectively. The cytotoxic effect observed depended only on concentration, not the time on both cell lines. *A. esculentus* extracts at different doses (50 to 1000 µg/mL) showed lower toxicity on normal L929 cells than the cancer C6 cells. The toxicity of extract on cancer cells may be due to the high metabolic rate of cells. Hence, cancer cells have a greater possibility of absorbing *A. esculentus* extracts than normal cells. There are additional reasons why cancer cell lines may have lower IC₅₀ values than normal cell lines, including strong extract interaction and

a high rate of cell division **Hata! Başvuru kaynağı bulunamadı..** The results are consistent with numerous studies that found that flavonoid substances may induce cancer cell apoptosis [38–39].

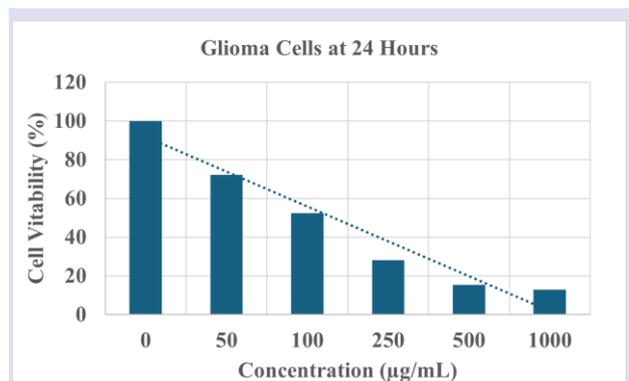


Figure 5. Cytotoxic activity (expressed as the percentage of viable cells) on C6 cell lines after 24 h of incubation with of *A. esculentus* seed extracts

Conclusion

The study indicates that the extract isolated from *A. esculentus* seeds has substantial anticancer and antioxidant activity at all concentrations. The toxicity results of extract at different doses (50 to 1000 µg/mL) showed on L929 cells lower than cancer cells. Due to the metabolic rate of cells, cancer cells may have a greater possibility of absorbing *A. esculentus* extracts than normal cells. The anticancer effect has been observed dose-dependent. The biological properties might be due to the synergistic actions of bioactive compounds such as a considerable amount of total phenolics and flavonoids. Many previous works have reported a highly significant relationship between total phenol content and antioxidant activity. The seed extract's antioxidant potential may be due to the presence of vitamins such as vitamin E. So, it could be concluded that the extract can act as a source of anticancer drugs and can be used to improve health status. There should also be a dependency on the dose use of ethanol, which offers the highest total phenolic contents. However, more research is needed to understand and implement these findings into clinical practice.

Further studies should also be conducted to evaluate the *A. esculentus* seed extract, which needs to be tested in animal models and clinical studies to understand better its effectiveness and possible side effects in cancer treatment. Additionally, researchers showed that the possibility of green synthesized nanoparticles with the help of *A. esculentus* can potentially act as anticancer and antibacterial agents. Therefore, scientists and healthcare professionals must continue to support studies and discoveries in this field. Investigating the potential of plant extracts in cancer treatment may contribute to developing more effective and natural treatment methods.

Conflicts of interest

The authors declare no conflict of interest.

References

- [1] Roy A., Shrivastava S.L., Mandal S.M., Functional properties of Okra *Abelmoschus esculentus* L. (Moench): traditional claims and scientific evidences, *Plant Sci. Today, Journal*, 1(3) (2014) 121-130.
- [2] Petropoulos S., Fernandes A., Barros, L., Ciric A., Sokovic M., Ferreira I.C.F.R., The chemical composition, nutritional value and antimicrobial properties of *Abelmoschus esculentus* seeds, *Food Funct. J.*, 8 (2017) 4733-4743.
- [3] Yasir H., A Review on The Potential Effects of Plant Metabolites, *International Journal of Pharmacy Research & Technology (IJPR)*, 13(1) (2023) 43-59.
- [4] Karagecili H., İzol E., Kirecci E., Gulcin İ., Determination of Antioxidant, Anti-Alzheimer, Antidiabetic, Antiglaucoma and Antimicrobial Effects of Zivzik Pomegranate (*Punica granatum*)—A Chemical Profiling by LC-MS/MS, *Life*, 13(3) (2023) 735.
- [5] Durazzo A., Lucarini M., Novellino E., Souto E. B., Daliu P., Santini A., Bioactive components beneficial properties focused on antidiabetic role for sustainable health applications, *Molecules*, 24(1) (2019) 24-38.
- [6] Islam M.T., Phytochemical information, and pharmacological activities of Okra (*Abelmoschus esculentus*): a literature based review, *Phytother Res.*, 33(1) (2019) 72-80.
- [7] Xia F., Zhong Y., Li M., Chang Q., Liao Y., Liu X., Pan R., Antioxidant and anti-fatigue constituents of okra, *Nutrients*, 7(10) (2015) 8846–8858.
- [8] Gemede H.F., Haki G.D., Beyene F., Woldegiorgis A.Z., Rakshit S.K., Proximate, mineral, and antinutrient compositions of indigenous okra (*Abelmoschus esculentus*) pod accessions: implications for mineral bioavailability, *Food Sci. Nutr.*, 4(2) (2016) 223–233.
- [9] Ngoc T.H., Ngoc Q.N., Van A.T.T., Phung N.V., Hypolipidemic effect of extracts from *Abelmoschus esculentus* L. (Malvaceae) on tyloxapol-induced hyperlipidemia in mice, *Journal of Pharm. Sci.*, 35 (2008) 42–46.
- [10] Sabitha V., Ramachandran S., Naveen K.R., Panneerselvam K., Antidiabetic and antihyperlipidemic potential of *Abelmoschus esculentus* (L.) Moench. in streptozotocin-induced diabetic rats, *J. Pharm. Bioallied. Sci.*, 3(3) (2011) 397–402.
- [11] Valentová K., Vrba J., Bancířová M., Ulrichová J., Křen V., Isoquercitrin: pharmacology, toxicology, and metabolism, *Food and Chemical Toxicology*, 68 (2014) 267–282.
- [12] Chen F., Chen X., Yang D., Isoquercitrin inhibits bladder cancer progression in vivo and in vitro by regulating the PI3K/Akt and PKC signaling pathways, *Oncology Reports*, 36(1) (2016) 165–172.
- [13] Chen Q., Li P., Li P., Xu Y., Li Y., Tang B., Isoquercitrin inhibits the progression of pancreatic cancer in vivo and in vitro by regulating opioid receptors and the mitogen-activated protein kinase signaling pathway, *Oncology Reports*, 33(2) (2015) 840–848.
- [14] Amado N. G., Predes D., Fonseca B. F., Isoquercitrin suppresses colon cancer cell growth in Vitro by targeting the Wnt/ β -catenin signaling pathway, *Journal of Biological Chemistry*, 289(51) (2014) 35456–35467.
- [15] Monte L.G., Santi-Gadelha T., Reis L.B., Braganhol E., Prietsch R.F., Dellagostin O.A. E., Lacerda R.R., Gadelha C.A., Conceição F.R., Pinto L.S., Lectin of *Abelmoschus esculentus* (okra) promotes selective antitumor effects in

- human breast cancer cells, *Biotechnol. Lett.*, 36(3) (2014) 461-9.
- [16] Chaemsawang W., Prasongchean W., Papadopoulos K., Ritthidej G., Sukrong S., Wattanaarsakit P., The Effect Of Okra (*Abelmoschus Esculentus* (L.) Moench) Seed Extract On Human Cancer Cell Lines Delivered In Its Native Form And Loaded In Polymeric Micelles, *Int. J. Biomater.*, (2019) 9404383.
- [17] Hayaza S., Anticancer activity of okra raw polysaccharides extracts against human liver cancer cells, *Tropical Journal of Pharmaceutical Research*, 18(8) (2019) 1667-1672.
- [18] Paulmurugan, R., and Massoud, T. F., Glioblastoma resistance to chemotherapy: molecular mechanisms and innovative reversal strategies, *Academic Press*, 1st Edition Volume 15 (2021) 828
- [19] Wu, W., Klockow, J. L., Zhang, M., Lafortune, F., Chang, E., Jin, L., Daldrup-Link, H. E., Glioblastoma multiforme (GBM): An overview of current therapies and mechanisms of resistance, *Phar. Research*, 171 (2021) 105780.
- [20] Kumar, A., Kumar, P. and Nadendla, R., A review on: *Abelmoschus esculentus* (Okra), *Int. Research J. of Phar. and Applied Sci.*, 3(4) (2013) 129-132.
- [21] Abamor E.S., A New approach to the treatment of leishmaniasis: quercetin-loaded polycaprolactone nanoparticles, *Journal of the Turkish Chemical Society Section A: Chemistry*, 5(3) (2018) 1071-1082.
- [22] Shi J., Nawaz H., Pohorly J., Mittal G., Kakuda Y. and Jiang Y., Extraction of polyphenolics from plant material for functional foods-engineering and technology, *Food Rev. Int.*, 21(1) (2005) 139-166.
- [23] Mansuroğlu B, Derman S, Kizilbey K, Ateş S.C, and Akdeste Z.M., Synthetic peptide vaccine for Foot-and-Mouth Disease: synthesis, characterization and immunogenicity, *Turkish Journal of Biochemistry*, 45(6) (2020) 859-868.
- [24] Patle T. K., Shrivastava K., Kurrey R., Upadhyay S., Jangde R. and Chauhan R., Phytochemical screening and determination of phenolics and flavonoids in *Dillenia pentagyna* using UV-vis and FTIR spectroscopy, *Spectrochimica Acta Part A: Mol. and Bio. Spect.*, 242 (2020) 118717.
- [25] Astutiningsih C., Kristanti J., Ariani W.L., Quercetin Compound of Okra Fruit (*Abelmoschus esculentus* L.) as Potential of Antioxidant and Antimicrobial, *Journal Biodjati.*, 8(2) (2023) 295-305.
- [26] Doreddula S.K., Bonam S.R., Gaddam D.P., Rao Desu B.S., Ramarao N., Pandey V., Phytochemical analysis, antioxidant, antistress, and nootropic activities of aqueous and methanolic seed extracts of ladies finger (*Abelmoschus esculentus* L.) in mice, *Sci. World Journal*, (2014) 519848.
- [27] Archana G., Sabina K., Babuskin S., Radhakrishnan K., Fayidh M.A., Babu P.A.S., Sivarajan M., Sukumar M., Preparation and characterization of mucilage polysaccharide for biomedical applications, *Carbohydr. Polym.*, 98(1) (2013) 89-94.
- [28] Dimopoulou M., Ritzoulis C., Panayiotou C., Surface characterization of okra hydrocolloid extract by inverse gas chromatography (IGC), *Colloid. Surface. A: Physicochem. Eng. Aspects*, 475 (2015) 37-43.
- [29] Cotrim M.A.P., Mottin A.C., Ayres E., Preparation and Characterization of Okra Mucilage (*Abelmoschus esculentus*) Edible Films, *Macromol. Symp.*, 367 (2016) 90-100
- [30] Zhang, W., Xiang, Q., Zhao, J., Mao, G., Feng, W., Chen, Y., Zhao, T. Purification, structural elucidation and physicochemical properties of a polysaccharide from *Abelmoschus esculentus* L (okra) flowers. *International journal of biological macromolecules*, 155 (2020) 740-750.
- [31] Alpaslan, D., Turan, A., Erşen Dudu, T., & Aktas, N. Evaluation of particle biosynthesis, p (Okra) particle bioactivity, and drug release properties using *Abelmoschus esculentus* (okra) plant extract. *Polymer Bulletin*, 81(5) (2024) 4557-4572.
- [32] Woumbo C.Y., Kuate D., Tamo D.G.M. and Womeni H.M., Antioxidant and antidiabetic activities of a polyphenol rich extract obtained from *Abelmoschus esculentus* (okra) seeds using optimized conditions in microwave-assisted extraction (MAE), *Front. Nutr.*, 28 (9) (2022) 1030385.
- [33] Mensor L., Menezes F., Leitao G., Screening of Brazilian plant extracts for antioxidant activity by the use of DPPH free radical method, *Phytotherapy Res.*, 15(2) (2001) 127-130.
- [34] Graham J., Agbenorhevi J. and Kpodo F., Total Phenol content and antioxidant activity of Okra seeds from different genotypes, *J. Food Nutr.*, 5(3) (2017) 90-94.
- [35] Augustyniak A., Bartosz G., Cipak A., Duburs G., Horáková L., Luczaj W., Majekova M., Odysseos A.D., Rackova L., Skrzydlewska E., Stefek M., Strosová M., Tirzitis G., Venskutonis P.R., Viskupicova J., Vranka P.S., Zarković N., Natural and synthetic antioxidants: an updated overview, *Free Radic. Res.*, 44(10) (2010) 1216-1262.
- [36] Manee S. and Kaewsrichan J., Cosmeceutical benefit of *Abelmoschus esculentus* L. seed extract, *J. of Pharm. Res. International*, 19(6) (2017) 1-11.
- [37] Ranjbar M., Dastani M., Khakdan F., Evaluation of antimicrobial, cytotoxicity, and catalytic activities of CuO-NPs synthesized by *tanacetum parthenium* extract, PREPRINT (Version 1) available at Research Square (2021)
- [38] Tedasen A., Sukrong S., Sritularak B., Srisawat T., and Graidist P., Trihydroxy-6,8-diprenylisoflavone and lupalbigenin, active components of *Derris scandens*, induce cell death on breast cancer cell lines, *Biomedicine & Pharmacotherapy*, 81 (2016). 235-241.
- [39] Palko-Łabuz A., Środa-Pomianek K., Wesołowska O., Kostrzewa-Susłow E., Uryga A., and Michalak K., MDR reversal and pro-apoptotic effects of statins and statins combined with flavonoids in colon cancer cells, *Biomedicine & Pharmacotherapy*, 109 (2019) 1511-1522.

Sisyrinchium micranthum (Iridaceae), a New Alien Record from Türkiye

Özgür Eminağaoğlu^{1,a}, Hayal Akyıldırım Beğen^{2,b,*}

¹ Department of Forest Engineering, Faculty of Forestry, Artvin Çoruh University, 08000 Artvin, Türkiye

² Health Services Vocational School, Artvin Çoruh University, 08000 Artvin, Türkiye

*Corresponding author

Research Article

History

Received: 24/09/2024

Accepted: 20/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Sisyrinchium is one of the largest genera in the Iridaceae family, including over 200 taxa. It is native to regions ranging from the Hawaiian Islands through temperate and subtropical areas of the Americas, extending down to the Falkland Islands. *S. angustifolium* Mill. is only recorded from Artvin, Kemalpaşa (Türkiye) from this genus. A newly alien species, *S. micranthum* Cav., has been reported as a new taxon for the Flora of Turkey from the eastern black sea region of Türkiye. *S. micranthum* is morphologically distinguished from *S. angustifolium* by urceolate-campanulate perianth, broadly acuminate tepal, 0.7–2 mm wide of node, 2–10 cm first internode, broadly acuminate petal apex and size of the capsule. Diagnostic characters, description and photographs are given in the manuscript. The distribution and habitat in Türkiye of the new record is determined. Taxonomic and conservation status of this species are evaluated.

Keywords: Aliien, Iridaceae, morphology, *Sisyrinchium*, Türkiye.

oeминаgaoğlu@artvin.edu.tr

<https://orcid.org/0000-0003-0064-0318>

h.akyıldırım@artvin.edu.tr

<https://orcid.org/0000-0003-2028-5827>

Introduction

The Iridaceae family includes 69 accepted genera [1]. *Sisyrinchium* L., a genus within the Iridaceae family, encompasses a diverse range of species, taxonomically making it the most complex genus in this family. The estimated number of species in *Sisyrinchium* is thought to range from approximately 60 to 212, depending on the taxonomist [1-3]. The genus is distributed across all biomes but occurs mainly in the South and Southeastern regions, particularly in the Atlantic Forest and Pampa (Fig 1) [4] with biome delimitation according to [5]. In Brazil, 71 species and 9 subspecies are currently recognized [4]. Only one species (*S. angustifolium*) was determined from Türkiye [6]. Most *Sisyrinchium* species are native to the America, with several becoming invasive weeds globally [7-9].

These plants are perennial herbs featuring bulbs or corms. Their leaves are typically clustered at the base of the stem, compressed laterally, mostly linear, basal sheathing and arranged equitant manner. Recent discoveries of new *Sisyrinchium* taxa have highlighted the need for comprehensive taxonomic studies to understand the genus' diversity fully. *Sisyrinchium* needs additional research to address subgeneric classifications, precisely identify taxa, and clarify synonyms for ambiguous taxa [10-13]. Based on earlier taxonomic research, the genus *Sisyrinchium* has been classified into two subgenera: *Sisyrinchium* and *Echthronema* [2,11,14].

S. micranthum Cav. is native to South America and Mexico [15,16]. It is introduced to Alabama, Albania, Arkansas, Azores, Canary Is., China Southeast, Dominican Republic, East Himalaya, Fiji, Florida, Galápagos, Georgia, Hawaii, India, Iran, Italy, Japan, Korea, Louisiana, Madagascar, Malaya, Mauritius, Mississippi, New

Caledonia, New South Wales, Norfolk, North Carolina, Puerto Rico, Queensland, Réunion, Samoa, Sardegna, South Australia, South Carolina, Spain, Tasmania, Texas, Tibet, Tubuai [1,17-19] (Fig. 1).

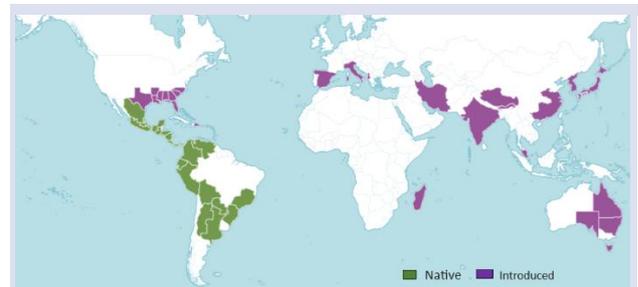


Figure 1. Worldwide distribution map of *S. micranthum* (IPNI, 2024).

This study presents a detailed examination of new record of *S. micranthum* based on voucher specimens, original characterization and contemporary literature [10, 15, 17] To facilitate comparison, images of *S. angustifolium* and an depiction of *S. micranthum* are provided.

Materials and Methods

Sisyrinchium specimens were collected from Karadeniz Center, Ortahisar (Trabzon), Fındıklı (Rize), Kemalpaşa (Artvin) Türkiye in May-October 2024 (Fig. 2). Plant material for this study was collected exclusively from three sites located around of the eastern black region, Türkiye. The collected materials were critically studied. Five specimens were measured for the morphological analyses.



Figure 2. Collected locations of *Sisyrinchium micranthum* samples from Türkiye; A-Trabzon, B- Rize, C-Artvin

Morphological characteristics of *S. micranthum* were determined by examining of these collected specimens and by comparing them to data from previous studies [6, 14, 20] and herbarium materials at ARTH (Artvin Çoruh University Herbarium). All collected specimens have been deposited at ARTH. Population and habitat data are based on field observations.

Results and Discussions

Sisyrinchium micranthum Cav., Diss. 6: 144, tab. 191, 1788. (Fig. 3, 4).

Sisyrinchium micranthemum Pers. in Syn. Pl. 1: 50 (1805); *Sisyrinchium iridifolium* Kunth in F.W.H.von Humboldt, A.J.A.Bonpland & C.S.Kunth, Nov. Gen. Sp. 1: 324 (1816); *Marica iridifolia* (Kunth) Ker Gawl. in Bot. Reg. 3: t. 229 (1817); *Marica micrantha* (Cav.) Ker Gawl. in Irid. Gen.: 22 (1827); *Sisyrinchium dichroum* Poepp. ex Klatt in C.F.P.von Martius & auct. suc. (eds.), Fl. Bras. 3(1): 537 (1871); *Sisyrinchium fimbriatum* Dombey ex Klatt in C.F.P.von Martius & auct. suc. (eds.), Fl. Bras. 3(1): 537 (1871); *Bermudiana iridifolia* (Kunth) Kuntze in Revis. Gen. Pl. 2: 699 (1891); *Bermudiana micrantha* (Cav.) Kuntze in Revis. Gen. Pl. 2: 700 (1891); *Bermudiana bermudiana* var. *micrantha* (Cav.) Kuntze in Revis. Gen. Pl. 3(3): 307 (1898); *Sisyrinchium rosulatum* E.P.Bicknell in Bull. Torrey Bot. Club 26: 228 (1899); *Sisyrinchium exile* E.P.Bicknell in Bull. Torrey Bot. Club 28: 573 (1901); *Sisyrinchium brownii* Small in Contr. NewYork Bot. Gard. 327: 330 (1931); *S. micranthum* f. *flavum* Ravenna in Onira 5: 57 (2001); *S. micranthum* f. *eburneochraceum* Ravenna in Onira 5: 57 (2001); *S. micranthum* f. *purpureum* Ravenna in Onira 5: 57 (2001); *S. micranthum* f. *luteum* Ravenna in Onira 5: 58 (2001); *Sisyrinchium micranthum* subsp. *scudiculare* Ravenna in Onira 5: 57 (2001).

A herbaceous annual plant typically has thin tufts, but it can also be a short-lived rosulate perennial that turns yellowish green when dry. Stems 10–15(–25) cm, compressed, usually 2–3 in a tuft, up to 10 in more vigorous individuals, rarely simple, glabrous, with 1–2 nodes. Leaf blades 15–50(–70) × 1–2.5 mm, linear-lanceolate, glabrous, straight, acuminate, sparsely scabrous on margins. Inflorescences terminal, with 2–5 pedunculated flowers, spathes 15–25(–30) mm long,

conspicuously compressed, keeled, glabrous, usually entire, with thin, 0.2–0.3 mm wide, scarious margins. Perianth urceolate-campanulate basally, spreading in the upper half, tepals 5–10 mm long, acute to aristate, the spreading part pale lavender, tinged purple at the base, urceolate part yellowish to ochroleucous with purple patterns or strips, hairy outside, filaments connate basally, occasionally to 1/2 their length, the ovary ovate, up to 1 mm long, green, patent-hairy.

Capsules 2–3 mm, globose, tan with purplish sutures, pedicels spreading to arcuate. Seeds 0.5–1 mm, ± globose to slightly compressed, black, surface rugulose to finely alveolate.



(a)



(b)

(c)

(d)

Figure 3. *Sisyrinchium micranthum*; a-habit, b- seed capsule, c- flowers, d- leaves

Flowering: Between March and July.

Specimens examined: Türkiye: A7 Trabzon, Pelitli, Merkez, in front of the Trabzon airport, 32 m, 15 July 2024, 40°59'40"N, 39°46'45"E, 0-5 m, OEmin. 22926 (ARTH 17303); A8 Rize, Ardeşen, Fındıklı, beachside, 01 October 2024, 41°16'29"N, 41°08'41"E, 0-9 m, OEmin. 23809 (ARTH 18326); A8 Artvin, Kemalpaşa, Fındıklı, beachside, 03 October 2024, 41°30'09"N, 41°32'02"E, 10-35 m, OEmin. 23810 (ARTH 18327).

Distribution: South America, Mexico Australia, Iran, India, Japan, China.

Ecology: *S. micranthum* distributed on roadsides at 0-10 m in association *Oxalis corniculata* L., *Plantago major* L. *Trifolium repens* L., *Prunella vulgaris* L., *Duchesnea*

indica (Andrews) Focke, *Hydrocotyle ramiflora* Maxim., *Juncus inflexus* L., *Potentilla erecta* (L.) Rausch.

Conservation status: *S. micranthum* is known from eastern black sea region. The species' natural habitat is endangered by transportation, human activities, and the development of roads [21].

Sisyrinchium micranthum taxon was collected for the first time from the eastern black sea region of Türkiye and with this study, it was included as a new record in the flora of Türkiye.



Figure 4. Tepal apex forms; a-*S. micranthum*, b- *S. angustifolium*

A key to distinguish *Sisyrinchium micranthum* from *S. angustifolium*, which has been previously recorded in the flora of Türkiyey, has been established. [6,22]

Distinguish Key to the species of genus *Sisyrinchium* in Türkiye (Fig. 5-7)

- 1 Perianth stellate-rotate, widely spreading from base; tepals abruptly acuminate, cristate, 5–10 mm, pale blue; filaments connate ± entirely; capsule ca. 4-7 mm.....*S. angustifolium*

- 2 Perianth urceolate-campanulate basally; tepals broadly acuminate, 7.5–12.5 mm, pale lavender in the spreading part; filaments connate to 1/2 of their length; capsule ca. 2-3 mm.....*S. micranthum*

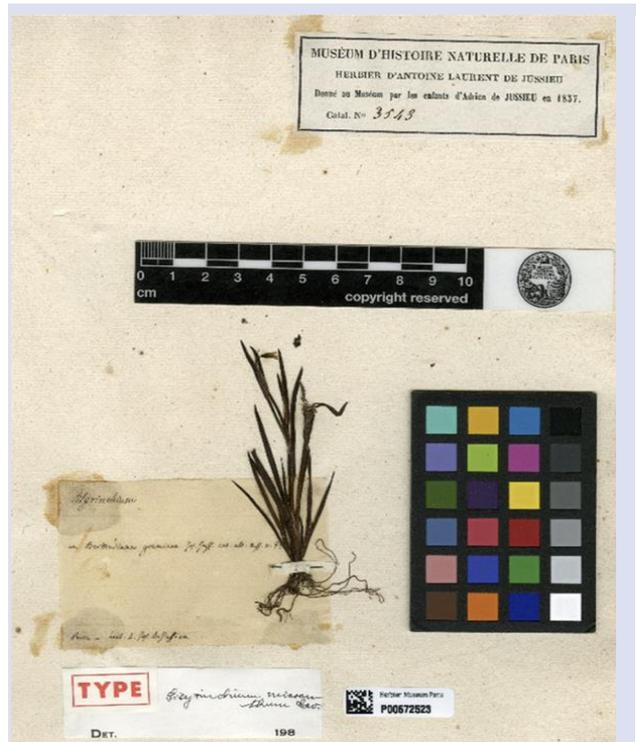


Figure 5. Type specimens of *S. micranthum* in National Museum (France).



Figure 6. Herbarium specimens of *S. micranthum* (ARTH 17303)

Sisyrinchium micranthum are new record and a newly naturalized species for the Türkiye flora and an alien species and it is thought that it probably entered our country via Georgia and grows in the Eastern Black Sea region, as it can easily expand its distribution area.



Figure 7. Herbarium specimens of *S. angustifolium* (ARTH 16705)

Table 1. Distinctive characters of *S. angustifolium* and *S. micranthum*

| | <i>S. angustifolium</i> | <i>S. micranthum</i> |
|-----------|--|--|
| Habit | Up to 45 cm | 15 cm |
| Leaves | 3-5 mm broad | 1-3 mm broad |
| Stems | 1-2 nodes, 2.3-5 mm wide, first internode 10-30 cm, usually longer than leaves | 1-2(-3) nodes, 0.7-2 mm wide, first internode 2-10 cm, usually shorter than leaves |
| Flowers | 1-4 | 2-5 |
| Perianth | stellate-rotate | urceolate-campanulate basally |
| PetalApex | abruptly acuminate, cristate, acuminate to acute | acute, rarely aristate, broadly acuminate, occasionally erose |
| Tepal | pale blue to violet, occasionally white, bases yellow; outer tepals 7.7-12.5 mm | maroon or pink to lavender-rose with purple stripes, or yellow with rosy purple bases; outer tepals 5-11mm |
| Filaments | usually more or less adnate to the tube, connate ± entirely, stipitate-glandular basally | connate basally, occasionally to 1/2 their length |
| Capsules | 4 -7 mm long, ± globose | 2-4 mm, globose, tan with purplish sutures |
| Seeds | 0.5-1.2 mm, globose to obconic, lacking obvious depression, rugulose | 0.5-1 mm, ± globose to slightly compressed, black, surface rugulose to finely alveolate |

Conclusions

Sisyrinchium micranthum is distributed near roadsides on the eastern blacksea region of Türkiye. Once the flowering period is over, there may be difficulties in species identification as they resemble Poaceae species. It can be damaged by human and transportation activities on the beach and roadside. Its populations need to be observed and monitored in order to continue its existence and their habitats. This species, like the species of the genus *Iris*, *Crocus* and *Gladiolus* in the Iridaceae family, possesses decorative and aesthetic features. For this reason, it is recommended to cultivate and exhibition it in botanical gardens. Since its natural habitats are under threat, this approach will also ensure its conservation.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] IPNI, International Plant Names Index. Available at: <http://www.ipni.org>, The Royal Botanic Gardens, Kew, Harvard University Herbaria & Libraries and Australian National Herbarium. Retrieved July 02, 2024.
- [2] Dani J.G., Inacio C.D., Souza-Chies T.T., OntheTrail of Morphological Traits: Morphometry Helps to Clarify Morphological Variation in Section Viperella (*Sisyrinchium*), *Plants*, 13 (2024) 2183.
- [3] Barker C., World checklist of Iridaceae. The Board of Trustees of the Royal Botanic Gardens, Kew. Available at: <http://apps.kew.org/wcsp>. Retrieved June 28, 2024.
- [4] Eggers L., Inácio C.D., *Sisyrinchium*. In: Flora do Brasil, Jardim Botânico do Rio de Janeiro. Available at: <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB8066>. Retrieved June 28, 2023
- [5] IBGE, Biomas e Sistema Costeiro-Marinho do Brasil. Available at: <https://www.ibge.gov.br/geociencias/cartas-e-mapas/informacoes-ambientais/15842-biomas.html?=&t=downloads>. Retrieved June 8, 2021.
- [6] Eminağaoğlu Ö., Özcan M., Systematics of *Sisyrinchium angustifolium* Mill. (Iridaceae) a newly recorded species from Turkey, *Bangladesh J. Plant Taxon*, 21(2) (2014) 175-180.
- [7] Esler A.E., Astridge S.J., The naturalisation of plants in urban Auckland, New Zealand 2. Records of introduction and naturalisation, *New J. Bot.*, 25 (1987) 523-537.
- [8] Miyawaki S., Washitani I., Invasive alien plant species in riparian areas of Japan: the contribution of agricultural weeds, revegetation species and aquacultural species, *Glob. Environ. Res.*, 8 (2004) 89-101.
- [9] Wu S-H., Hsieh C.F., Chaw S.M., Rejmánek M., Plant invasions in Taiwan: insights from the flora of casual and naturalized alien species, *Diversity and Distributions*, 10 (2004) 349-362.
- [10] Alves T.L.S., Chauveau O., Eggers L., de Souza-Chies T.T., Species discrimination in *Sisyrinchium* (Iridaceae): assessment of DNA barcodes in a taxonomically challenging genus, *Mol. Ecol. Resour.*, 14 (2014) 324-335.
- [11] Cocucci A.A., Vogel S., Oil-producing flowers of *Sisyrinchium* species (Iridaceae) and their pollinators in southern South America, *Flora*, 196 (2001) 26-46.
- [12] Goldblatt P., Rudall P., Henrich J.E., The genera of the *Sisyrinchium* Alliance (Iridaceae: Iridoideae): phylogeny and relationships, *Syst. Bot.*, 15 (1990) 497-510.
- [13] Souza-Chies T.T., dos Santos E.K., Eggers L., Flores A.M., Alves E.M.S., Fachinetto J., Lustosa J., Corrêa L.B., Tacuatiá L.O., Piccoli P., Miz R.B., Studies on diversity and evolution of Iridaceae species in southern Brazil, *Genet. Mol. Biol.*, 35 (2012) 1027-1035.
- [14] Cholewa A.F., Henderson D.M., Flora of North America Editorial Committee. Flora of North America north of Mexico. *FNA*, 26 (2002) 351-371.
- [15] Cavanilles A.J., *Sisyrinchium micranthum*, *Cav Sexta Dissertatio Botanica Diss*, 6 (1788) 345.
- [16] Goldblatt P., Manning J.C., The Iris Family: Natural History and Classification, Timber Press, Portland, OR, (2008) 290.
- [17] Chauveau O., Eggers L., Raquin C., Silvério A., Brown S., Couloux A., Cruaud C., Kaltchuk-Santos E., Yockteng R., Souza-Chies T.T., Nadot S., Evolution of oil-producing trichomes in *Sisyrinchium* (Iridaceae): insights from the

- first comprehensive phylogenetic analysis of the genus, *Ann. Bot.* 107 (2011) 1287-1312.
- [18] Tacuatiá L.O., Souza-Chies T.T., Flores A.M., Eggers L., Siljak-Yakovlev S., Kaltchuk-Santos E., Cytogenetic and molecular characterization of morphologically variable *Sisyrinchium micranthum* (Iridaceae) in southern Brazil, *Bot. J. Linn.*, 169 (2012) 350-364.
- [19] Shin H.W., Kim M.Y., Lee N.S., First report of a newly naturalized *Sisyrinchium micranthum* and a taxonomic revision of *Sisyrinchium rosulatum* in Korea, *Korean J. Pl. Taxon*, 46(3) (2016) 295-300.
- [20] Bicknell E.P., *Sisyrinchium rosulatum* E.P. Bicknell, *Bull. Torrey Bot. Club*, 26 (1899) 228.
- [21] IUCN, Standards and Petitions Subcommittee, Guidelines for Using the IUCN Red List Categories and Criteria. Version 14. Available at: <http://www.iucnredlist.org/documents/RedListGuidelines.pdf>. Retrieved June 9, 2023.
- [22] Eminağaoğlu Ö., *Sisyrinchium angustifolium* Mill. In: Güner A. (Ed), *Resimli Türkiye Florası*, ANG Vakfı Nezahat Gökyiğit Botanik Bahçesi Yayınları, 2022, 827-830.

Utilization of Multiwalled Carbon Nanotubes as Adsorbent for Removal of Acid Violet 17 Dye: Equilibrium, Kinetics and Thermodynamic Studies

Neşe Keklikcioğlu Çakmak^{1,a}, İlknur Şentürk^{2,b,*}

¹ Department of Chemical Engineering, Faculty of Engineering, Sivas Cumhuriyet University, Sivas 58140, Türkiye

² Department of Environmental Engineering, Faculty of Engineering, Sivas Cumhuriyet University, Sivas 58140, Türkiye

*Corresponding author

Research Article

History

Received: 14/01/2024

Accepted: 07/09/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

The adsorption capacity and separation efficiency of multiwalled carbon nanotube (MWCNT) were investigated for the removal of Acid Violet 17 (AV 17) dye in an aqueous solution. The effect of reaction conditions, such as contact time, AV 17 initial concentration, pH, adsorbent dosage, and temperature, on the adsorption capacity of MWCNTs was investigated. It was found that a contact time of 180 min, an adsorbent dosage of 0.8 g/L, and 100 mg/L AV 17 concentration are ideal conditions for maximum adsorption capacity (119.45 mg/g). AV 17 dye adsorption on MWCNTs followed a pseudo-second-order kinetic model ($R^2 = 0.9998$) and the Langmuir isotherm model ($R^2 = 0.9813$) suggesting that adsorption was the uniform and homogenous process. The maximum adsorption capacity from the Langmuir isotherm model was determined to be 322.58 mg/g. Adsorption of AV 17 dye was determined to be spontaneous and thermodynamically favorable since the values of ΔG° were negative. Positive ΔH° and ΔS° values suggest that the process is endothermic and randomness. The research results showed the MWCNTs could be used successfully to remove dye for the water treatment process.

Keywords: Multiwalled carbon nanotube, Acid violet 17, Adsorption, Adsorption mechanism.

^a nkeklkcioglu@cumhuriyet.edu.tr ^{ORCID} <https://orcid.org/0000-0002-8634-9232>

^b ilknursenturk@cumhuriyet.edu.tr ^{ORCID} <https://orcid.org/0000-0002-8217-2281>

Introduction

Dyes, one of the largest and most important groups of chemicals, especially Azo group dyes, are widely used for textile and leather dyeing [1,2]. Dyes contain different compounds with toxic, carcinogenic, and mutagenic properties as well as unknown environmental behaviors [3]. For this reason, dyes originating from industrial discharges affect the natural beauty of the streams or rivers where the discharge is made and also have highly toxic and harmful effects on the flora and fauna in aquatic life [1]. Approximately 70% of all commercially used dyes are azo dyes, and approximately 20% of these dyes can enter wastewater and cause an environmental problem. Due to the health and environmental concerns associated with synthetic dyes, it is necessary to fully treat the dyes before discharge into natural water bodies [4].

The chemical oxygen demand is raised when dyes are present in the water structure. Light penetration into the water is impeded by its impact on the photosynthetic activities of aquatic plants. Dyes are hazardous even at very low doses, and treating them is highly challenging due to their complexity. Due to the stability of dyes against heat, light, and microbial attack, the desired efficiency cannot be achieved with conventional wastewater treatment methods. Therefore, appropriate technologies should be developed to protect the ecosystem [3]. Different methods such as sorption, chemical flocculation, flotation, sedimentation, chemical oxidation, filtration technologies, photocatalytic and electrochemical oxidation, and biological techniques are

used to remove dyes from the solution before discharge into the natural environment. However, these techniques have some limitations such as secondary pollution formation and high cost. Among the different applicable techniques, adsorption is widely used for the removal of dyes from wastewater. The sorption process has a definite advantage over other methods used for dye recovery because it is a simple, selective, economical, and efficient process, does not generate hazardous by-products, and is flexible in operation [5,6].

Acid Violet 17 (AV 17), the model dye for this research, is an anionic azo dye. It is widely used in the paper, textile, food, and cosmetics industries. Since it is a dye related to the trimethyl methane class, it shows carcinogenic properties. Therefore, to avoid hazards caused by AV 17 in wastewater, this wastewater must be treated [4].

There are some previous studies conducted by researchers for the treatment of AV 17 dye. For example, Şentürk and Alzein [7] studied the adsorption of AV 17 with pistachio shells activated with 10% N H₂SO₄ and obtained 93.04% removal efficiency. The maximum Langmuir adsorption capacity (q_m) was determined as 26.455 mg/g for 160 mg/L AV 17. In another study, biosorbents obtained from the activation of fallen leaves of *Ficus racemosa* with NaOH and H₂SO₄ were used for the adsorption of AV 17 dye. The q_m value for raw biosorbent, H₂SO₄-activated biosorbent, and NaOH-activated biosorbent was determined as 45.25, 61.35, and 119.05 mg/g respectively [8]. Thinakaran vd., [9] studied the

removal of AV 17 dye by activated carbon prepared from sunflower seed shells. The q_m value of the adsorbent was 65.78 mg/g. Jain vd., [5] synthesized a new adsorbent from *Salvadora persica* powder by chemical activation to remove AV 17 from aquatic solution. The q_m values for raw and activated adsorbent were determined as 74.28 and 141.30 mg/g, respectively.

Carbon nanotubes (CNTs) have become a novel adsorbent for the removal of dyes from wastewater in recent times. Carbon nanotubes, which consist of both single-walled and multiwalled carbon nanotubes (SWCNTs and MWCNTs), are hexagonal lattice-shaped cylindrical nanostructures with a large specific surface area (100–1000 m²g⁻¹) that are primarily made of carbon atoms in sp² hybridization [10]. The physical characteristics of the nanomaterial and the structure of the adsorbed dye determine how dyes adsorb on carbon nanotubes. Because of CNTs unique chemical structure, large specific surface area, small sizes, layered and hollow structure, abundant adsorption sites, easy attachment of functional groups, and the ability to be decorated with other nanoparticles. They are a desirable substitute for eliminating dyes, fluoride, heavy metals, and other organic and inorganic pollutants from wastewater and water [11,12].

The present investigation aims to explore the adsorption behavior and removal capacity of MWCNTs, and other operating parameters affecting the adsorption of AV 17 from aquatic solution. In batch experimental studies, parameters known to affect adsorption efficiency were investigated. Isotherm, kinetic, and thermodynamic models were calculated to investigate the adsorption mechanism, to understand the process kinetics, and to see the effect of temperature on the process. So far, many researchers have reported the removal of different dyes and pollutants from aqueous solution by a multiwalled carbon nanotube nanocomposite. They investigated adsorption kinetics, the adsorption capacity of the adsorbent, and the effects of adsorption dosage and solution pH values on removal efficiency. However, these studies were carried out at lower concentrations with cationic azo dyes. This study differs in terms of investigating the treatment of high concentrations of AV 17, an anionic azo dye, with MWCNT.

Material and Method

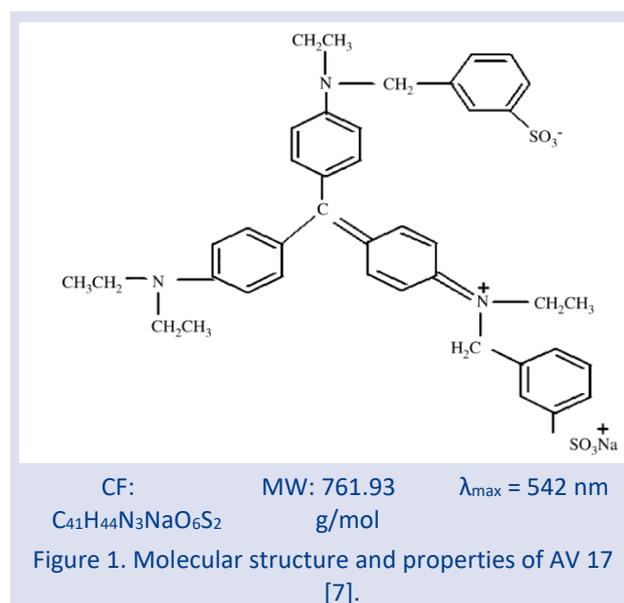
Adsorbent

Analytical-grade chemicals were all employed in this investigation. Multi walled carbon nanotubes (MWCNTs) were used exactly as they were received from Nanografi (Turkey), with an average outside diameter of 8–28 nm.

Preparation of the Dye Solution

Acid Violet 17, an anionic azo dye, was selected for this study. The chemical structure and properties of AV 17 dye are shown in Figure 1. The dye was weighed and dissolved in distilled water and the stock dye solution at a

concentration of 1 g/L was diluted to the desired working concentrations and used in batch experiments.



Adsorption Experiments

Experimental studies were carried out in a 10 mL glass tube at room temperature (25°C). Batch experiments consisted of contact time, pH, adsorbent dosage, initial dye concentration, and temperature studies. The time-dependent behavior of dye adsorption was studied by varying the adsorption time. The natural pH of the prepared dye solution was measured as 6.52. Glass tubes containing the prepared mixture were placed in an incubator shaker (Gerhardt, Germany) and stirred at 150 rpm at 25°C. Each test tube was then analyzed separately at the time intervals determined for the analysis to determine the concentrations of unadsorbed dye. Adjustments required during the pH study were carried out with dilute HCl and/or NaOH solutions. Other experimental conditions are given in Table 1.

Table 1. Batch experimental conditions

| Experiment | Solution pH | Adsorbent dosage (g/L) | Initial AV 17 conc. (mg/L) | Contact time (hours) | Temperature (°C) |
|-------------------------|-------------|------------------------|----------------------------|----------------------|------------------|
| Adsorbent concentration | 6.52 | 0.1-1.0 | 100 | 3 | 25 |
| Effect of pH | 2-12 | 0.8 | 100 | 3 | 25 |
| Effect of time | 6.52 | 0.8 | 100 | 5-330 min | 25 |
| Effect of concentration | 6.52 | 0.8 | 100-300 | 3 | 25 |
| Effect of temperature | 6.52 | 0.8 | 200 | 3 | 25-35-45-55 |

After the contact time was determined in all experimental studies, the loaded MWCNT adsorbent was centrifuged at 4000 rpm for 5 min and separated from the solution. The filtrate was then analyzed using a UV/Vis

spectrophotometer (Spectroquant Pharo 300, Merck) at a maximum wavelength of 542 nm and the concentrations of AV 17 remaining in the solution were determined. The adsorption efficiency (%) and adsorption capacity (q_e , mg/g) of the adsorbent used in the study were calculated as given below.

$$\text{Adsorption efficiency, } R (\%) = \frac{C_0 - C_e}{C_0} * 100 \quad (1)$$

$$\text{Adsorption capacity } (q_e) = [(C_0 - C_e) * V] / m \quad (2)$$

Where,

q_e : Adsorption capacity of the adsorbent (mg/g),

V: Solution volume (L),

C_0 : Initial AV 17 concentration in the solution (mg/L),

C_e : Concentration of AV 17 remaining in the solution after treatment (mg/L),

m: Adsorbent quantity (g).

Results and Discussion

Characterizations of MWCNTs

SEM and EDX analysis

The surface morphology of MWCNTs was characterized by scanning electron microscopy (SEM).

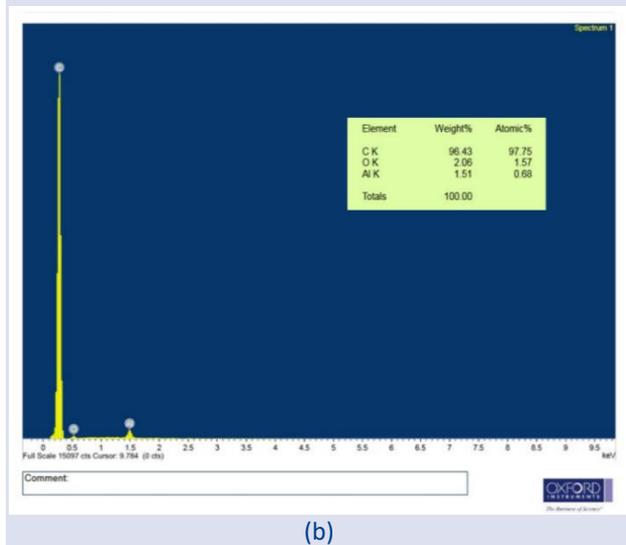
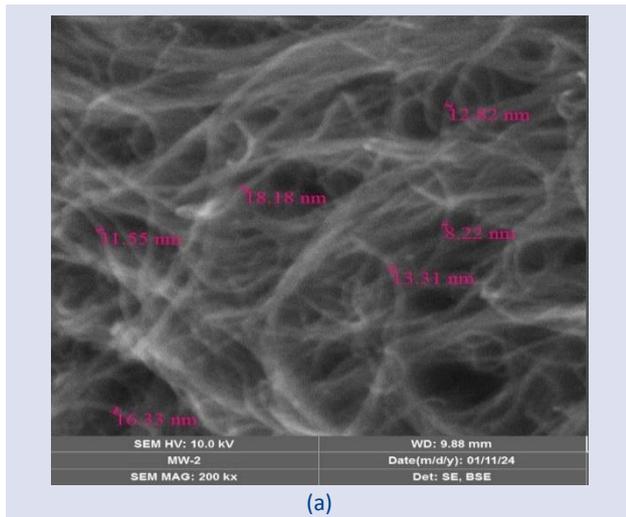


Figure 2. (a) SEM image of MWCNT (b) EDX spectra of MWCNT before the adsorption of AV 17 dye.

Figure 2 shows the MWCNTs' SEM and EDX images. The image indicates clearly that the MWCNTs are curved, twisted together, and cylindrical. As seen in Figure 2(a), MWCNTs have lengths ranging from several to tens of nanometers and diameters between 8 and 18 nm.

The presence of groups on the surface of carbon nanotubes is quantitatively analyzed using energy-dispersive X-ray spectroscopy (EDX) measurement. Figure 2(b) demonstrates the results of the MWCNTs. The data indicates that the sample contains oxygen in addition to carbon.

Figure 3 shows the SEM image and EDX result of MWCNT after AV 17 dye adsorption. When compared with the pre-adsorption image of MWCNT, it was observed that the spaces in the fibrous structure of MWCNT were filled. In the EDX result after adsorption, Na, Al and S peaks caused by AV 17 ($C_{41}H_{44}N_3NaO_6S_2$) dye are observed. SEM and EDX results show that MWCNT adsorbed AV 17 dye.

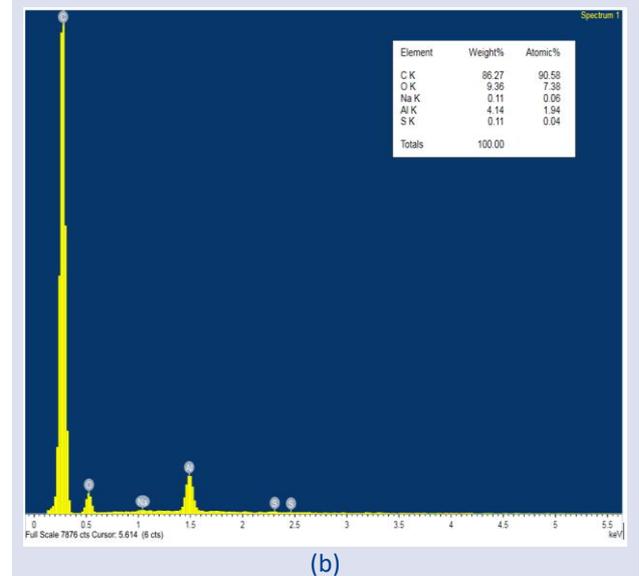
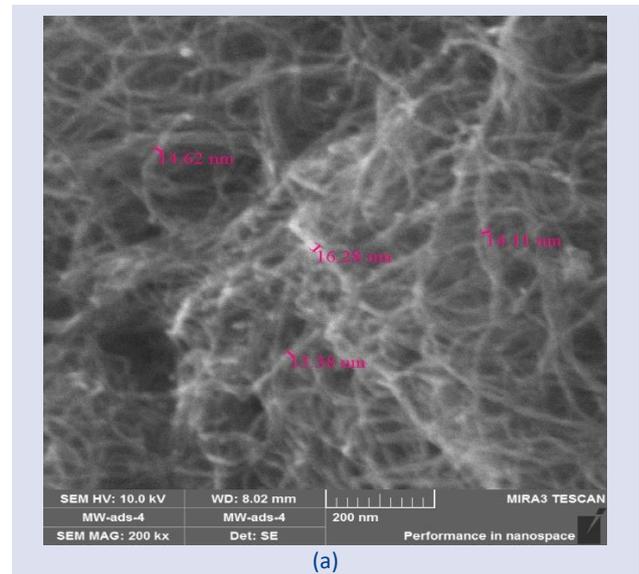


Figure 3. (a) SEM image of MWCNT (b) EDX spectra of MWCNT after the adsorption of AV 17 dye.

XRD analysis

The XRD pattern of the MWCNTs is shown in Figure 4. The diffraction patterns of typical graphite (002), (100), (004), and (110) are attributed to the intense diffraction

peak around $2\theta = 26^\circ$ and the low intense diffraction peaks about 44° , 54° , and 79° in the pattern. The MWCNT XRD graph and results are consistent with previous research.

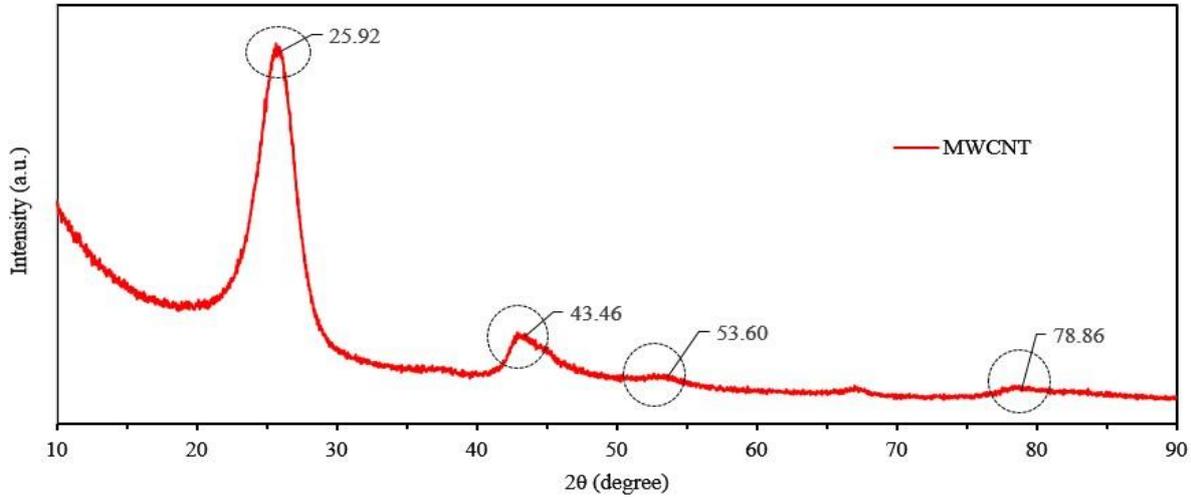


Figure 4. XRD patterns of MWCNT.

Adsorption studies

To investigate the adsorption behavior of AV 17 dye on the MWCNT adsorbent selected for this study, the effect of parameters such as contact time, pH, adsorbent amount, dye concentration, and temperature were investigated and the results were presented in detail.

Effect of adsorbent dosage

The unquestionably effective parameter in the adsorption process is the amount of adsorbent used in the solution. The adsorption of dye molecules becomes easier as the active adsorbent sites increase when the adsorbent dosage increases [13]. However, a further increase in the amount of adsorbent after a certain dose does not show a

significant change in dye adsorption. In the study investigating Reactive Black 5 adsorption with $Fe_3O_4@Granite$ magnetic adsorbent, results supporting this information were reported [13].

To investigate the effect of adsorbent dosage on AV 17 dye removal, the results obtained in the adsorption study at 100 mg/L initial AV 17 dye concentration are given in Figure 5. As the amount of adsorbent increased between 0.1 and 1.0 g/L, the dye removal increased because the effective surface area also increased. However, it was observed that the increase in dye removal was quite insignificant in adsorbent applications above 0.8 g/L (Figure 5). Therefore, the most suitable adsorbent amount was determined as 0.8 g/L MWCNT.

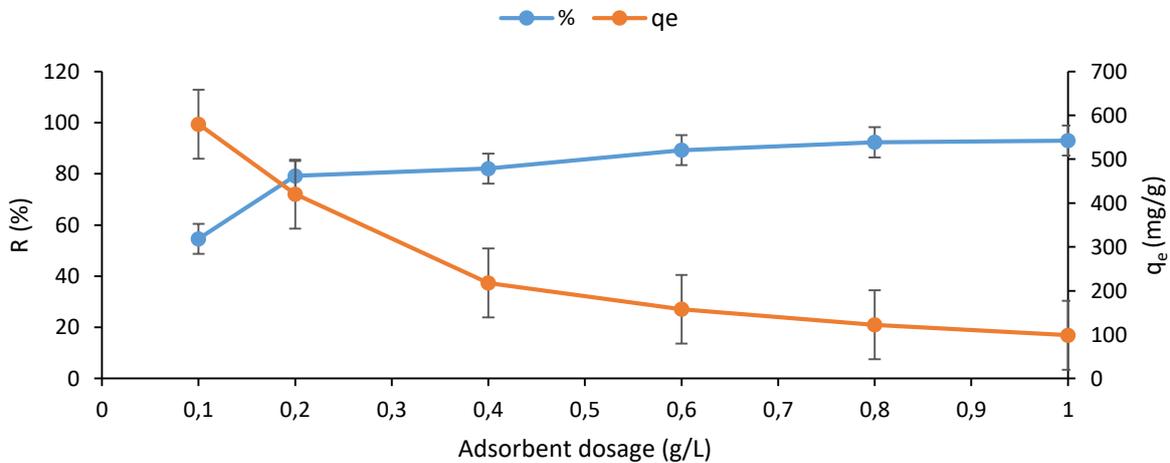


Figure 5. Effect of adsorbent amount on adsorption capacity and removal efficiency.

Effect of initial dye concentration

To investigate the effect of initial AV 17 dye concentration on adsorption, the five different dye concentrations (100, 150, 200, 250, and 300 mg/L) with 0.8 g/L adsorbent dosage were studied (Figure 6). After the study, it was observed that MWCNT treatment efficiency decreased with increasing dye concentration. At 100 mg/L AV 17 concentration, 91.64% dye removal efficiency was obtained, while the efficiency was 73.3% at 300 mg/L dye concentration. It is seen that the decrease in yield is faster after 250 mg/L AV 17 concentration. As the dye concentration increases, the treatment efficiency decreases as the pores and surface saturation of the adsorbent increase [14]. However, adsorption capacity increased with increasing dye concentration. The MWCNT adsorbent's adsorption capacity rose from 121.62 mg/g to 270.36 mg/g when the AV 17 dye concentration was raised from 100 mg/L to 250 mg/L.

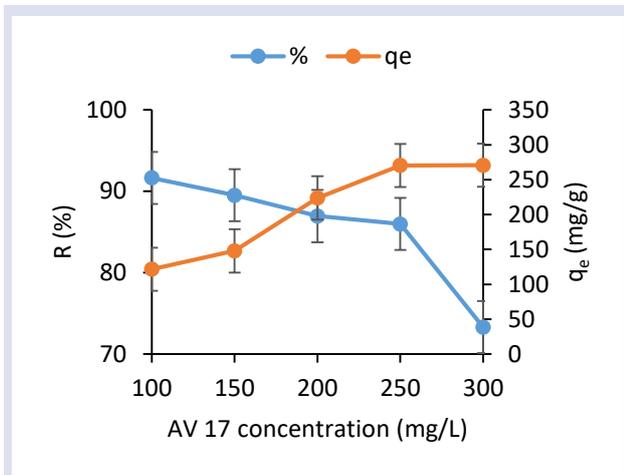


Figure 6. Effect of AV 17 dye concentration on adsorption capacity and removal efficiency.

Effect of contact time

The effect of different contact times on adsorption with MWCNT at 100 mg/L AV 17 concentration under the experimental conditions given in Table 1 was investigated and the results are shown in Figure 7. As can be seen from Figure 7, the treatment started in the first minutes of the contact time and a rapid treatment was realized for the first 120 minutes. At the end of 180 minutes, it was observed that the adsorption process reached

equilibrium. The removal efficiency of MWCNT adsorbent was 90% and the adsorption capacity was 119.45 mg/g at the assumed equilibrium moment. The adsorption rate, which occurs in the first minutes and then stops or slows down, shows that there is plenty of free space on the MWCNT surface in the first place, but as the contact time increases, the active sites required for dye adsorption are filled with dye molecules and decreased. A similar situation was observed in the sorption removal of cationic dyes from aquatic solution using magnetic MWCNT by Song, Shi, et al., [15].

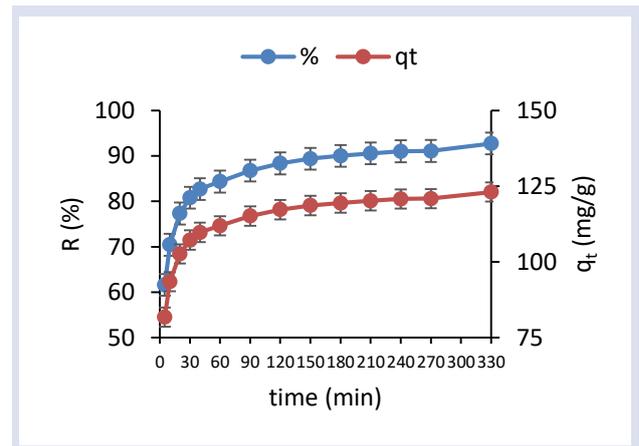


Figure 7. Impact of contact time on removal efficiency and adsorption capacity.

Effect of initial pH

The pH of the solution is a crucial factor in adsorption since it has a significant impact on the degree of ionization, the structure of the dye molecules, and the surface charge of the adsorbents [16]. Therefore, to determine the effect of solution pH on the adsorption capacity of MWCNT used as an adsorbent, the pH value was varied between 2-12 and its effect on AV 17 removal was monitored (Figure 8). The results in Figure 8 show that the amount of AV 17 adsorption by MWCNT did not significantly increase with increasing pH. While qe was 122.53 mg/g in the solution adjusted to pH 6, this value was calculated as 126.32 mg/g at pH 12. The treatment efficiency increased from 92.32% to 95.18% when the pH increased from 6 to 12. Therefore, it was decided that it was appropriate to work at the natural pH value of the dye solution without pH adjustment.

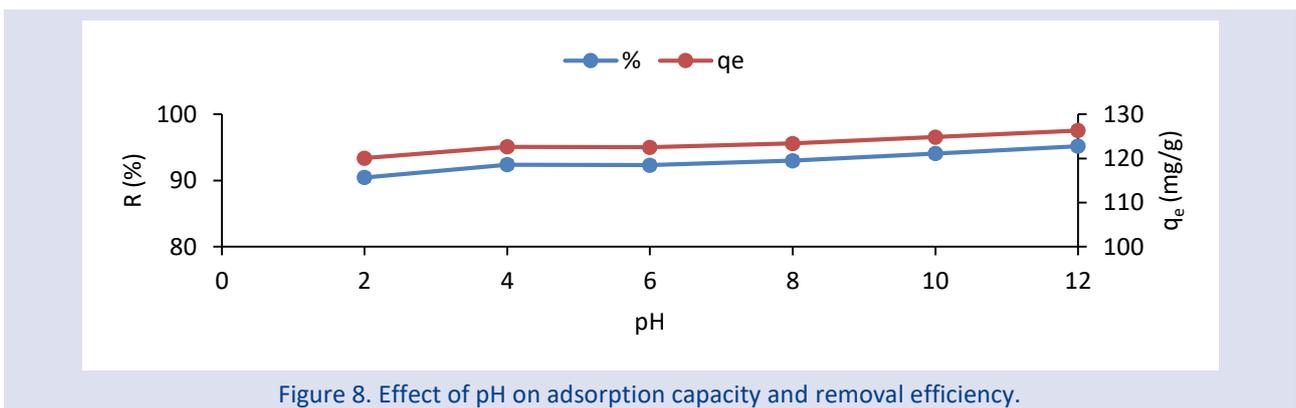


Figure 8. Effect of pH on adsorption capacity and removal efficiency.

Effect of temperature

Temperature is one of the parameters that significantly affect the sorption capacity of sorbents. The effect of temperature on AV 17 adsorption was studied at four different temperature values (25, 35, 45, and 55°C) (Figure 9). As seen in Figure 9, the adsorption efficiency and capacity increased with increasing temperature. Because high temperatures can expand the pores and/or activate the adsorbent surface [17].

Considering the cost increase that will occur if the process water temperature is increased, the process can be operated without the need for heating, especially at low concentrations, but it is more appropriate to heat the process water to the range of 35-45°C at higher dye concentrations for treatment efficiency. The increase in efficiency with increasing temperature indicates the presence of an endothermic reaction between AV 17 and MWCNT.

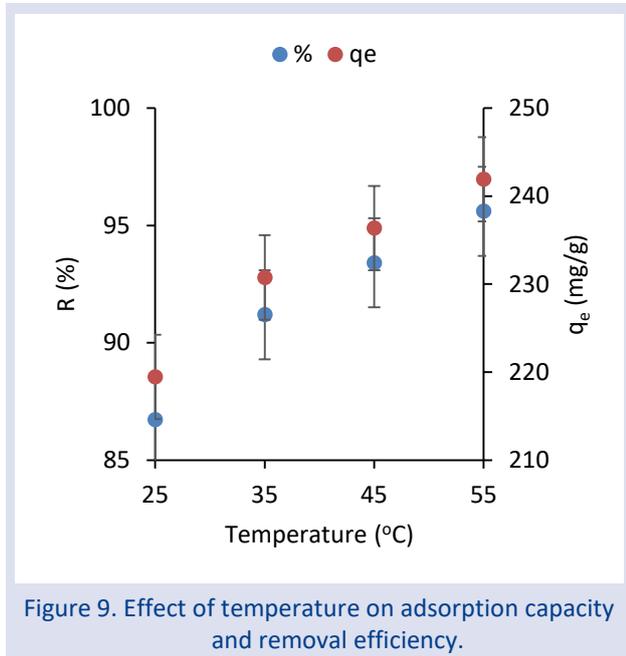


Figure 9. Effect of temperature on adsorption capacity and removal efficiency.

Adsorption kinetics

Adsorption kinetic models were used to investigate their interaction and to evaluate the mass transfer of AV 17 dye on MWCNT. To determine the contact time required to complete the adsorption process, the time range of 0-330 min was studied. Adsorption kinetic models provide important information needed to model the adsorption process, predict its rate, and design it. The adsorption kinetics of AV 17 dye with MWCNT adsorbent were investigated according to the pseudo-second-order kinetics (Figure 10a), intraparticle diffusion (Figure 10b), and Elovich (Figure 10c) models. The kinetic parameters were calculated by substituting the data obtained from Figure 7 into the equations given in Table 2. As seen in Table 2, a higher R² value (0.9998) was obtained for the pseudo-second-order reaction kinetics model. It is also seen that the calculated theoretical q_e value (q_{e, cal}) is close to the experimental q_e value (q_{e, exp}).

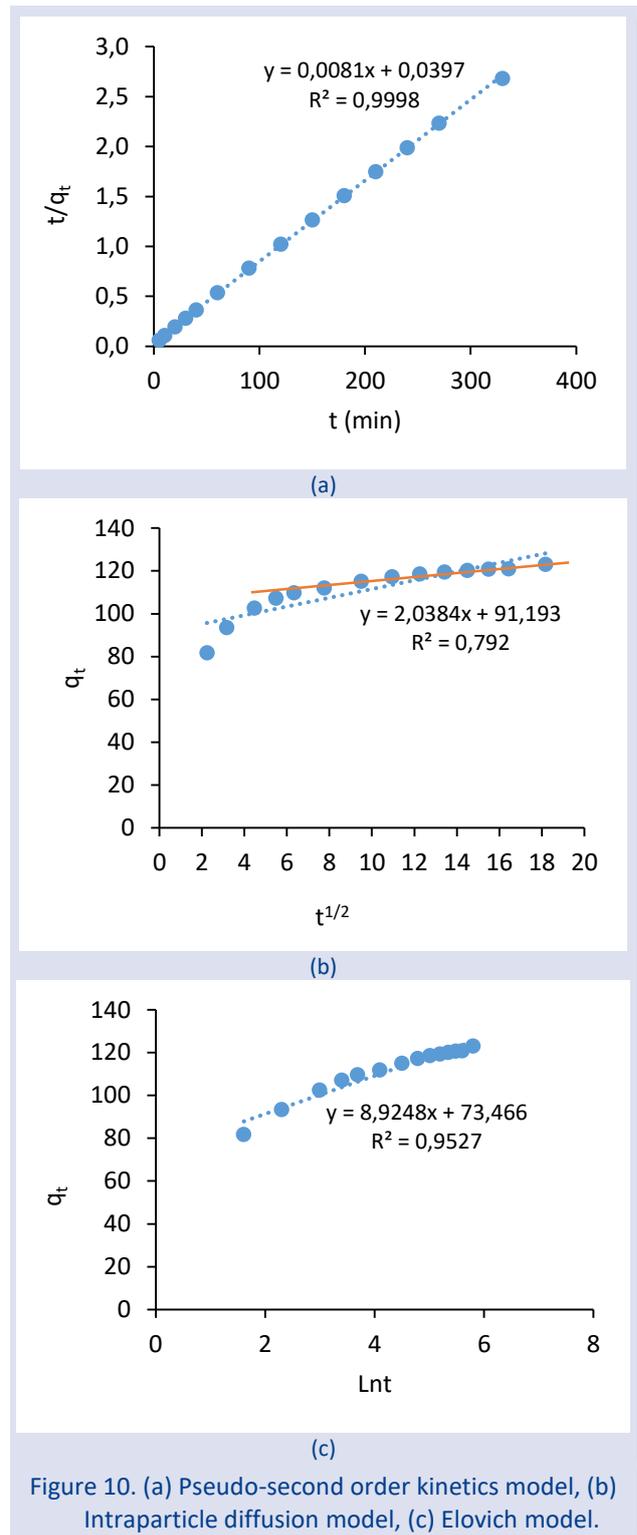


Figure 10. (a) Pseudo-second order kinetics model, (b) Intraparticle diffusion model, (c) Elovich model.

To investigate the stepwise adsorption process and analyze the diffusion mechanism, the intraparticle diffusion model was applied and the obtained graph is given in Figure 10(b). From Figure 10(b), it can be seen that the straight line between q_t and t^{1/2} is nonlinear from the origin and the AV 17 adsorption of MWCNT adsorbent consists of a two-step process. The first stage's quick rise is an instantaneous diffusion of the adsorption stage at the outermost surface. The intraparticle diffusion mechanism regulates the rate during the second step, which is the progressive adsorption stage [21].

Table 2. Kinetic parameters for the adsorption of AV 17 on MWCNT

| Kinetic Models | Equations | Reference |
|-------------------------------|--|--|
| Pseudo-second-order model | $t/q_t = (1/k_2 \cdot q_e^2) + (t/q_e)$ (3) | |
| | q_e (mg/g) | $q_{e, cal} = 123.45$ $q_{e, exp} = 119.45$ |
| | k_2 (mg/g.dk) | $1.652 \cdot 10^{-3}$ |
| | R^2 | 0.9998 |
| Intraparticle diffusion model | $q_t = (k_d \cdot t^{1/2}) + C$ (4) | |
| | k_d (mg/(g dk ^{1/2})) | 2.0384 |
| | C (mg/g) | 91.193 |
| | R^2 | 0.7920 |
| Elovich model | $q_t = (1/\beta) \cdot (\ln(\alpha\beta)) + (1/\beta) \cdot (\ln t)$ (5) | |
| | α (mg/g dk) | 33554.47 |
| | β (g/mg) | 0.112 |
| | R^2 | 0.9527 |

As a result, both phases work together to control the adsorption rate; intraparticle diffusion is not the only step that limits the rate. The slope of the second step is smaller than the slope of the first. This could be connected to the reduction in active sites on MWCNT that occurs as dye molecules diffuse over longer times [15].

Adsorption isotherms

The relationship between the amount of pollutant adsorbed on the adsorbent and the concentration of the pollutant in the solution at equilibrium can be explained using adsorption isotherms. Adsorption isotherm analysis is a vital tool for figuring out an adsorbent's maximal adsorption capacity. Isotherm models are used to determine the adsorption capacity as well as the physical or chemical nature of the adsorption process. In this study, Langmuir, Freundlich, and Temkin isotherm models were used to assess the adsorption data and explain the adsorption mechanism.

The Langmuir isotherm assumes that once a pore on the adsorbent surface is filled, no further sorption will occur there. The Langmuir model is a theoretical model of monolayer chemical adsorption [22]. To discover the equilibrium distribution between solid and liquid, employ the Langmuir adsorption isotherm. The Langmuir isotherm equation is given in Equation 6 and the graph formed from this equation is represented in Figure 11(a).

$$q_e = \frac{q_m \times K_L \times C_e}{1 + K_L \times C_e} \tag{6}$$

Where,

q_e : Amount of pollutant adsorbed per unit weight of adsorbent at equilibrium (mg/g),

K_L : Constant connected to the free energy of adsorption (L/mg),

q_m : Maximum capacity for adsorption (mg/g),

C_e : Concentration of pollutants in solution at the equilibrium time (mg/L).

According to the Freundlich isotherm, adsorption processes take place on heterogeneous surfaces. The concentration of the pollutant adsorbed on the adsorbent surface increases as the concentration of the pollutant in the solution increases. Equation 7 provides the Freundlich isotherm equation, and Figure 11(b) shows the graph that results from this equation.

$$q_e = K_F \cdot C_e^{\frac{1}{n}} \tag{7}$$

Where,

K_F : Constant showing the adsorbent's adsorption capability (mg^{1-(1/n)}L^{1/n}g⁻¹),

n : A constant that represents the intensity of adsorption,

C_e : Concentration of the substance left in solution following adsorption (mg/L).

The heat of adsorption is inversely proportional to the increasing adsorbent surface, and the energy distribution is uniform up to the maximal binding energy of adsorption, according to the Temkin isotherm model [21]. This isotherm considers the interaction between adsorbent and adsorbate and ignores extremely high and low concentration values. This model states that the heat of adsorption (ΔH , a function of temperature) of all molecules in the layer drops linearly, not logarithmically, as the surface area of the adsorbent grows. This adsorption isotherm model covers only the intermediate concentration range [23]. Equation 8 provides the Temkin isotherm equation, and Figure 11(c) shows the graph that results from this equation.

$$q_e = B \ln(A C_e) \tag{8}$$

Where,

q_e : Quantity of adsorbate adsorbed at equilibrium (mg/g),

C_e : Concentration of solution at equilibrium (mg/L),
 B: Constant connected to the adsorption heat ($B=RT/b$)

b: Temkin constant connected to the adsorption heat (J/mol), T: Absolute temperature (K), R: Gas constant (8.314 J/mol K),

A: Temkin isotherm constant (L/g).

Table 3. lists the model constants and correlation coefficients that were determined using isotherm plots. The adsorption fits the Langmuir isotherm model ($R^2 = 0.9813$) when the correlation coefficient values listed in Table 3 are analyzed. This demonstrates that the monolayer adsorption mechanism is the mode of adsorption.

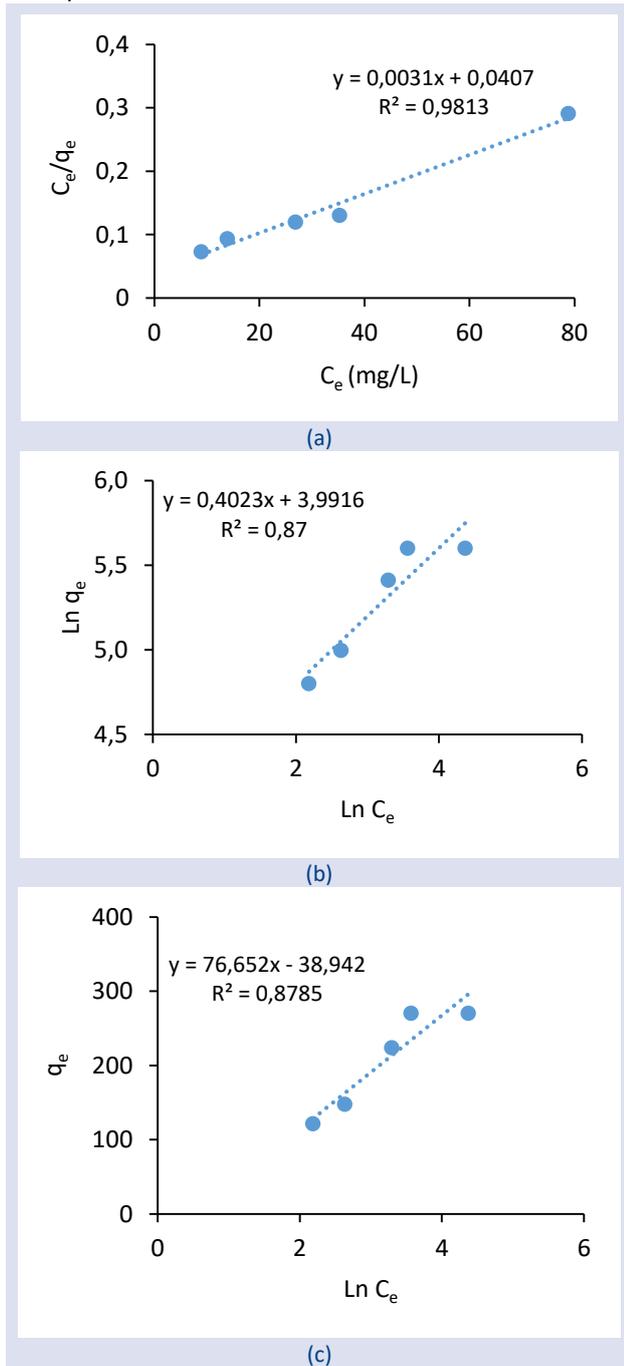


Figure 11. (a) Langmuir, (b) Freundlich, (c) Temkin isotherm models.

$$R_L = 1/(1+bC_0) \tag{9}$$

The varied distribution of dye molecules on the adsorption sites could be the cause of this. Equation 9 provides the calculation of R_L , a dimensionless constant derived from the Langmuir model. $0 < R_L < 1$: Favorable and spontaneous adsorption, $R_L > 1$: Not favorable adsorption, $R_L = 1$: The linear process, $R_L = 0$: Irreversible adsorption [24].

Where C_0 is the initial dye concentration.

Since the R_L value was obtained in the range of 0 to 1 for all dye concentrations (0.116 for 100 mg/L AV 17; 0.042 for 300 mg/L AV 17), it can be concluded that the adsorption process is effective in removing AV 17 dye and a strong bond is formed between MWCNT and dye. Table 3 displays the coefficients of adsorption intensity (n) and adsorption capacity (K_F) that were determined using the Freundlich model.

The n should take values ranging from 1 to 10 for a preferable adsorption process. Furthermore, $n > 1$ denotes the formation of multiple layers on the adsorbent surface. The $1/n$ represents the surface heterogeneity factor. Surface heterogeneity rises as $1/n$ gets closer to zero [25]. When we look at Table 3, the n value determined from the Freundlich model is 2.485 and the $1/n$ value is 0.4. This indicates that AV 17 adsorption on the adsorbent is preferred and the MWCNT adsorbent surface is heterogeneous.

Table 3. Isotherm parameters for the adsorption of AV 17 on MWCNT

| Langmuir | q_m (mg/g) | b (L/mg) | R^2 |
|------------|--------------|-------------------------------------|--------|
| | 322,58 | 0,076 | 0,9813 |
| Freundlich | K_F (mg/g) | n ((mg/g)·(L/mg) ^{1/n}) | R^2 |
| | 54,14 | 2,485 | 0,87 |
| Temkin | A (L/g) | B (J/mol) | R^2 |
| | 0,6 | 76,652 | 0,8785 |

Furthermore, the n value is employed to denote the chemical or physical nature of the process. In chemical processes, $n < 1$ denotes whereas in physical processes, $n \geq 1$ [26]. The physical mechanism was responsible for the realization of AV 17 adsorption on MWCNT, as indicated by the n constant (2.485) obtained for this work.

Table 4. compares the maximal Langmuir adsorption capacities (q_{max} (mg/g)) of MWCNT for the adsorption of different pollutants. Table 4 confirms that this nanocomposite is an effective applicant for AV 17 dye-containing wastewater treatment. When compared with other adsorbate materials to be treated, it is seen that higher efficiency per unit adsorbent is obtained from this study. In their study, Moussavi and Fazli [27] also examined Acid violet 17 dye decolorization by multi-walled carbon nanotubes from aqueous solution. However, it differs from this study in terms of working conditions, research method, and results obtained. The adsorption capacity obtained after 180 minutes of contact time at a concentration of 25 mg/L AV17 at pH 4 and a dose of 0.4 g/L MWCNT is about 4.5 mg/g. This result is considerably lower than the unit adsorption capacity obtained in this study as shown in Table 4.

Table 4. Adsorption capacities determined from the Langmuir isotherm model for different adsorbate materials uptake by MWCNT

| Adsorbate material | pH | Adsorbate conc. (mg/L) | Sorbent dose (g/L) | Time | q _m (mg/g) | References |
|------------------------|-------|------------------------|--------------------|--------|-----------------------|------------|
| Phenol | 4.65 | 50 | 2 | 56 min | 64.60 | [28] |
| Cr(VI) | 5-6 | 100 | 0.15 | 6 h | 24 | [29] |
| Direct Blue 53 | 2 | 300 | 1.5 | 3 h | 334.8 | [30] |
| Congo Red | | 50-400 | 2.5 | | 148.08 | [31] |
| Reactive green HE4BD | | 50-400 | 2.5 | | 151.88 | [31] |
| Methylene blue | 6 | 20 | 1.0 | 1 h | 19.6 | [32] |
| Golden yellow MR | | 50-400 | 2.5 | | 141.618 | [31] |
| Triclosan | 3, 10 | 8 | 0.05 | | 166.83 | [33] |
| Methyl orange | 2 | 20 | 0.3 | 2 h | 50.25 | [34] |
| Phenol | 7 | 25 | 5 | 5 h | 32.25 | [35] |
| Ni(II) | | | | | 6.09 | [35] |
| Maxilon blue dye (GRL) | 6 | 50 | 0.1 | 2 h | 187.69 | [36] |
| Methylene Blue Dye | | 40-100 | | 1 h | 55.18 | [37] |
| MB | 6 | 10 | 0.2g | | 59.7 | [38] |
| AR 183 | | | | | 45.2 | [38] |
| AV 17 | 6.52 | 200 | 0.8 | 3 h | 322.58 | * |

Adsorption thermodynamics

Adsorption thermodynamics was examined to examine the adsorption mechanism of AV 17 on MWCNT. Table 5 presents the results of the computations for three thermodynamic parameters: the change in Gibbs free energy (ΔG°), the change in enthalpy (ΔH°), and the change in entropy (ΔS°). From the slope (ΔH°) and shift (ΔS°) of the $1/T$ plot vs $\ln K_c$ in Figure 12, the values of ΔH° and ΔS° are computed.

$$\Delta G^\circ = -RT \ln K_c \tag{10}$$

$$\Delta G^\circ = \Delta H^\circ - T \Delta S^\circ \tag{11}$$

$$\ln K_c = \Delta S^\circ / R - \Delta H^\circ / RT \tag{12}$$

Where,

R = Universal gas constant; T = Absolute temperature (K)

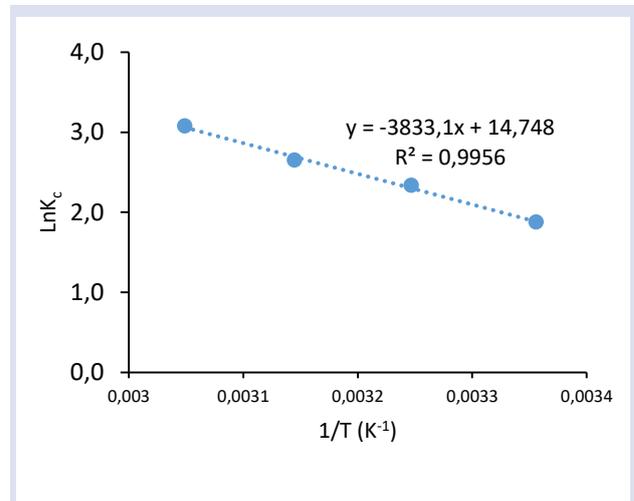


Figure 12. The estimation of thermodynamic parameters.

Table 5. Thermodynamic parameters for adsorption of AV 17 on MWCNT at different temperatures

| | ΔG° (kJ/mol) | | | | ΔS° (kJ/mol.K) | ΔH° (kJ/mol) |
|-----------------------|---------------------------|--------|--------|--------|-----------------------------|---------------------------|
| Temperature (K) | 298 | 308 | 318 | 328 | | |
| C ₀ (mg/L) | | | | | | |
| 200 | -4,651 | -5,986 | -7,010 | -8,399 | 0,123 | 31,870 |

From Table 5., it is seen that the ΔG° value decreases as the temperature increases. Negative ΔG° values indicate that adsorption is spontaneous and thermodynamically favorable. Also, absolute values of ΔG° less than 20 kJ/mol indicate that dye adsorption on MWCNT is mainly physical adsorption. Positive ΔH° values indicate that the reaction is endothermic. In addition, an absolute ΔH° value of less than 84 kJ/mol indicates that the process is physical adsorption [39]. Positive ΔS° values are also related to the randomness of the whole system during the adsorption process and the adsorbent's AV 17 gravity force [40].

Conclusions

Many parameters such as dye concentration, pH, contact time, adsorbent dose and temperature were

found to be effective in the removal of AV 17 using MWCNT. It was determined that the treatment was more effective at natural pH conditions. This is very beneficial at the application scale as there is no additional cost for pH adjustment. The kinetic and isotherm modeling of AV 17 dye adsorption with MWCNT showed agreement with pseudo-second-order kinetics and the Langmuir isotherms. The thermodynamic study presented a favorable and spontaneous adsorption process. MWCNT showed a high adsorption capacity ($q_m = 322.58$ mg/g). This is a very good value compared to other results obtained in the literature for AV 17 dye removal. All the results obtained showed that MWCNT can be used effectively for azo group dye removal.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Saleem M., Pirzada T., Qadeer R., Sorption of Some Azo-Dyes on Wool Fiber from Aqueous Solutions, *Colloids Surf. A: Physicochem. Eng. Asp.*, 260 (1–3) (2005) 183–188.
- [2] Saleem M., Pirzada T., Qadeer R., Sorption of Acid Violet 17 and Direct Red 80 Dyes on Cotton Fiber from Aqueous Solutions, *Colloids Surf. A: Physicochem. Eng. Asp.*, 292 (2–3) (2007) 246–250.
- [3] Jain S.N., Gogate P.R., Adsorptive Removal of Acid Violet 17 Dye from Wastewater Using Biosorbent Obtained from NaOH and H₂SO₄ Activation of Fallen Leaves of *Ficus racemosa*, *J. Mol. Liq.*, 243 (2017) 132–143.
- [4] Kodavatiganti S., Bhat A.P., Gogate P.R., Intensified Degradation of Acid Violet 7 Dye using Ultrasound Combined with Hydrogen Peroxide, Fenton, and Persulfate, *Sep. Purif. Technol.*, 279 (2021) 119673.
- [5] Jain S.N., Shaikh Z., Mane V.S., Vishnoi S., Mawal V.N., Patel O.R., et al., (2019) Nonlinear Regression Approach for Acid Dye Remediation Using Activated Adsorbent: Kinetic, Isotherm, Thermodynamic and Reusability Studies, *Microchem J.*, 148 (2019) 605–615.
- [6] Üstün-Odabaşı S., Utilizing Sustainable Hemp Biomass as an Eco-Friendly for Potentially Toxic Elements Removal from Water, *Mater. Res. Express*, 11 (2) (2024) 025104.
- [7] Şentürk İ., Alzein M., Adsorption of Acid Violet 17 onto Acid-Activated Pistachio Shell: Isotherm, Kinetic and Thermodynamic Studies, *Acta Chim. Slov.*, 67(1) (2020) 55–69.
- [8] Jain S.N., Gogate P.R., Adsorptive Removal of Acid Violet 17 Dye from Wastewater using Biosorbent Obtained from NaOH and H₂SO₄ Activation of Fallen Leaves of *Ficus Racemosa*, *J. Mol. Liq.*, 243 (2017) 132–143.
- [9] Thinakaran N., Baskaralingam P., Pulikesi M., Panneerselvam P., Sivanesan S., Removal of Acid Violet 17 from Aqueous Solutions by Adsorption onto Activated Carbon Prepared from Sunflower Seed Hull, *J. Hazard. Mater.*, 151 (2) (2008) 316–322.
- [10] Gupta V.K., Agarwal S., Bharti A.K., Sadegh H., Adsorption Mechanism of Functionalized Multi-Walled Carbon Nanotubes for Advanced Cu (II) Removal, *J. Mol. Liq.*, 230 (2017) 667–673.
- [11] Al-Johani H., Salam M.A., Kinetics and Thermodynamic Study of Aniline Adsorption by Multi-Walled Carbon Nanotubes from Aqueous Solution, *J. Colloid Interface Sci.*, 360 (2), (2011) 760–767.
- [12] Zhou W., Zhou X., Song W., Wang C., Zhang H., Huang X., Removal of Benzohydroxamic Acid from Aqueous Solutions using Multi-Walled Carbon Nanotubes/Iron-Doped Hydroxyapatite Composites: Synthesis, Adsorption Performance, and Characteristics, *J. Environ. Chem. Eng.*, 11 (6) (2023) 111259.
- [13] Topal Canbaz G., Fe₃O₄@Granite: A Novel Magnetic Adsorbent for Dye Adsorption, *Processes*, 11 (9) (2023) 2681.
- [14] Egbosuba T.C., Egwunyenga M.C., Tijani J.O., Mustapha S., Abdulkareem A.S., Kovo A.S., et al., Activated Multi-Walled Carbon Nanotubes Decorated with Zero Valent Nickel Nanoparticles for Arsenic, Cadmium and Lead Adsorption from Wastewater in a Batch and Continuous Flow Modes, *J. Hazard. Mater.*, 423 (2022) 126993.
- [15] Song G., Shi Y., Wang H., Li A., Li W., Sun Y., et al., Effective Sorptive Removal of Five Cationic Dyes from Aqueous Solutions by Using Magnetic Multi-Walled Carbon Nanotubes, *Water Sci. Technol.*, 85 (7) (2022) 1999–2014.
- [16] Song G., Tong L., Chen S., Zhang J., Zhang Y., Wang H., et al., Facile Fabrication of Amino Functionalized Magnetic Multi-walled Carbon Nanotubes for Removal of Congo Red from Aqueous Solutions, *Water Air Soil Pollut.*, 233 (12) (2022) 534.
- [17] Abutaleb A., Imran M., Zouli N., Khan A.H., Hussain S., Ali M.A., et al., Fe₃O₄-Multiwalled Carbon Nanotubes-Bentonite as Adsorbent for Removal of Methylene Blue from Aqueous Solutions, *Chemosphere*, 316 (2023) 137824.
- [18] Ho Y.S., McKay G., Pseudo-Second Order Model for Sorption Processes, *Process Biochem.*, 34 (5) (1999) 451–465.
- [19] Weber W.J., Morris J.C., Kinetics of Adsorption on Carbon from Solution, *J. Sanit. Eng. Div.*, 89 (2) (1963) 31–60.
- [20] Chien S.H., Application of Elovich Equation to the Kinetics of Phosphate Release and Sorption in Soils, *Soil Sci. Soc. Am. J.*, 44 (2) (1980) 265–268.
- [21] Balarak D., McKay G., Utilization of MWCNTs/Al₂O₃ as Adsorbent for Ciprofloxacin Removal: Equilibrium, Kinetics and Thermodynamic Studies, *J. Environ. Sci. Health A Tox., Part A*, 56 (3) (2021) 324–333.
- [22] Wang J., Guo X., Adsorption Isotherm Models: Classification, Physical Meaning, Application and Solving Method, *Chemosphere*, 258 (2020) 127279.
- [23] Edet U.A., Ifelebuegu A.O., Kinetics, Isotherms, and Thermodynamic Modeling of the Adsorption of Phosphates from Model Wastewater Using Recycled Brick Waste, *Processes*, 8 (6) (2020) 665.
- [24] Chandrasekaran A., Patra C., Narayanasamy S., Subbiah S., Adsorptive Removal of Ciprofloxacin and Amoxicillin from Single and Binary Aqueous Systems using Acid-Activated Carbon from *Prosopis juliflora*, *Environ. Res.*, 188 (2020) 109825.
- [25] Şentürk İ., Yıldız M.R., Highly Efficient Removal from Aqueous Solution by Adsorption of Maxilon Red GRL dye using Activated Pine Sawdust, *Korean J. Chem. Eng.*, 37 (6) (2020) 985–999.
- [26] Şentürk İ., Effective Adsorption of Congo Red by Eco-Friendly Granite-Modified Magnetic Chitosan Nanocomposite (G@Fe₃O₄@CS), *Biomass Conv. Bioref.*, (2023) <https://doi.org/10.1007/s13399-023-04826-1>
- [27] Moussavi S.P., Mohammadian Fazli M., Acid Violet 17 Dye Decolorization by Multi-walled Carbon Nanotubes from Aqueous Solution, *JHEHP*, 1 (2) (2016) 110–117.
- [28] Dehghani M.H., Mostofi M., Alimohammadi M., McKay G., Yetilmezsoy K., Albadarin A.B., et al., High-Performance Removal of Toxic Phenol by Single-Walled and Multi-Walled Carbon Nanotubes: Kinetics, Adsorption, Mechanism and Optimization Studies, *J. Ind. Eng. Chem.*, 35 (2016) 63–74.
- [29] Ahmadpour A., Eftekhari N., Ayati A., Performance of MWCNTs and a Low-Cost Adsorbent for Chromium (VI) Ion Removal, *J. nanostructure chem.*, 4 (2014) 171–178.
- [30] Prola L.D., Machado F.M., Bergmann C.P., de Souza F.E., Gally C.R., Lima E.C., et al., Adsorption of Direct Blue 53 Dye from Aqueous Solutions by Multi-Walled Carbon Nanotubes and Activated Carbon, *J Environ Manage.*, 130 (2013) 166–175.
- [31] Mishra A.K., Arockiadoss T., Ramaprabhu S., Study of Removal of Azo Dye by Functionalized Multi Walled Carbon Nanotubes, *Chem. Eng. J.*, 162 (3) (2010) 1026–1034.

- [32] Ahmed S.H., Rasheed E.A.A., Rasheed, L.A.A., Abdulrahim, F.R. Decolorization of Cationic Dye from Aqueous Solution by Multiwalled Carbon Nanotubes, *J. Ecol. Eng.*, 25 (2) (2024) 72–84.
- [33] Zhou S., Shao Y., Gao N., Deng J., Tan C., Equilibrium, Kinetic, and Thermodynamic Studies on the Adsorption of Triclosan onto Multi-Walled Carbon Nanotubes, *Clean - Soil Air Water*, 41 (6) (2013) 539–547.
- [34] Yao Y., Bing H., Feifei X., Xiaofeng C., Equilibrium and Kinetic Studies of Methyl Orange Adsorption On Multiwalled Carbon Nanotubes, *Chem Eng J.*, 170 (1) (2011) 82–89.
- [35] Abdel-Ghani N.T., El-Chaghaby G.A., Helal F.S., Individual and Competitive Adsorption of Phenol and Nickel onto Multiwalled Carbon Nanotubes, *J. Adv. Res.*, 6 (3) (2015) 405–415.
- [36] Alkaim A.F., Sadik Z., Mahdi D.K., Alshrefi S.M., Al-Sammarraie A.M., Alamgir F.M., et al., Preparation, Structure and Adsorption Properties of Synthesized Multiwall Carbon Nanotubes for Highly Effective Removal of Maxilon Blue Dye, *Korean J. Chem. Eng.*, 32 (2015) 2456–2462.
- [37] Ceroni L., Benazzato S., Pressi S., Calvillo L., Marotta E., Menna E., Enhanced Adsorption of Methylene Blue Dye on Functionalized Multi-Walled Carbon Nanotubes, *Nanomaterials*, 14 (6) (2024) 522.
- [38] Wang S., Ng C.W., Wang W., Li Q., Hao Z., Synergistic and Competitive Adsorption of Organic Dyes on Multiwalled Carbon Nanotubes, *Chem. Eng. J.*, 197 (2012) 34–40.
- [39] Lin R., Liang Z., Yang C., Zhao Z., Cui F., Selective Adsorption of Organic Pigments on Inorganically Modified Mesoporous Biochar and Its Mechanism Based on Molecular Structure, *J. Colloid Interface Sci.*, 573 (2020) 21–30.
- [40] Şentürk İ., Şartlandırılmış Zeolit - Klinoptilolit Minerali ile Bakır Gideriminin Araştırılması, *Karadeniz Fen Bilimleri Dergisi*, 13 (1) (2023) 97–113.

Investigation of Electrochemical Coating of Polypyrrole in The Presence of Sodium Molybdate and Sodium Tungstate

Asuman Ünal^{1,a,*}¹Department of Chemistry, Faculty of Science, Çankırı Karatekin University, Çankırı, Türkiye

*Corresponding author

Research Article

History

Received: 11/06/2024

Accepted: 02/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Polymer composites are considered alternative material because they combine the advantages of all their components. Furthermore, polymer composites can exhibit unique electrochemical and physical properties, enabling their use in applications where metals may not be suitable. Polypyrrole, as a conductive polymer, is a promising candidate for synthesizing new composite materials due to its high electroactivity and ease of processing. Incorporating certain organic and inorganic species into polypyrrole matrices can enhance its electrochemical and physical properties. In the study, polypyrrole films were electrochemically coated in the presence of varying concentrations of sodium molybdate and sodium tungstate. The effects of these metal oxides on the growth of polypyrrole were analyzed, and the electrochemical properties of the resulting composite films were examined in monomer-free solutions. The findings indicate that molybdate and tungstate reduce film's electroactivity, alter redox peaks, and lead to the formation of a new electrochemical film. The maximum anodic peak charge density for polypyrrole, polypyrrole/WO₃ and polypyrrole/MoO₄ films was 12.4 mC cm⁻², 3.57 mC cm⁻² and 3.14 mC cm⁻², respectively. Additionally, the results demonstrated that an optimal amount of sodium tungstate enhances charge transfer while maintaining reversibility in redox reactions.

Keywords: Polypyrrole, Sodium tungstate, Sodium molybdate, Electrochemical deposition.✉ asumanunal@karatekin.edu.tr  <https://orcid.org/0000-0002-8850-7150>

Introduction

Conductive polymers are interesting materials because, unlike metals, their conductivity and insulating properties can be reversibly modified. In other words, while their neutral state is insulating, they can easily be converted into conductive materials through electrochemical or chemical doping [1]. Due to these unique properties, conductive polymers are widely used in energy storage devices [2–4], corrosion protection [5,6], and various sensors, gas, chemical, and biosensors [1,7]. Additionally, conductive polymers exhibit color changes in response to changes in conductivity, making them suitable for use in electrochromic devices [1,8,9]. Polyaniline [10–12], polypyrrole [13,14] and polythiophene [15,16] are the most common conductive polymers [1]. Polypyrrole stands out for its high electroactivity, strong environmental stability, and facile synthesis, which does not require acidic or basic media [17]. It is synthesized from pyrrole monomer in an appropriate electrolyte solution through chemical or electrochemical methods [18]. The electrochemical method is often preferred for producing polymer and polymer composite films, as it allows for precise control of film thickness and electrochemical properties [18]. Polypyrrole has been utilized various fields, including controlled drug release systems [19], corrosion prevention [13] and optical devices [20,21].

In recent decades, the fabrication of polymer composite films has gained significant attention,

especially when incorporating certain types of species, such as organic monomers, carbon nanotubes, graphene, or inorganic species. These species can improve the chemical and physical properties of the polymer films [22]. For example, studies have shown that chemically copolymerizing pyrrole and aniline results in a crystal structure distinct from their individual homopolymers and demonstrates selective ammonia-sensing capabilities [23,24]. Additionally, Khan *et al.* reported that a polyaniline/NiO composite exhibits improved conductivity and thermal stability compared to pure polyaniline [22]. In other words, polymer composites offer a great opportunity to produce new polymer films with enhanced electrochemical properties.

Transition metal oxides like molybdate and tungstate are promising candidates for synthesizing polymer composite materials due to their high specific energy capacity, low operating voltage, multiple reversible redox states, abundant availability, environmental friendliness, and strong physical chemical stability [16,25–29]. In the present study, the electrochemical deposition of polypyrrole was investigated in the presence of various concentrations of sodium tungstate and sodium molybdate. The resulting polymer composite films were analyzed for their electrochemical properties.

Experimental Work

Chemicals and Reagents

Pyrrole was purchased from Sigma-Aldrich and used without distillation. Other reagents, including potassium nitrate (KNO_3) from Sigma Aldrich, and sodium tungstate ($\text{NaWO}_4 \cdot 2\text{H}_2\text{O}$, 99.5%) and sodium molybdate ($\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$) from Carlo Erba, were also purchased. All solutions were prepared with deionized water.

Instrumentation

Cyclic voltammetry (CV) experiments were conducted using a Gamry Instruments Interface 1000E. A glassy carbon electrode (0.07 cm^2) served as the working electrode, a platinum flag electrode (0.06 cm^2) as the counter electrode and an Ag/AgCl electrode was used as the reference electrode.

Preparation and Characterization of Polymer Composite Films

Polypyrrole and its composites were synthesized by cyclic voltammetry using pyrrole monomer in 0.1 M KNO_3 . Initially, a pure polypyrrole film was synthesized from a solution containing 0.1 M pyrrole and 0.1 M KNO_3 at a scan rate of 10 mV s^{-1} for 10 cycles. Subsequently, polymer composites were synthesized in the presence of various concentrations of sodium molybdate and sodium tungstate (0.05 M, 0.1 M, 0.2 M, 0.3 M). Electrochemical coating of the polymer and polymer composites was carried out at a scan rate of 10 mV s^{-1} over several cycles. Characterization was performed in 0.1 M KNO_3 monomer-free solutions at a scan rate of 50 mV s^{-1} over multiple cycles.

Results and Discussions

Deposition of Polymer and Polymer Composites

The electrochemical behavior of a polymer-modified electrode depends on the type of conductive polymer the nature of electrode and electrolyte, and the applied voltage during the polymerization [30]. The coating of polypyrrole and polypyrrole composite films was carried out under identical conditions to specifically compare the effects of metal oxides, $\text{NaWO}_3 \cdot 2\text{H}_2\text{O}$ and $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$.

Figure 1a shows the electrochemical coating of polypyrrole formed from 0.1 M pyrrole and 0.1 M KNO_3 at a scan rate of 10 mV s^{-1} over 10 cycles. The increase in current density with each cycles indicates an increase in electroactive sites within the film. Polypyrrole exhibits a broad anodic peak between -0.36 V and 0.48 V , and a broad cathodic peak between -0.74 V and 0.63 V , along with an irreversible peak at 0.8 V , consistent with previous reports [14]. After deposition, an insoluble black polypyrrole film forms on the electrode surface, imparting unique electroactive properties to the modified electrode.

Figure 1b shows CV curve of the polypyrrole film in a monomer-free 0.1 M KNO_3 solution at a scan rate of 50 mV s^{-1} over 10 cycles. The anodic peak appears between

0.45 V and 0.75 V with a peak current density of 6.28 mA cm^{-2} , while the cathodic peak occurs between -0.76 V and 0.76 V with a peak current density of -5 mA cm^{-2} . The anodic and cathodic charge densities of the polymer film were determined to be 12.4 mC cm^{-2} and 10.6 mC cm^{-2} . When oxidized, the polypyrrole film becomes conductive and upon reduction, it returns to an insulating state.

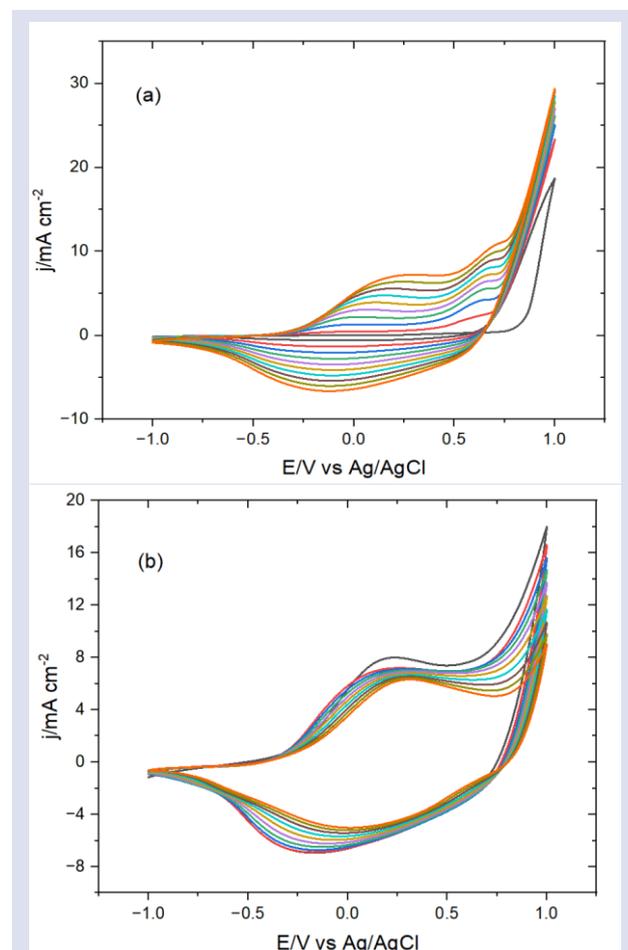


Figure 1. (a) Electrochemical coating of polypyrrole film from 0.1 M pyrrole and 0.1 M KNO_3 (b) CV curve of the polypyrrole film in monomer-free 0.1 M KNO_3 at a scan rate of 50 mV s^{-1}

For the synthesis of the PPy/ WO_3 polymer composite, electrochemical coating of polypyrrole was applied at a scan rate of 10 mV s^{-1} in the presence of varying concentrations of sodium tungstate (0.05 M, 0.1 M, 0.2, 0.3 M), with the resulting CV curves shown in Figure 2. In the pure polypyrrole growth curve, the oxidation and reduction peak current densities are approximately 7.14 mA cm^{-2} , with an irreversible peak current density of 28.6 mA cm^{-2} . For the polypyrrole/ WO_3 composites, the irreversible peak current density was greatly reduced, and the redox peaks almost disappeared, except in the case of polypyrrole/ WO_3 (0.05M). This indicates that higher concentrations of sodium tungstate block the electroactive sites within the polymer matrix. Conversely, a low concentration (0.05 M) does not substantially affect the oxidation and reduction peak positions.

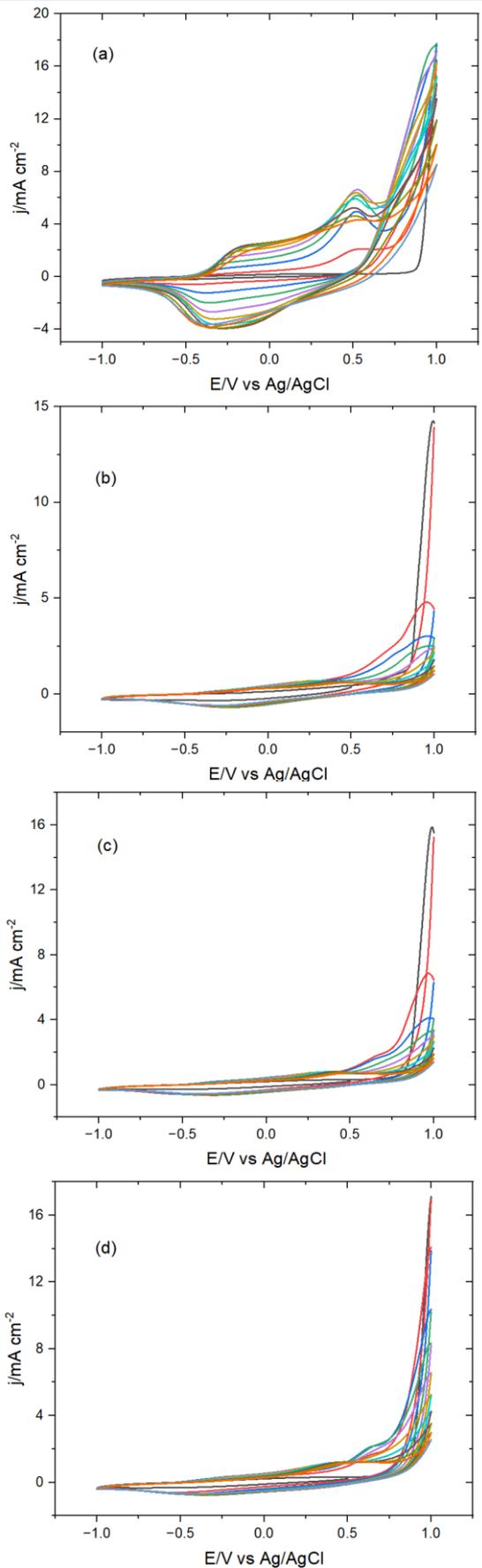


Figure 2. Electrochemical coating of polypyrrole was applied at a scan rate of 10 mV s^{-1} in the presence of various concentrations of sodium tungstate (a) 0.05 M (b) 0.1 M (c) 0.2 M and (d) 0.3 M

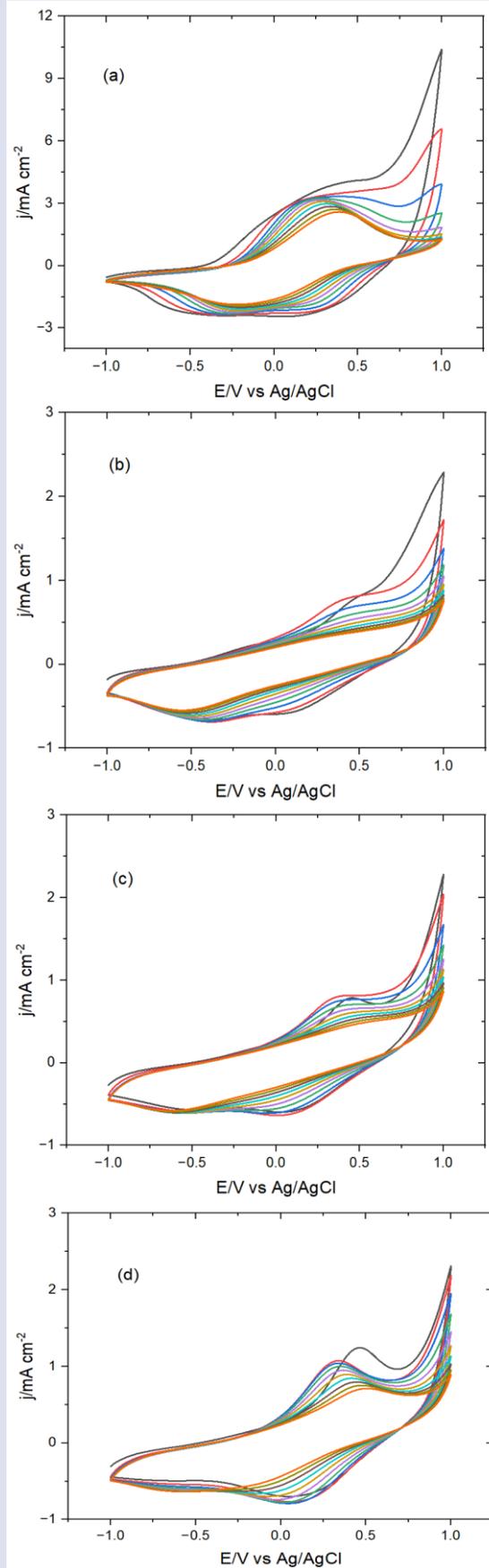


Figure 3. CV curves of polypyrrole/ WO_3 composite films in 0.1 M KNO_3 monomer-free solutions at a scan rate of 50 mV s^{-1} (a) polypyrrole/ $\text{WO}_3(0.05)$ (b) polypyrrole/ $\text{WO}_3(0.1)$ (c) polypyrrole/ $\text{WO}_3(0.2)$ (d) polypyrrole/ $\text{WO}_3(0.3)$

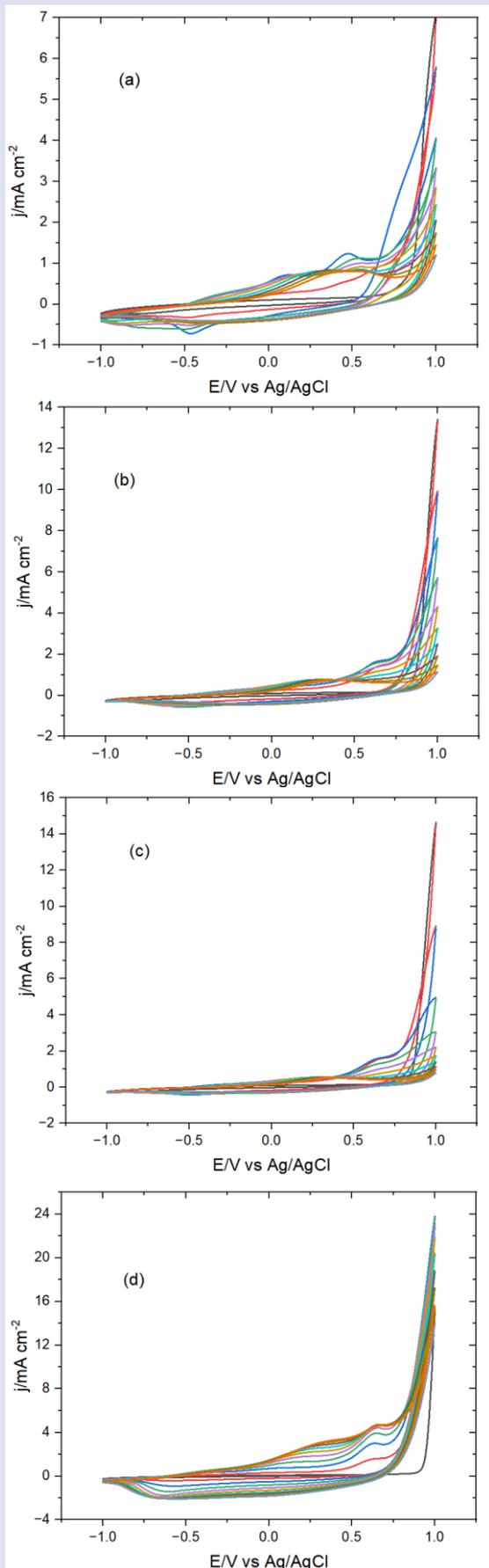


Figure 4. Electrochemical coating of polypyrrole was applied at a scan rate of 10 mV s^{-1} in the presence of various concentrations of sodium molybdate (a) 0.05 M (b) 0.1 M (c) 0.2 M and (d) 0.3 M

The electroactivity of the resulting composite films was characterized in a monomer-free 0.1 M KNO_3 solutions at a scan rate of 50 mV s^{-1} , as illustrated in Figure 3. Figure 3a indicates that the existence of sodium tungstate imparts more reversible behavior to the film compared to pure polypyrrole (see Figure 1b). Adding 0.05 M sodium tungstate does not alter the redox characteristics of the polypyrrole film, but the anodic charge density decreases from 12.4 mC cm^{-2} to 3.57 mC cm^{-2} . While this is a significant reduction, the addition of sodium tungstate introduces reversible behavior. However, with continuous cycling, the film's behavior becomes unstable, suggesting that tungstate ions block the electroactive sites. Increasing the concentration of sodium tungstate in polymer matrix further reduces the electroactivity of the polypyrrole and shifts the anodic peak potential to more positive values. These findings show that tungstate ions block electroactive sites, as evidenced by disappearance of redox peaks after several cycles. Thus, the presence of sodium tungstate makes polypyrrole a recyclable but low-electroactive composite

The electrochemical coating of polypyrrole was also studied in the presence of sodium molybdate, with the growth curves shown in Figure 4. The results are similar to those Figure 2, as both the current density values decrease, and the redox peaks disappear with higher concentrations of sodium molybdate. Interestingly, however, the highest concentration of sodium molybdate (0.3 M) produces broad peaks and an increase in current density values.

The electrochemical properties of the polypyrrole/ MoO_4 composite film were examined in 0.1 M KNO_3 at a scan rate of 50 mV s^{-1} , with the resulting cyclic voltammograms shown in Figure 5. In the presence of higher concentration of sodium molybdate, the oxidation peak of the polypyrrole composite films shifted to more positive potentials, and their electroactivity became unstable and decreased with continued cycling. However, a high concentration of sodium molybdate (0.3 M) provided a more stable and reversible behavior. The anodic peak charge density of this composite film was 3.14 mC cm^{-2} , matching that pure polypyrrole, but with the added advantage of improved reversibility.

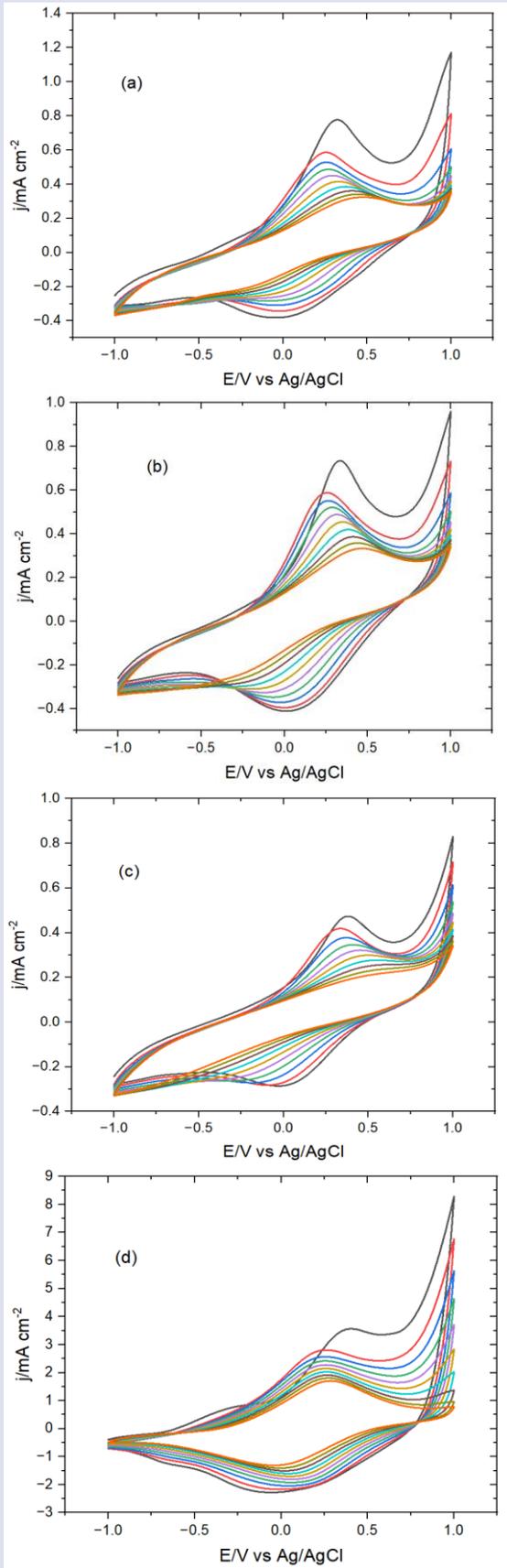


Figure 5. CV curves of polypyrrole/MoO₄ composite films in 0.1 M KNO₃ monomer free solutions at a scan rate of 50 mV s⁻¹ (a) polypyrrole/MoO₄(0.05) (b) polypyrrole/MoO₄(0.1) (c) polypyrrole/MoO₄(0.2) (d) polypyrrole/MoO₄(0.3)

Table 1. shows the charge density values passing through oxidation and reduction redox, and as can be seen, the charge densities values for composite films decrease, but it is also noteworthy that composite films gain reversible properties.

Table 1. Charge density values of polypyrrole and its composites during redox reactions.

| <i>Polymer Composite Film</i> | <i>Anodic Peak Charge Density (mC cm⁻²)</i> | <i>Cathodic Peak Charge Density (mC cm⁻²)</i> |
|-------------------------------------|--|--|
| Polypyrrole | 12.4 | 10.6 |
| Polypyrrole/WO ₃ (0.05) | 3.57 | 3.57 |
| Polypyrrole/WO ₃ (0.1) | 0.86 | 1.00 |
| Polypyrrole/WO ₃ (0.2) | 1.00 | 1.14 |
| Polypyrrole/WO ₃ (0.3) | 1.28 | 1.28 |
| Polypyrrole/MoO ₄ (0.05) | 0.57 | 0.57 |
| Polypyrrole/MoO ₄ (0.1) | 0.57 | 0.57 |
| Polypyrrole/MoO ₄ (0.2) | 0.43 | 0.43 |
| Polypyrrole/MoO ₄ (0.3) | 3.14 | 2.43 |

Conclusion

In this study, various polypyrrole composite films were synthesized in the existence of NaWO₃.2H₂O and Na₂MoO₄.2H₂O. The results indicate that while both sodium tungstate and sodium molybdate provide polypyrrole with reversible redox properties, they influence the polymer matrix differently depending on their concentrations. At a lower concentrations of sodium tungstate (0.05 M), the anodic charge density is 3.57 mC cm⁻², which is lower than that of pure polypyrrole but exhibits excellent redox reversibility. However, higher concentrations of sodium tungstate block the electroactive sites in polypyrrole, leading to decreased current density and a shift in peak potential to more positive values. In contrast, for sodium molybdate, the highest concentration (0.3 M) results in significantly higher anodic current density, while lower concentrations have minimal impact on the polypyrrole film. The anodic charge density of the optimized polypyrrole/MoO₄ composite film is 0.22 mC cm⁻², which is lower than that pure polypyrrole (0.87 mC) but offers improved reversibility in redox reactions. Overall, both composite films allow for the synthesis of polypyrrole films with enhanced redox reversibility, although their electrochemical stability is limited. Further studies could focus on creating electroactive, reversible chemical sensors for various applications by increasing the thickness of polypyrrole through different coating techniques.

Conflict of interest

There is no conflict of interest in this work.

Acknowledgement

We thank Ankara Yıldırım Beyazıt University for all laboratory equipment provided by Yıldız Research Group and Prof. Dr. Abdullah YILDIZ.

References

- [1] Inzelt G., Conducting Polymers: Past, Present, Future, *Journal of Electrochemical Science and Engineering*, 8(1) (2017) 3-37.
- [2] Fan L.Z., Maier J., High-Performance Polypyrrole Electrode Materials for Redox Supercapacitors, *Electrochemistry Communications*, 8(6) (2006) 937-940.
- [3] Muthulakshmi B., Kalpana D., Pitchumani S., Renganathan N.G., Electrochemical Deposition of Polypyrrole for Symmetric Supercapacitors, *Journal of Power Sources*, 158(2) (2006) 1533-1537.
- [4] Mi H., Zhang X., Ye X., Yang S., Preparation and Enhanced Capacitance of Core-Shell Polypyrrole/Polyaniline Composite Electrode for Supercapacitors, *Journal of Power Sources*, 176(1) (2008) 403-409.
- [5] Shinde V., Sainkar S.R., Patil P.P., Corrosion Protective Poly(o-Toluidine) Coatings on Copper, *Corrosion Science*, 47(6) (2005) 1352-1369.
- [6] Gelling V.J., Wiest W.W., Tallman D.E., Bierwagen G.P., Wallace G.G., Electroactive-Conducting Polymers for Corrosion Control, *Journal of Solid-State Electrochemistry*, 6 (2001) 85-100.
- [7] De León P.C., Campbell S.A., Smith J.R., Walsh F.C., Conducting Polymer Coatings in Electrochemical Technology Part 2 – Application Areas, *Transactions of the IMF*, 86(1) (2008) 35-40.
- [8] Mu S., Pronounced Effect of The Ionic Liquid on The Electrochromic Property of The Polyaniline Film: Color Changes in The Wide Wavelength Range, *Electrochimica Acta*, 52(28) (2007) 7827-7834.
- [9] Ohsaka T., Kunimura S., Oyama N., Electrode Kinetics of Poly (o-Aminophenol) Film Prepared by Electro-Oxidative Polymerization of o-Aminophenol and Its Electrochromic Properties, *Electrochimica Acta*, 33(5) (1988) 639-645.
- [10] Syed A.A., Dinesan M.K., Review: Polyaniline-A Novel Polymeric Material, *Talanta*, 38(8) (1991) 815-837.
- [11] Karthikeyan M., Satheeshkumar K.K., Elango K.P., Defluoridation of Water Via Doping of Polyanilines, *Journal of Hazardous Materials*, 163(2-3) (2009) 1026-1032.
- [12] Wei Y., Focke W.W., Wnek G.E., Ray A., MacDiarmid A.G., Synthesis and Electrochemistry of Alkyl Ring-Substituted Polyanilines, *The Journal of Physical Chemistry*, 93(1) (1989) 495-499.
- [13] González M.B., Saidman S.B., Electrodeposition of Polypyrrole On 316l Stainless Steel for Corrosion Prevention, *Corrosion Science*, 53(1) (2011) 276-282.
- [14] Baker C.K., Reynolds J.R., A Quartz Microbalance Study of The Electrosynthesis of Polypyrrole, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 251(2) (1988) 307-322.
- [15] Leclerc M., Faïd K., Electrical and Optical Properties of Processable Polythiophene Derivatives: Structure-Property Relationships, *Advanced Materials*, 9(14) (1997) 1087-1094.
- [16] Szymanska D., Rutkowska I.A., Adamczyk L., Zoladek S., Kulesza P.J., Effective Charge Propagation and Storage in Hybrid Films of Tungsten Oxide and Poly(3,4-Ethylenedioxythiophene), *Journal of Solid-State Electrochemistry*, 14 (2010) 2049-2056.
- [17] Li T., Zhou Y., Dou Z., Ding L., Dong S., Liu N., Qin Z., Composite Nanofibers by Coating Polypyrrole on The Surface of Polyaniline Nanofibers Formed In Presence Of Phenylendiamine As Electrode Materials In Neutral Electrolyte, *Electrochimica Acta*, 243 (2017) 228-238.
- [18] Campbell S.A., Li Y., Breakspear S., Walsh F.C., Smith J.R., Conducting Polymer Coatings in Electrochemical Technology Part 1 – Synthesis and Fundamental Aspects, *Transactions of the IMF*, 85(5) (2007) 237-244.
- [19] Alshammary B., Walsh F.C., Herrasti P., Ponce de Leon C., Electrodeposited Conductive Polymers for Controlled Drug Release: Polypyrrole, *Journal of Solid-State Electrochemistry*, 20 (2016) 839-859.
- [20] Amiri M., Alizadeh N., Highly Photosensitive Near Infrared Photodetector Based on Polypyrrole Nanoparticle Incorporated with CdS Quantum Dots, *Materials Science in Semiconductor Processing*, 111 (2020) 104964.
- [21] Bu C., Tai Q., Liu Y., Guo S., Zhao X., A Transparent and Stable Polypyrrole Counter Electrode for Dye-Sensitized Solar Cell, *Journal of Power Sources*, 221 (2013) 78-83.
- [22] Khan M.A., Ishrat U., Dar A.M., Ahmad A., Structural, Electrical, Optical and Analytical Applications of Newly Synthesized Polyaniline-Based Nickel Molybdate Composite, *Journal of Alloys and Compounds*, 636 (2015) 124-130.
- [23] Hammad A.S., Noby H., Elkady M.F., El-Shazly A.H., In-situ Polymerization of Polyaniline/Polypyrrole Copolymer Using Different Techniques, *IOP Conference Series Materials Science and Engineering*, 290 (2018) 012001.
- [24] Chaudhary V., Kaur A., Enhanced and Selective Ammonia Sensing Behaviour of Poly(aniline-co-pyrrole) Nanospheres Chemically Oxidative Polymerized At Low Temperature, *Journal of Industrial and Engineering Chemistry*, 26 (2015) 143-148.
- [25] Kumar R., Gupta P.K., Agrawal A., Nagarale R.K., Sharma A., Hydrothermally Synthesized Reduced Graphene Oxide-NiWO₄ Nanocomposite For Lithium-Ion Battery Anode, *Journal of The Electrochemical Society*, 164 (2017).
- [26] Wang X.X., Li Y., Liu M.C., Bin Kong L., Fabrication and Electrochemical Investigation of MWO₄ (M = Co, Ni) Nanoparticles As High-Performance Anode Materials for Lithium-Ion Batteries, *Ionics*, 24 (2018) 363-372.
- [27] Park J.S., Cho J.S., Kang Y.C., Scalable Synthesis of NiMoO₄ Microspheres with Numerous Empty Nanovoids as An Advanced Anode Material for Li-Ion Batteries, *Journal of Power Sources*, 379 (2018) 278-287.
- [28] Peng T., Liu C., Hou X., Zhang Z., Wang C., Yan H., Lu Y., Liu X., Luo Y., Control Growth of Mesoporous Nickel Tungstate Nanofiber and Its Application as Anode Material for Lithium-Ion Batteries, *Electrochimica Acta*, 224 (2017) 460-467.
- [29] Balaraju J.N., Raman N., Manikandanath N.T., Nanocrystalline Electroless Nickel Poly-Alloy Deposition: Incorporation of W And Mo, *Transactions of the Institute of Metal Finishing*, 92(3) (2014) 169-176.
- [30] Unal, A.R. Hillman, K.S. Ryder, S. Cihangir, Highly Efficient Defluoridation of Water Through Reusable Poly(aniline-co-o-Aminophenol) Copolymer Modified Electrode Using Electrochemical Quartz Crystal Microbalance, *Journal of The Electrochemical Society*, 168(2) (2021) 022502.

The Effects of a Continuous Positive Airway Pressure (CPAP) Therapy on Oxidative Stress in Patients with Obstructive Sleep Apnea

Mustafa Murat Ozturk ^{1,a,*}, Muhammet Fevzi POLAT ^{2,b}, Yavuz Selim Intepe ^{3,c}, Bulent Ciftci ^{3,d}

¹ Medical Services and Techniques Programs, Vocational School of Health Services, Kırşehir Ahi Evran University, Kırşehir, Türkiye

² Department of Medical Biochemistry, Faculty of Medicine, Yozgat Bozok University, Yozgat, Türkiye

³ Department of Chest Diseases, Faculty of Medicine, Yozgat Bozok University, Yozgat, Türkiye

*Corresponding author

Research Article

History

Received: 25/07/2024

Accepted: 08/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Obstructive sleep apnea (OSA) is a sleep disorder that is prevalent and can have major health consequences. The primary objective of this study was to determine whether oxidative stress is induced in OSA patients. It was further aimed to assess effectiveness of a continuous positive airway pressure (CPAP) therapy on decreasing total oxidant status (TOS) and total antioxidant status (TAS) in OSA patients. This study included 70 patients with sleep disorders. Considering the apnea-hypopnea index (AHI) score confirmed by polysomnography, the subjects were divided into two groups as OSA patients (n = 35) and control group including patients with simple snoring problem (n = 35). OSA patients received a CPAP therapy for one month. Blood samples were collected from both groups of patients to measure TAS and TOS levels before and after the CPAP therapy. TAS and TOS levels were significantly higher (p < 0.001), whereas minimum oxygen saturation (minSpO₂) and mean oxygen saturation (meanSpO₂) were significantly lower in the OSA patients in comparison to those in the control group (P < 0.001). Positive correlations were observed between AHI and TAS levels as well as between AHI and TOS levels. Moreover, positive correlations were found between TAS and average stress index (OSI), TOS and OSI, as well as between minSpO₂ and meanSpO₂. In contrast, negative correlations were observed between AHI and minSpO₂ and meanSpO₂, and between TAS and OSI, minSpO₂ and meanSpO₂. One month of CPAP therapy in OSA patients caused a decrease in TAS and TOS levels, and an increase in OSI, minSpO₂ and meanSpO₂ values (p < 0.001). According to the current study, OSA patients have elevated oxidative stress. One month of CPAP therapy seems to have a positive impact on the antioxidant status remarkably, and led to improvement in oxidative stress.

Keywords: CPAP therapy, Obstructive sleep apnea, Oxidative stress, TAS, TOS.

^a mustafa.ozturk@ahievran.edu.tr

^b <https://orcid.org/0000-0002-2302-0018>

^c dyavuz76@yahoo.com

^d <https://orcid.org/0000-0002-5697-5291>

^e polatfevzi@hotmail.com

^f <https://orcid.org/0000-0002-9818-8763>

^g bulentciftci@gmail.com

^h <https://orcid.org/0000-0003-2703-3817>

Introduction

It is well known that obstructive sleep apnea (OSA) is one of the most common sleep disorders which may have serious health consequences. According to the literature data, 9% to 38% of the adult population has OSA with an apnea hypopnea index (AHI) \geq 5 events/h. OSA is further reported to be more common in men and older age groups [1]. The high incidence of OSA in the population raises serious concerns whether OSA is a public health emergency problem [2]. Thus, it is important to understand physio-pathological mechanisms underlying OSA related diseases.

One of the most frequent pathological abnormalities documented in patients with OSA is hypoxia and hypercapnia. Both conditions are induced by concurrent upper airway collapse occurring as episodes in sleep. Increased respiratory effort and increased sympathetic activity are the concomitant adjustments of the body which may increase the risk of circulatory problems such as cardiovascular diseases and also other health problems like obesity, diabetes mellitus, depression and dementia [3]. It was discovered that elevated oxidative stress is associated with the physio-pathological mechanism

behind cardiovascular disorders in OSA patients [4, 5]. High levels of reactive oxygen species (ROS) and an imbalanced mechanism between ROS and antioxidant systems are thought to contribute to microvascular damage in cases of oxidative stress [6]. Moreover, although the pathophysiological mechanisms of OSA have not been completely defined yet, ischemia-reperfusion due to hypoxia episodes cause increased level of ROS and oxidative stress. Those changes may further damage the vascular endothelium [7, 8]. Therefore, treating OSA is crucial to lower oxidative stress and the comorbidities that go along with it. The common clinical practice used in treating OSA is Continuous Positive Airway Pressure (CPAP) therapy [9].

Previous studies report contradictory findings regarding the relationship between oxidative stress and OSA. While the study by Svatikova et al. suggests that increased oxidative stress is not evident in otherwise healthy OSA patients [10], the study by Lavie et al. indicates that oxidative stress is an important underlying cause of cardiovascular diseases in OSA patients [11]. Additionally, there is lack of scientific evidence

regarding the changes related to oxidative stress in OSA patients following CPAP therapy. The first aim of this study therefore is to determine total oxidant status (TOS) and total antioxidant status (TAS) levels in OSA patients and compare them with control subjects. It was further aimed to clarify the relationship between oxidative stress and sleep parameters. The second aim was to evaluate into potential changes in TOS and TAS levels following a month of CPAP therapy.

Materials and Methods

Subjects

The Ethical Committee gave its approval to the study's protocol (2017-KAEK-189_2019.09.25_23). The study's subjects included a total of 70 patients who were checked out at the sleep center for OSA. Before the study, each patient provided written informed consent. The subjects had no known comorbidities, and did not use cigarettes or alcohol. Patients were suspected of having OSA if they had one of the symptoms, which include snoring, apnea, excessive daytime sleepiness, or choking during sleep.

Polysomnography

All individuals underwent an overnight polysomnographic examination using a Philips Respironics Alice 6 Sleep Diagnostic System (Germany). The polysomnography consisted of continuous polygraphic recordings of electroencephalography, electrooculography, electrocardiography, and electromyography. Nasal and oral airflow and tracheal sounds were recorded using a microphone. Thoracic and abdominal respiration were measured using an impedance belt. Transcutaneous oxygen saturation of each subject was monitored continuously with a pulse oximeter. Video recordings of the subjects were performed for full-night and positional changes in sleep cycles were recorded. When the test subject awoke in the morning, all recordings associated with the test protocol were stopped. A computerized polysomnographic system and a manual scoring process were used in order to collect the data. Scores for respiratory events were determined using AASM (American Academy of Sleep Medicine) guidelines from 2007 [12]. Accordingly, airflow stopping for at least 10 seconds was used to define apnea. A 30% or greater decrease in airflow lasting for at least 10 seconds along with a perceptible reduction of 4% or more oxygen saturation, a 50% or greater decrease in airflow lasting for at least 10 seconds along with a perceptible reduction of 3% or more oxygen saturation, or an electroencephalogram arousal were all taken into consideration to determine hypopnea.

Two groups were established considering the apnea-hypopnea index (AHI) scores of patients. The AHI was scored as follows: AHI \geq 5 events/h was defined as OSA (n = 35) and AHI < 5 events/h recognized as control group (n = 35). One month of CPAP therapy (use < 4 h per night and < 5 days per week) was applied to the patients diagnosed with OSA (therapy group, n = 35).

After an overnight polysomnography, fasting morning blood samples from the OSA and control group participants were taken. Same procedure was repeated for the treatment group who came to the control after 1-month of CPAP therapy for analysis of serum TOS and TAS values as well as for calculation of average stress index (OSI) value. The OSI value, which shows the oxidant/antioxidant balance in the organism, is calculated by dividing TOS values by TAS values [13]. Blood samples were centrifuged at 3000 rpm for 10 minutes in a Biobase brand (China) device, and aliquots of serum were kept tightly closed at -20 °C until analysis.

Analysis of the Total Oxidant-Antioxidant Status

An automated colorimetric test approach for TAS developed by Erel [14] was applied to measure the total antioxidant capacity of blood serum. With hydrogen peroxide, the Fe²⁺-o-dianisidine complex produces an OH radical by a Fenton type reaction. At the reducing low pH, the potent ROS interact with the colorless o-dianisidine molecule to produce yellow-brown dianisidyl radicals. More color is formed as a result of dianisidyl radical participation in additional oxidation processes. Antioxidants present in the samples, however, block these oxidation processes and prevent the development of color. In Biobase Elisa automatic analyzers (China), this reaction is monitored spectrophotometrically. In terms of mmol Trolox Eq/L, the TAS results are presented.

An automated colorimetric testing technique developed by Erel [15] was used to measure plasma TOS levels. The ferrous ion-o-dianisidine complex is converted to ferric ion by the sample's oxidants. This reaction is accelerated by about a threefold factor by the glycerol in the medium. In the acidic media, ferric ions combine with xylenol orange to generate a colorful product. This color is identified spectrophotometrically and is proportional to the amount of oxidant present in the sample. Hydrogen peroxide equivalent per liter ($\mu\text{mol H}_2\text{O}_2$ Eq/L) are used to express the TOS results.

Statistical Analysis

The SPSS for Windows (14.1, SPSS Inc., Chicago, IL, USA) package software was utilized for all statistical analyses. Descriptive statistics for the data were calculated and displayed as 'arithmetic mean \pm standard error of mean'. Before the significance tests, the data were analyzed with the Shapiro-Wilk test in terms of conformity to the normal distribution according to the parametric test assumptions. The difference between the control and OSA groups was examined with the Student t test, and the difference in terms of the means before and after CPAP therapy was examined with the paired-sample t test. The relationship between the variables acquired in the OSA group prior to CPAP therapy was examined using Pearson correlation analysis to identify its strength and direction. If a p value for any statistical analysis was less than 0.05, the results were considered significant.

Results

Subjects were divided into two groups as patients diagnosed with OSA ($n = 35$), and control group with simple snoring problem ($n = 35$). The average age of OSA patients and the control subjects were 47 ± 1.37 years old and 43.31 ± 1.72 , respectively. In both groups, majority of the OSA patients were male ($n = 21$, 60%). The distribution of age and gender did not significantly differ between groups. Compared to the control subjects, TAS and TOS values were significantly greater whereas OSI levels significantly lower in OSA patients ($p < 0.001$). In comparison to the control subjects, minimum oxygen saturation (minSpO_2) and mean oxygen saturation (meanSpO_2) were significantly lower in the OSA patients ($p < 0.001$). Table 1 shows the demographics and the parameters of oxidative stress of both groups.

In the OSA patients, one month of CPAP therapy resulted in lower TAS and TOS levels and higher OSI, minSpO_2 , and meanSpO_2 values ($p < 0.001$, Table 2). When patients who received CPAP therapy for 1 month and control subjects were compared, it was observed that CPAP treatment was quite effective and the values of OSA patients approached the values of control subjects, and no statistical difference could be found between these two groups in terms of the parameters examined (Table 2).

Table 1. Comparison of demographic variables and oxidative stress parameters in control subjects and OSA patients

| Items | Controls (n = 35) | OSA (n = 35) | p-Value |
|---|-------------------|------------------|-----------|
| Age (years) | 43.31 ± 1.72 | 47 ± 1.37 | 0.098 |
| Gender (male: female) | 21: 14 | 21: 14 | 1 |
| AHI (events/h) | 2.77 ± 0.23 | 40.46 ± 3.93 | < 0.001 |
| TAS (mmol Trolox Eq/L) | 0.88 ± 0.02 | 1.34 ± 0.04 | < 0.001 |
| TOS ($\mu\text{mol H}_2\text{O}_2$ Eq/L) | 8.19 ± 0.23 | 9.98 ± 0.22 | < 0.001 |
| OSI | 0.94 ± 0.02 | 0.77 ± 0.03 | < 0.001 |
| minSpO_2 (%) | 88.03 ± 0.35 | 78.97 ± 1.36 | < 0.001 |
| meanSpO_2 (%) | 92.34 ± 0.33 | 87.71 ± 0.6 | < 0.001 |

Table 2. Comparisons of OSA patients after 1-month of CPAP therapy

| Items | Before CPAP | After CPAP | p-Value |
|---|------------------|------------------|-----------|
| TAS (mmol Trolox Eq/L) | 1.34 ± 0.04 | 0.89 ± 0.02 | < 0.001 |
| TOS ($\mu\text{mol H}_2\text{O}_2$ Eq/L) | 9.98 ± 0.02 | 8.6 ± 0.03 | < 0.001 |
| OSI | 0.77 ± 0.03 | 0.97 ± 0.02 | < 0.001 |
| minSpO_2 (%) | 78.97 ± 1.36 | 88.09 ± 0.34 | < 0.001 |
| meanSpO_2 (%) | 87.71 ± 0.6 | 92.43 ± 0.34 | < 0.001 |

Table 3. Comparisons of control subjects and OSA patients after 1-month CPAP therapy

| Items | Controls | OSA after CPAP | p-Value |
|---|------------------|------------------|---------|
| TAS (mmol Trolox Eq/L) | 0.88 ± 0.02 | 0.89 ± 0.02 | 0.656 |
| TOS ($\mu\text{mol H}_2\text{O}_2$ Eq/L) | 8.19 ± 0.23 | 8.6 ± 0.23 | 0.203 |
| OSI | 0.94 ± 0.02 | 0.97 ± 0.02 | 0.195 |
| minSpO_2 (%) | 88.03 ± 0.35 | 88.09 ± 0.34 | 0.908 |
| meanSpO_2 (%) | 92.34 ± 0.33 | 92.43 ± 0.34 | 0.856 |

Discussion

In our study, we compared polysomnographic records and oxidative stress status of OSA patients with those of the control subjects. We also evaluated the changes in the parameters following a CPAP therapy lasted for 1 month in OSA patients. The specific relationships between physiopathogenic mechanisms in OSA remain unclear. The repeated hypoxia and reoxygenation cycles caused by repeated episodes of breathing cessation during sleep may induce the oxidative stress response which leads to physiopathological changes in OSA patients including metabolic alterations, cardiovascular and neural disorders [16].

A limited number of studies report contradictory results on the TAS/TOS levels in OSA and further on the effectiveness of CPAP treatment to decrease the parameters related to oxidative stress. Kang et al. reported that OSA does not change the TAS and TOS levels and further 1-night of CPAP treatment does not improve the oxidative stress [17]. Similarly, some other authors did not find any differences in parameters indicating increased oxidative stress in OSA patients [10, 11, 18-20]. In contrast, enhanced neutrophil superoxide release [21] and low TAS, vitamin A, and E levels [22] were reported in OSA patients compared with control subjects. In this study, impaired protective systems for oxidative stress were evident in OSA patients. Furthermore, in parallel to findings of the previous study, the higher values of TAS and TOS, and the lower values of OSI noted in OSA group, which suggests that the balance between oxidant and antioxidant levels was negatively affected. Changes in TAS, TOS and OSI show that oxidative stress is evident in OSA patients. These conflicting findings could be caused by a number of variables, including the presence of comorbidities and the use of drugs by OSA patients. Moreover, most studies lack data from control subjects on different parameters (e.g., body mass index and obesity). The timing of oxidative stress measurements may also affect the results obtained from different studies. All these factors may alter the oxidative stress status of the subjects and are among the limitations of the study. Differentiating the acute and chronic effects of hypoxemia is even more important, as chronic exposure to high oxidative stress may have different effects on oxidative stress parameters in sleep apnea and OSA patients [10].

Discussions over the benefits of CPAP therapy to reduce oxidative stress have been contentious for years. Although some authors reported that CPAP therapy was not effective to decrease oxidative stress as well as the level of antioxidant enzymes [10, 23], other studies demonstrated that CPAP therapy is highly effective to decrease the level of oxidative stress [24, 25]. Kang et al., for example, reported that one night of CPAP treatment had no effect on antioxidant status [17]. In our study, however, it was observed that all parameters (TAS, TOS, OSI, minSpO_2 , meanSpO_2) improved following a long-term therapy with positive airway pressure in OSA patients and approached the values of control subjects. Considering

the parallel findings of this study and the study by Barceló et al. which reported that TAS value returned to normal after 12 months of CPAP therapy [22], one may suggest that positive airway therapy can be effective when conducted for long-term.

In this study, not only TOS but also TAS levels were greater in patients with OSA in comparison to control group in contrast to findings by Barceló et. al. [22]. Although increased TAS level in OSA patients may be explained by the compensatory mechanism of oxidative stress in the body as stated by Verit and Erel [15], the reasons for these contradictory results regarding TAS and TOS levels should be considered carefully. Some authors suggest that different results related to oxidative stress caused by OSA may be explained by the differences in study protocols such as including/excluding patient with comorbid diseases and treated with drug therapies [17]. To our knowledge, this study is the only study investigating the relationship between oxidative stress and OSA as well as the long-term effect of CPAP therapy on oxidative stress parameters in patients having OSA without any comorbidities while.

In our study, the increase in OSI value after CPAP titration suggests that the balance between oxidant and antioxidant levels is positively affected.

Conclusion

The current study shows that high levels of TAS and TOS, which signify higher oxidative stress, are linked to OSA. CPAP therapy lasted for one-month improved antioxidant status and oxidative stress remarkably. More research is needed to find out the root cause of oxidative stress in OSA patients and its connection to variables like AHI, the length of intermittent hypoxia, the severity of OSA, smoking, as well as to comorbidities like obesity and hypertension. Furthermore, future studies investigating the effectiveness of CPAP therapy in the presence of different factors would contribute to the relevant literature.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Senaratna C.V., Perret J.L., Lodge C.J., Lowe A.J., Campbell B.E., Prevalence of obstructive sleep apnea in the general population: A systematic review, *Sleep Medicine Reviews*, 34 (2017) 70-81.
- [2] Strange C., Richard C.L., Shan S., Phillips B.A., Kanotra S., A population-based estimate of the health care burden of obstructive sleep apnea using a STOP-BAG questionnaire in South Carolina, *Journal of Clinical Sleep Medicine*, 17(3) (2021) 367-374.
- [3] Fiedorczuk P., Stróżyński A., Olszewska E., Is the oxidative stress in obstructive sleep apnea associated with cardiovascular complications? - Systematic Review, *Journal of Clinical Medicine*, 9(11) (2020) 3734.
- [4] Cai H., Harrison D.G., Endothelial dysfunction in cardiovascular diseases: the role of oxidant stress, *Circulation Research*, 87(10) (2000) 840-844.
- [5] Landmesser U., Harrison D.G., Oxidant stress as a marker for cardiovascular events: Ox marks the spot, *Circulation*, 104 (2001) 2638-2640.
- [6] Mathur S., Devaraj S., Jialal I. Accelerated atherosclerosis, dyslipidemia, and oxidative stress in end-stage renal disease, *Current Opinion in Nephrology and Hypertension* 2002; 11(2)() 141-147.
- [7] Gozal D., Kheirandish-Gozal L., Cardiovascular morbidity in obstructive sleep apnea: oxidative stress, inflammation, and much more, *American Journal of Respiratory and Critical Care Medicine*, 177(4) (2008) 369-375.
- [8] Suzuki Y.J., Jain V., Park A.M., Day R.M., Oxidative stress and oxidant signaling in obstructive sleep apnea and associated cardiovascular diseases, *Free Radical Biology and Medicine*, 40(10) (2006) 1683-1692.
- [9] Maniaci A., Iannella G., Cocuzza S., Vicini C., Magliulo G., Oxidative stress and inflammation biomarker expression in obstructive sleep apnea patients, *Journal of Clinical Medicine*, 10(2) (2021) 277.
- [10] Svatikova A., Wolk R., Lerman L.O., Juncos L.A., Greene E.L., Oxidative stress in obstructive sleep apnoea, *European Heart Journal*, 26(22) (2005) 2435-2439.
- [11] Lavie L., Oxidative stress - a unifying paradigm in obstructive sleep apnea and comorbidities, *Progress in Cardiovascular Diseases*, 51(4) (2009) 303-312.
- [12] Berry R.B., Budhiraja R., Gottlieb D.J., Gozal D., Iber C., American Academy of Sleep Medicine. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events, Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine, *Journal of Clinical Sleep Medicine*, 8(5) (2012) 597-619.
- [13] Ghiselli A., Serafini M., Natella F., Scaccini C., Total antioxidant capacity as a tool to assess redox status: critical view and experimental data, *Free Radical Biology and Medicine*, 29(11) (2000) 1106-1114.
- [14] Erel O., A novel automated direct measurement method for total antioxidant capacity using a new generation, more stable ABTS radical cation, *Clinical Biochemistry*, 37(4) (2004) 277-285.
- [15] Erel O., A novel automated method to measure total antioxidant response against potent free radical reactions, *Clinical Biochemistry*, 37(2) (2004) 112-119.
- [16] Badran M., Ayas N., Laher I., Cardiovascular complications of sleep apnea: role of oxidative stress, *Oxidative Medicine and Cellular Longevity*, (2014) 985258.
- [17] Kang H.H., Kim I.K., Lee S.H., Total oxidant and antioxidant status in patients with obstructive sleep apnea and the effect of continuous positive airway pressure, *Chronobiology in Medicine*, 1(3) (2019) 126-129.
- [18] Alzogaibi M.A., Bahammam A.S., Lipid peroxides, superoxide dismutase and circulating IL-8 and GCP-2 in patients with severe obstructive sleep apnea: a pilot study, *Sleep and Breathing*, 9(3) (2005) 119-126.
- [19] Öztürk L., Mansour B., Yüksel M., Yalçın A.S., Celikoğlu F., Lipid peroxidation and osmotic fragility of red blood cells in sleep-apnea patients, *Clinica Chimica Acta*, 332 (2003) 83-88.
- [20] Wali S.O., Bahammam A.S., Massaeli H., Pierce G.N., Iliskovic N., Susceptibility of LDL to oxidative stress in obstructive sleep apnea, *Sleep*, 21(3) (1998) 290-296.
- [21] Schulz R., Mahmoudi S., Hattar K., Sibelius U., Olschewski H., Enhanced release of superoxide from

- polymorphonuclear neutrophils in obstructive sleep apnea, Impact of continuous positive airway pressure therapy, *American Journal of Respiratory and Critical Care Medicine*, 162 (2000) 566-570.
- [22] Barceló A., Barbé F., de la Peña M., Vila M., Pérez G., Antioxidant status in patients with sleep apnoea and impact of continuous positive airway pressure treatment, *The European Respiratory Journal*, , 27(4) (2000) 756-760.
- [23] Stanek A., Brożyna-Tkaczyk K., Myśliński W., Oxidative Stress Markers among Obstructive Sleep Apnea Patients, *Oxidative Medicine and Cellular Longevity*, (2021) 9681595.
- [24] Celec P., Jurkovičová I., Buchta R., Bartík I., Gardlí R., Antioxidant vitamins prevent oxidative and carbonyl stress in an animal model of obstructive sleep apnea, *Sleep and Breathing*, 17(2) (2013) 867-871.
- [25] Karamanlı H., Özol D., Ugur K.S., Yıldırım Z., Armutçu F., Influence of CPAP treatment on airway and systemic inflammation in OSAS patients, *Sleep and Breathing*, 18(2) (2014) 251-256.

A Quantum Chemical Study on the Antioxidant Properties of Myricetin, Quercetin, and Kaempferol Using Density Functional Theory (DFT) and Molecular Docking

Mustafa Elik ^{1,a,*}¹ Department of Mathematics and Science Education, Faculty of Education, Sivas Cumhuriyet University, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 09/06/2024

Accepted: 11/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

melik@cumhuriyet.edu.tr<https://orcid.org/0000-0001-8245-4273>

ABSTRACT

This study explores the antioxidant properties of myricetin, quercetin, and kaempferol using density functional theory (DFT) and molecular docking analyses. The findings of this study not only highlight the significant antioxidant potential of flavonoids and provide a foundation for potential clinical applications and future research aimed at optimizing their pharmacokinetic properties for therapeutic use. The optimized geometries of these flavonoids were determined through DFT calculations at the B3LYP/6-31G(d,p) level. Key electronic properties, such as total electronic energy (E_t), zero-point energy (ZPE), enthalpy (H), entropy (S), dipole moment (μ), and HOMO-LUMO energy gaps (ΔE_{H-L}), were computed to evaluate molecular stability and reactivity. Molecular docking with proteins 1HD2, 1RAU, 3FZS, and 3GRS revealed strong binding affinities, particularly for myricetin, which achieved docking scores of -5.330 with 1HD2 and -8.652 with 3GRS, indicating a high potential for antioxidant activity. ADME analyses demonstrated that kaempferol exhibited the most favorable pharmacokinetic profile, with a human oral absorption rate of 63.637% and compliance with Lipinski's Rule of Five. Conversely, myricetin showed reduced bioavailability despite its strong binding affinity. This work emphasizes optimizing flavonoid pharmacokinetics to maximize their therapeutic potential while reinforcing their relevance in antioxidant therapy.

Keywords: Antioxidant activity, Density functional theory (DFT), Molecular docking, Flavonoids and pharmacokinetics (ADME).

Introduction

In the 21st century, advancing technology, environmental pollution, greenhouse effects, petrochemical products, X-UV radiation (photochemical), pharmaceuticals, and smoking have significantly increased our exposure to oxidative stress-inducing substances. These factors manifest their impact through the formation of free radicals. Additionally, biological systems constantly produce free radicals and other reactive oxygen species due to endogenous and exogenous stress factors. To protect against these stress factors, our body has an enzyme system that defends cells from oxidative damage [1-6]. The substances that enhance the effectiveness of these enzymes are known as antioxidants [7]. Antioxidants are obtained either endogenously within the body or exogenously through plants, fruits, and other foods. There are many different substances known for their antioxidant properties. Some of these are obtained through our daily diet, while the body produces others. Free radicals are molecules that attack cells and the immune system. In contrast, antioxidants are molecules that prevent the destructive effects of free radicals and inhibit chain reactions that can lead to various diseases and premature aging [8, 9].

By forming a protective shield against free radicals, antioxidants play a crucial role in preventing the chain reactions that can lead to many diseases and premature aging. The formation of these free radicals, which pose a significant risk to the body, is facilitated by petrochemical products, X and UV radiation, cigarette smoke, air

pollution, and even preservatives and additives in food and beverages [10, 11]. Therefore, understanding and enhancing the role of antioxidants in our bodies is paramount in preventing these diseases [12]. Another source of free radicals is oxygen. While oxygen is essential for all vital functions, the oxygen we breathe in can also be hazardous to human health. Oxygen is necessary to release energy from food, but surrounding molecules can become oxidized during its utilization in cells. This uncontrolled production of free radicals can damage crucial cellular components such as proteins, fats, and genetic material. As cells deteriorate, a chain of chemical reactions begins, generating more free radicals [13, 14]. Moreover, as the human body ages, its antioxidant defense systems gradually weaken, and the cells' ability to repair themselves diminishes. These detrimental developments increase the risk of diseases such as cancer and heart disease [15-18]. Flavonoids are a class of polyphenolic secondary metabolites found in plants and are widely consumed in the human diet. Their basic structure, shown in Figure 1, consists of three rings commonly referred to as rings A, B, and C [19, 20].

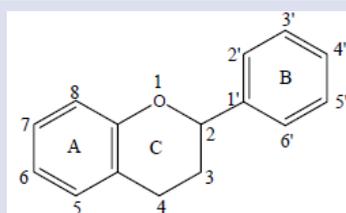


Figure 1. Basic Structure of Flavonoids

This study aims to elucidate the structural properties and activities of flavonols such as myricetin, quercetin, and kaempferol, known for their antioxidant activities, using molecular orbital methods. While significant work has been conducted on quercetin, there are fewer studies on myricetin and kaempferol and a need for more theoretical work explaining why they exhibit lower antioxidant activities. This gap in our understanding underscores the need for further research. If the relationship between these compounds can be elucidated through computational chemistry, it could guide experimental studies on newly discovered antioxidants, advancing our knowledge in this crucial area of biochemistry.

Computational chemistry plays a critical role in studying antioxidants, providing detailed insights into the molecular structure and behavior of flavonoids at the quantum mechanical level [21]. Methods such as density functional theory (DFT) allow for the precise optimization of molecular geometries and the calculation of electronic properties, stability, and reactivity. This enables the prediction of antioxidant efficiency and the identification of structural features contributing to their activity. By modeling solvent effects and exploring the impact of different substituents, computational studies can guide the design of new flavonoid derivatives with enhanced antioxidant properties, offering a cost-effective and efficient approach to advancing antioxidant research.

Computational Methods

Density Functional Theory (DFT) Study

All calculations for the compounds studied in this work were performed in the gas phase using Gaussian 2003 program [22]. The computational method utilized was Density Functional Theory (DFT), explicitly employing the closed-shell RB3LYP (Becke three-parameter hybrid functional) for non-radical species and the open-shell UB3LYP level for radicals. The basis sets used were 6-31G(d,p), which include polarization functions. Initial geometries of the compounds were generated using the GaussView 2.0 software package. The optimized geometrical structures were also obtained using the GaussView 2.0 software package [23].

Optimized geometries were determined at the B3LYP/6-31G(d,p) level for the compounds studied. Radicals were optimized at the UB3LYP/6-31G(d,p) level. The optimized geometries were verified as minimum energy structures by the absence of imaginary frequencies in the frequency calculations. Additionally, scan calculations were performed to determine the stability of the C, B, and A rings. Frequency calculations provided values for each compound's total electronic energy (E_T), internal correlation energy (E_{TC}), zero-point energy (ZPE), entropy (S), dipole moment (μ), and LUMO-HOMO energy gaps (E_{gap}). Structural parameters such as bond lengths, bond angles, and dihedral angles were also determined. Using these calculated values, the enthalpy change (ΔH) associated with OH bond dissociation,

relative energy change (ΔE), Gibbs free energy change (ΔG), and LUMO-HOMO energy gaps (E_{gap})

$$E_{TC} = E_T + ZPE$$

$$E = E_T + E_{TC}$$

$$H = E + PV = E + RT$$

$$G = H - TS$$

Molecular Docking

The potential binding modes and interactions of the ligands myricetin, quercetin, and kaempferol with antioxidant proteins urate oxidase (PDB ID: 1R4U), proline-rich tyrosine kinase 2 (PDB ID: 3FZS), and glutathione reductase (PDB ID: 3GRS) were analyzed using the Maestro 14.0, Schrödinger 2024–2 software suite [24]. The three-dimensional crystal structures of these proteins were obtained from the Protein Data Bank (PDB). Before performing the docking calculations, the proteins were prepared using the Protein Preparation module in Schrödinger Maestro 14.0. This preparation process included adding hydrogen atoms, assigning partial charges, building side chains, and completing any missing loops. Water molecules located more than 3 Å away from the binding site in the crystal structures were removed. The minimized energy configurations of the protein structures were achieved using the OPLS4 force field at a physiological pH of 7.0. Flavonoids with high phenolic content exhibit superior antioxidant activities, as highlighted in studies on phenolic compound-rich foods and beverages [25, 26]. These findings underscore the importance of optimizing flavonoid structures to maximize their therapeutic potential.

The ligands were prepared for molecular docking using the LigPrep [27] tool within the Schrödinger suite, applying the OPLS4 force field at pH 7.0 ± 2.0 . The receptor grid was generated by selecting any ligand atom to create the default grid box, which was set with a volumetric spacing of $20 \times 20 \times 20$ Å for all investigated proteins. The specific coordinates used for grid generation were x: 6.63, y: 42.27, and z: 34.64 for 1HD2; x: 31.20, y: 26.88, and z: 37.99 for 1R4U; x: -3.45, y: -3.29, and z: 12.44 for 3FZS; and x: 60.72, y: 51.36, and z: 18.87 for 3GRS. This comprehensive preparation ensures that the docking studies accurately reflect the potential interactions between the ligands and the protein binding sites, thereby providing reliable insights into their binding affinities and modes of action.

ADME Study

ADME (Absorption, Distribution, Metabolism, and Excretion) properties of myricetin, quercetin, and kaempferol were calculated using the advanced QikProp [28] tool in the Schrödinger program. The QikProp module generates relevant descriptors and uses them to carry out ADME estimations, employing Jorgensen's method to obtain pharmacokinetic properties and descriptors. This use of advanced tools and methods ensures the scientific rigor of the study, giving the audience confidence in the results.

By integrating these computational techniques, this study aims to comprehensively understand the antioxidant properties of myricetin, quercetin, and kaempferol. This comprehensive understanding is supported by molecular docking and DFT calculations, ensuring the audience is well-informed and knowledgeable about the topic.

Results and Discussion

Optimized Geometries and Electronic Properties

The ability of antioxidants to neutralize free radicals depends on the stabilization of their molecular structures and the protective functions of these structures. This study evaluates the effects of intramolecular hydrogen bonding in the radical and diradical forms of myricetin, quercetin, and kaempferol. Additionally, antioxidant activities were compared considering energy parameters (ΔE , ΔH , ΔG), dipole moment (μ), and hydrogen bond lengths. In this context, the superior properties of

quercetin diradical are particularly noteworthy. The molecular structure of flavonoids, particularly the number and position of hydroxyl groups, plays a pivotal role in their antioxidant activity, as demonstrated in recent studies (29, 30). These structural features enhance their ability to donate hydrogen atoms and interact with reactive oxygen species effectively.

The quercetin diradical (Q_{kdr}/2) stands out with its high energy stability (ΔE : 72.80 kcal/mol, ΔH : 73.39 kcal/mol) and significant dipole moment (6.6901 D). The hydrogen bonds in this compound (H^1 -Bond: 1.771 Å, H^2 -Bond: 1.944 Å) form a strong intramolecular network that maintains the structural integrity and stability of the molecule. This robust structure allows effective interactions with free radicals, producing high antioxidant activity for the quercetin diradical. These properties make the quercetin diradical a strong candidate for combating oxidative stress. As shown in Figure 1, quercetin diradical's optimized structure and hydrogen bonding network clearly illustrate its capability to stabilize and neutralize free radicals efficiently.

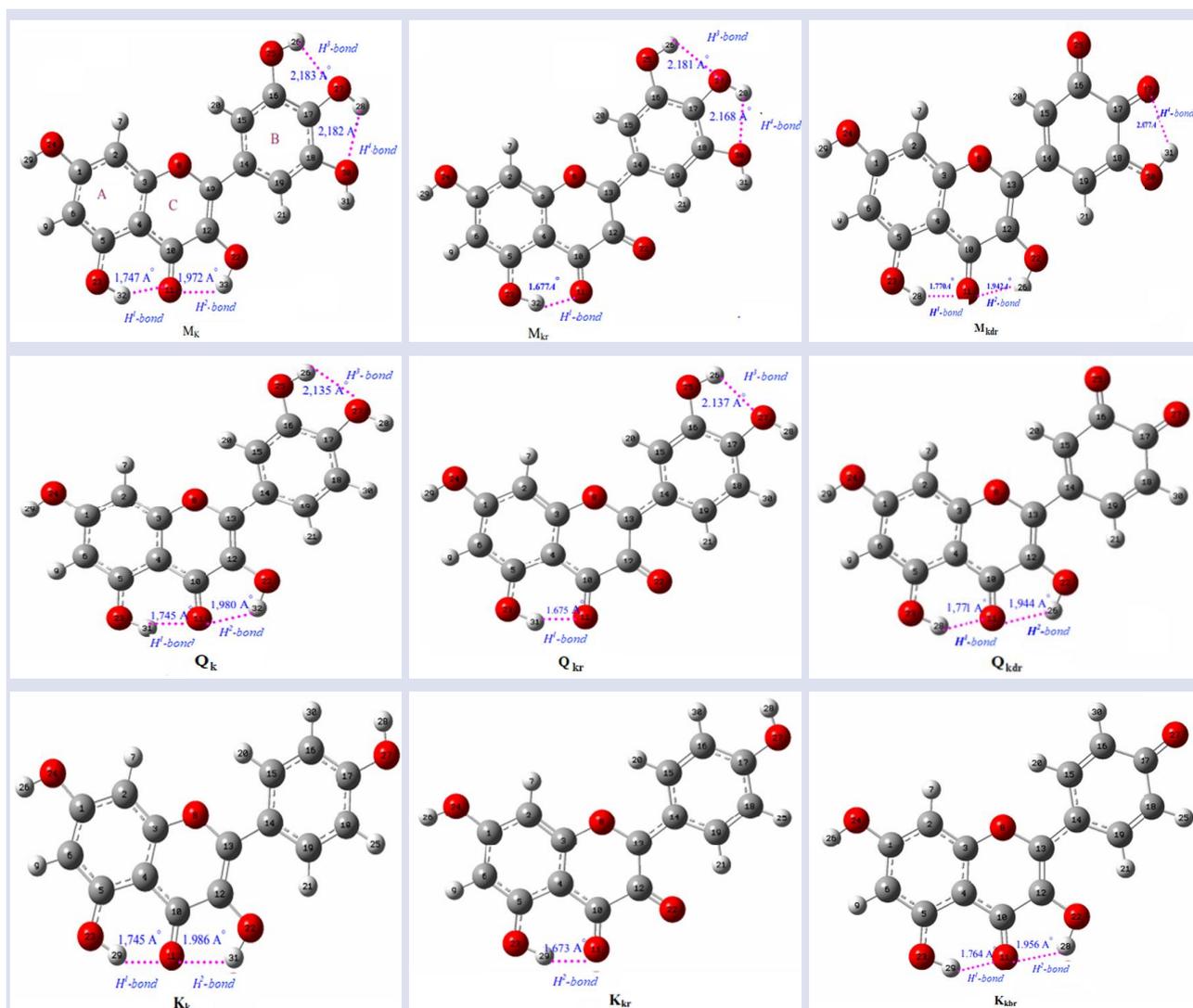


Figure 2. Optimized structures and intramolecular hydrogen bonds of myricetin, quercetin, and kaempferol and their radicals

Although the myricetin and kaempferol compounds also possess strong hydrogen bonds and good energy stability, they do not perform as well as quercetin in terms of dipole moment and the efficacy of hydrogen bonds. The myricetin radical (Mkr) and diradical (Mkdr) gain stability from their hydrogen bonds (Mkr: H¹-Bond: 1.677 Å, H²-Bond: 2.181 Å, H³-Bond: 2.168 Å; Mkdr: H¹-Bond: 1.770 Å, H²-Bond: 1.942 Å, H³-Bond: 2.118 Å), but their energy parameters and dipole moments are not as high as those of quercetin. Similarly, although kaempferol possesses strong hydrogen bonds (Kkr: H¹-Bond: 1.673 Å; Kbkr: H¹-Bond: 1.764 Å, H²-Bond: 1.956 Å), its energy stability and dipole moment are lower compared to the other two compounds.

In conclusion, the quercetin diradical compound has the highest antioxidant activity, attributed to its high energy stability, strong intramolecular hydrogen bonds, and significant dipole moment. Myricetin ranks second with its strong hydrogen bonds and good energy stability, while kaempferol ranks third. This study highlights the critical role of intramolecular hydrogen bonding and energy parameters in antioxidant efficacy, providing valuable insights for designing new compounds. These findings offer important clues for optimizing the structure and enhancing the effectiveness of antioxidants.

The DFT calculations performed at the B3LYP/6-31G(d,p) level, have been instrumental in determining the optimized geometries of myricetin, quercetin, and kaempferol. The absence of imaginary frequencies in the vibrational analysis has unequivocally confirmed these optimized structures as minimum energy configurations. Fundamental electronic properties, including total electronic energy (E_T), zero-point energy (ZPE), entropy (S), dipole moment (μ), and LUMO-HOMO energy gaps (ΔE_{L-H}), were meticulously calculated. These properties are pivotal for understanding the intrinsic stability and reactivity of the flavonoids. The antioxidant activity of flavonoids is directly influenced by their molecular structures, particularly the number and position of hydroxyl groups, which enhance their ability to donate hydrogen atoms and neutralize free radicals effectively.

Thermodynamic Parameters

The thermodynamic parameters derived from the optimized geometries and electronic properties of myricetin, quercetin, and Kaempferol provide critical insights into their antioxidant capacities. Specifically, we have computed the enthalpy change (ΔH) associated with OH bond dissociation, relative energy change (ΔE), and Gibbs free energy change (ΔG). These parameters quantitatively measure the flavonoids' ability to donate hydrogen atoms and neutralize free radicals, which is crucial for their potential therapeutic applications.

For myricetin radical (Mkr), the enthalpy change (ΔH) is 79.44 kcal/mol, and the relative energy change (ΔE) is 78.85 kcal/mol, indicating a significant amount of energy required for OH bond dissociation. The Gibbs free energy change (ΔG) is 71.21 kcal/mol, reflecting the reaction's spontaneity. Myricetin diradical (Mkdr/2) shows slightly lower values with ΔH of 71.23 kcal/mol, ΔE of 70.51 kcal/mol, and ΔG of 63.21 kcal/mol. These values suggest that myricetin possesses substantial thermodynamic stability and effective hydrogen-donating capability, especially in its radical form.

Quercetin radical (Qkr) exhibits a ΔH of 80.14 kcal/mol and an ΔE of 79.55 kcal/mol, the highest among the studied flavonoids, indicating a robust bond dissociation energy. The ΔG value for Qkr is 71.85 kcal/mol, demonstrating its significant free energy change. The quercetin diradical (Qkdr/2) presents an enthalpy change (ΔH) of 73.39 kcal/mol, a relative energy change (ΔE) of 72.80 kcal/mol, and a Gibbs free energy change (ΔG) of 65.27 kcal/mol. The notable thermodynamic stability of quercetin, particularly in its diradical form, highlights its superior antioxidant activity, supported by its strong hydrogen-bonding network (H¹-Bond: 1.771 Å, H²-Bond: 1.944 Å). As seen in Table 1, these values underscore quercetin's remarkable capability to stabilize and neutralize free radicals efficiently.

Kaempferol radical (Kkr) shows a ΔH of 80.22 kcal/mol and a ΔE of 79.63 kcal/mol, with a Gibbs free energy change (ΔG) of 71.86 kcal/mol, suggesting high energy stability. The Kaempferol B-ring radical (Kbkr) displays an enthalpy change (ΔH) of 80.66 kcal/mol, a relative energy change (ΔE) of 80.07 kcal/mol, and a Gibbs free energy change (ΔG) of 72.39 kcal/mol. These values indicate that Kaempferol has substantial antioxidant potential, albeit slightly lower than quercetin, due to its energy parameters and hydrogen-bond lengths (H¹-Bond: 1.764 Å, H²-Bond: 1.956 Å). Energy parameters such as ΔE , ΔH , and ΔG play a critical role in determining the antioxidant capacity of flavonoids, with lower ΔH and ΔG values corresponding to higher hydrogen-donating capabilities and stability against oxidative stress [31, 32].

In conclusion, the thermodynamic parameters reveal that quercetin, particularly in its diradical form (Qkdr/2), exhibits the highest antioxidant activity due to its superior energy stability and strong intramolecular hydrogen bonding. Myricetin follows, demonstrating significant hydrogen-donating capability and thermodynamic stability, especially in its radical form. Kaempferol ranks third while showing considerable antioxidant potential due to slightly lower energy stability and dipole moments. These findings provide valuable insights for designing new antioxidant compounds with enhanced stability and efficacy for therapeutic applications.

Table 1. Thermodynamic parameters and intramolecular hydrogen bonding lengths of myricetin, quercetin, and kaempferol, and their radicals

| Comp. | ΔE | ΔH | ΔG | μ | Δ_{L+H} | H ¹ -Bond | H ² -Bond | H ³ -Bond | H ⁴ -Bond | H ⁵ -Bond |
|--------|------------|------------|------------|--------|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Mkr | 78.85 | 79.44 | 71.21 | 3.6225 | 5.49 | 1.677 | - | - | 2.181 | 2.168 |
| Mkdr/2 | 70.51 | 71.23 | 63.21 | 5.1623 | 6.61 | 1.770 | 1.942 | 2.118 | - | - |
| Qkr | 79.55 | 80.14 | 71.85 | 4.9893 | 6.12 | 1.675 | - | 2.137 | - | - |
| Qkdr/2 | 72.80 | 73.39 | 65.27 | 6.6901 | 6.99 | 1.771 | 1.944 | - | - | - |
| Kkr | 79.63 | 80.22 | 71.86 | 6.9160 | 6.06 | 1.673 | - | - | - | - |
| Kbkr | 80.07 | 80.66 | 72.39 | 5.5238 | 6.40 | 1.764 | 1.956 | - | - | - |

To investigate the antioxidant activities of myricetin, quercetin, and kaempferol, we first determined the optimized structures of these compounds based on the structural parameters obtained from calculations. Upon examining the optimized geometric structures of these compounds, it is evident that the only structural difference lies in the number of hydroxyl groups on the B-ring. Myricetin has 3 OH groups, quercetin has 2 OH groups, and kaempferol has 1 OH group. According to the literature, the TEAC values for myricetin, quercetin, and kaempferol are 3.10, 4.70, and 1.34, respectively [33]. If there were a direct relationship between the number of hydroxyl groups and antioxidant activity, myricetin, with the most OH groups, would be expected to have the highest TEAC value.

Conversely, if having fewer OH groups resulted in higher antioxidant activity, kaempferol would have the highest TEAC value. However, quercetin has the highest TEAC value among the three. Therefore, there is no direct relationship between the number of OH groups and the TEAC value. These compounds exhibit antioxidant properties due to their ability to scavenge free radicals. As these compounds exhibit this effect, they donate hydrogen to free radicals, forming radicals and diradicals. These radicals can form chelate rings with trace metals present in our bodies, facilitating the easy release of hydrogen atoms, as shown in Figure 3. The released hydrogen atoms then bind to free radicals, neutralizing them [33]. As illustrated in Figure 4, the reactivity (energy) of the formed diradicals is crucial; the higher their reactivity, the stronger the chelate ring they create, thus enhancing their antioxidant capacity.

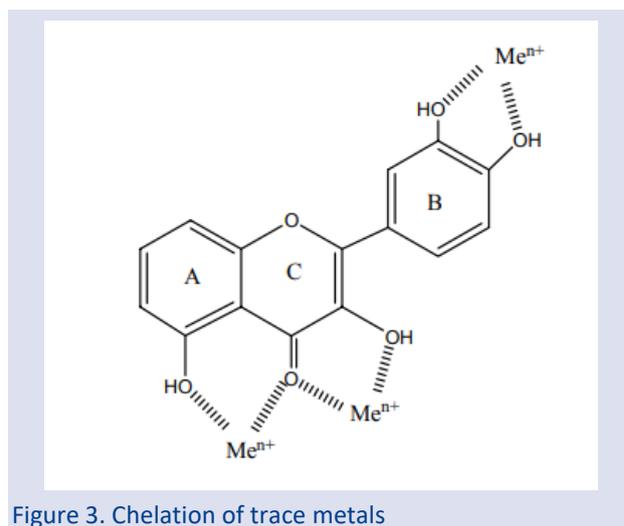


Figure 3. Chelation of trace metals

Additionally, to investigate the antioxidant activities of myricetin, quercetin, and kaempferol, we first determined the optimized structures of these compounds based on the structural parameters obtained from calculations. Upon examining the optimized geometric structures of these compounds, it is evident that the only structural difference lies in the number of hydroxyl groups on the B-ring. Myricetin has 3 OH groups, quercetin has 2 OH groups, and kaempferol has 1 OH group. According to the literature, the TEAC values for myricetin, quercetin, and kaempferol are 3.10, 4.70, and 1.34, respectively [34]. If there were a direct relationship between the number of hydroxyl groups and antioxidant activity, myricetin, with the most OH groups, would be expected to have the highest TEAC value.

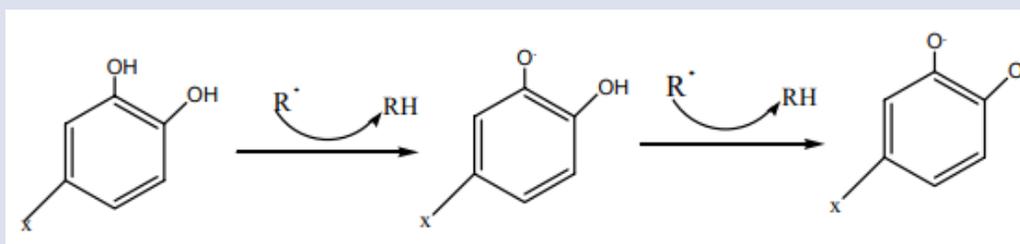


Figure 4. Scavenging of Reactive Free Radicals by Flavonoids

Conversely, if having fewer OH groups resulted in higher antioxidant activity, kaempferol would have the highest TEAC value. However, quercetin has the highest TEAC value among the three. Therefore, there is no direct relationship between the number of OH groups and the TEAC value. Among the flavonoids studied, quercetin's superior antioxidant activity is attributed to its optimal balance of electronic properties and strong hydrogen bonding network, highlighting its potential as a leading antioxidant candidate for therapeutic applications.

Molecular Docking

This study investigates the antioxidant properties of Myricetin, Quercetin, and Kaempferol, offering a comprehensive evaluation through molecular docking and

ADME (Absorption, Distribution, Metabolism, and Excretion) analyses. In the molecular docking analysis, the interactions of these flavonoids with four specific proteins—1HD2, 1R4U, 3FZS, and 3GRS—were meticulously examined. The results indicate that Myricetin possesses the highest binding affinity among all the tested proteins. Notably, Myricetin achieved shallow docking scores of -5.330 with 1HD2 and an impressive -8.652 with 3GRS, underscoring its robust binding properties. These findings strongly suggest that Myricetin forms stable interactions with the active sites of these proteins, demonstrating a high potential for antioxidant activity.

As illustrated in Figure 5, the detailed 3D visualization of the docking interactions between these flavonoids and

the respective proteins provides further insight into their binding modes. Each row of the figure represents a different protein (1HD2, 1R4U, 3FZS, and 3GRS), while each column shows the binding of Quercetin, Kaempferol, and Myricetin, respectively. The colorful representations highlight the complex interactions within the protein active sites, showcasing how Myricetin, Quercetin, and Kaempferol are positioned and stabilized. The figure vividly demonstrates Myricetin's superior binding affinity, particularly with 3GRS, where the intricate binding network is visible. These visualizations are crucial for understanding the molecular mechanisms that underpin the high docking scores and the potential antioxidant efficacy of these compounds.

Further emphasizing the robustness of these interactions, Figure 6 presents a detailed 2D visualization of the interactions between the receptor regions of the target proteins and the studied compounds. This figure highlights the specific binding modes and critical interactions, such as hydrogen bonds and hydrophobic contacts that these flavonoids form with the proteins. The visualization clearly shows the interaction sites and the nature of the interactions, providing insights into the molecular mechanisms underpinning the high docking scores. Such detailed interaction maps are crucial for understanding how these compounds stabilize within the protein active sites, facilitating their antioxidant action.

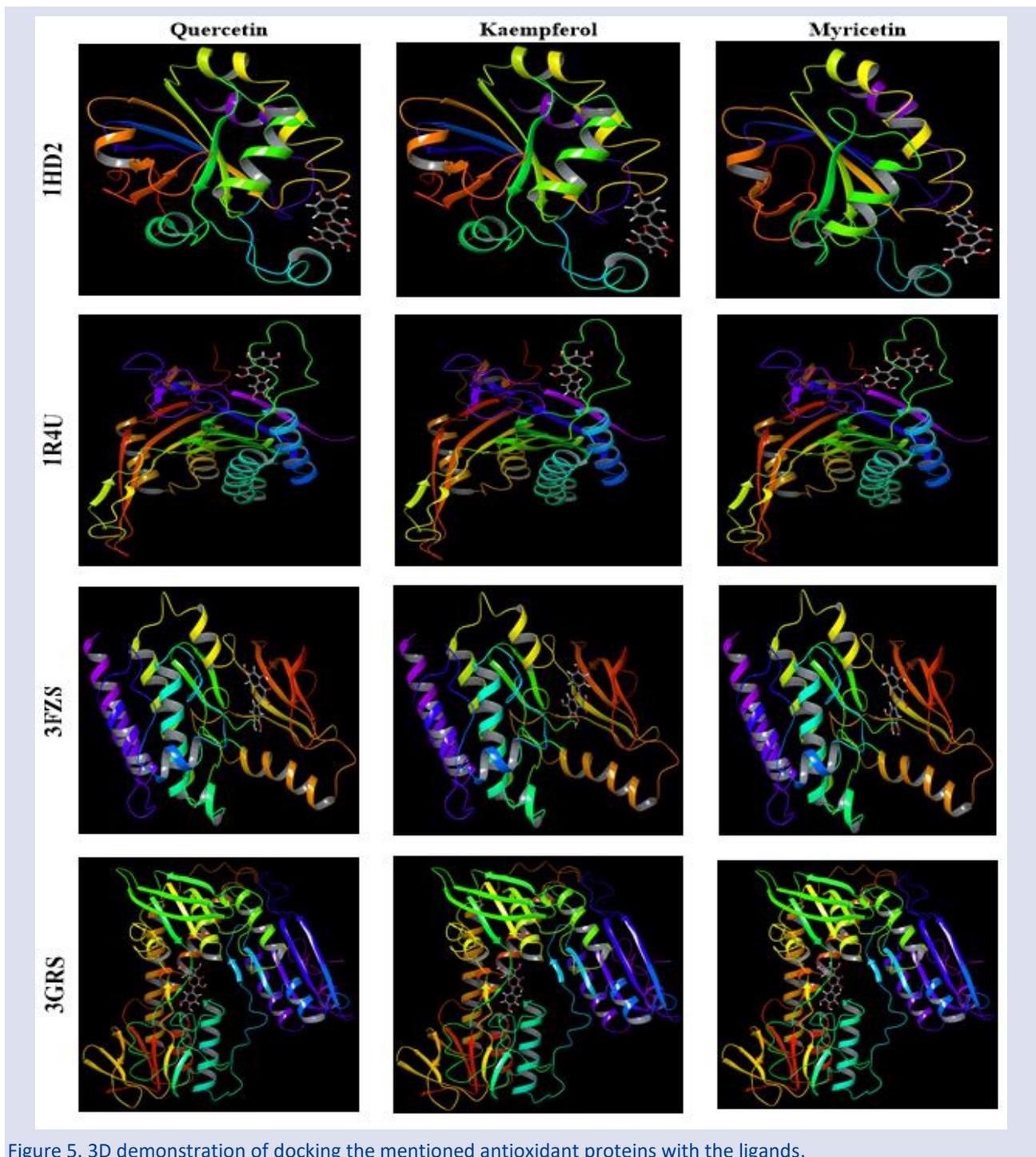


Figure 5. 3D demonstration of docking the mentioned antioxidant proteins with the ligands.

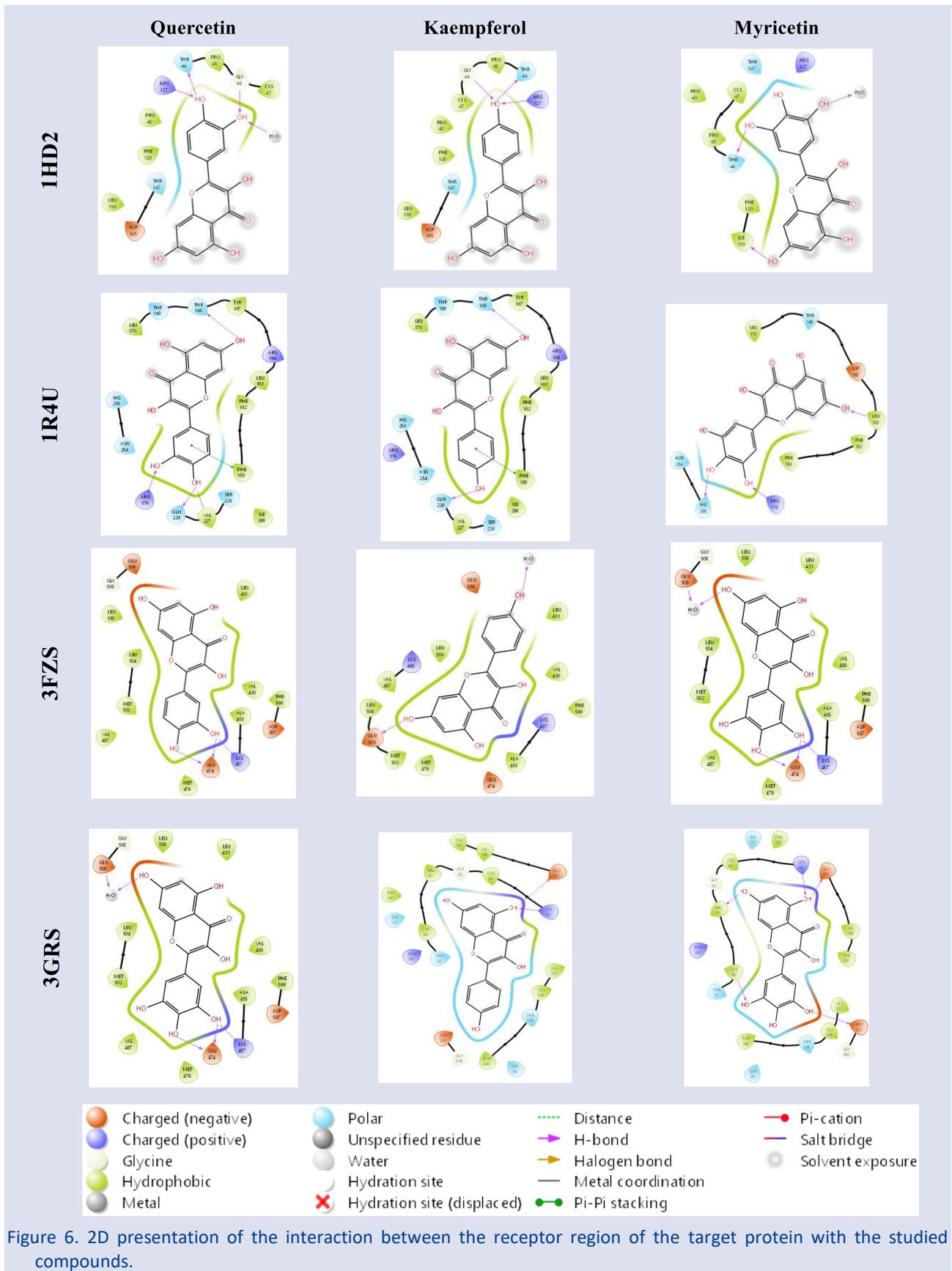


Figure 6. 2D presentation of the interaction between the receptor region of the target protein with the studied compounds.

As seen in Table 2, Myricetin consistently showed the highest docking scores compared to Quercetin and Kaempferol across all tested proteins, reinforcing its potential as a strong antioxidant agent. This superior performance is due to Myricetin's robust binding

properties, which enable it to form stable interactions with the active sites of these proteins. The high docking scores of Myricetin, particularly with proteins 1HD2 and 3GRS, underscore its ability to bind and stabilize within the protein structures effectively. These stable

interactions are crucial for neutralizing free radicals, thereby protecting cells from oxidative stress and contributing to Myricetin's high potential for antioxidant activity. The detailed interaction maps in Figures 1 and 2 illustrate how Myricetin fits into the protein binding sites, forming critical hydrogen bonds and hydrophobic

contacts that enhance its binding affinity. This strong binding capability, combined with its potential for effective free radical neutralization, makes Myricetin a promising candidate for further research and development in antioxidant therapies.

Table 2. The docking results obtained from interaction of the quercetin, kaempferol, and myricetin ligands with the mentioned proteins

| Protein | Ligand | Docking score | Glide hbond | Glide emodel | Glide ligand efficiency | Glide energy |
|---------|------------|---------------|-------------|--------------|-------------------------|--------------|
| 1HD2 | Quercetin | -4.427 | 0.000 | -39.799 | -0.201 | -32.473 |
| | Kaempferol | -4.324 | -0.618 | -34.030 | -0.206 | -27.642 |
| | Myricetin | -5.330 | -0.312 | -39.985 | -0.232 | -33.166 |
| 1R4U | Quercetin | -5.130 | -0.304 | -38.596 | -0.233 | -30.473 |
| | Kaempferol | -5.207 | -0.778 | -40.067 | -0.248 | -31.606 |
| | Myricetin | -4.363 | 0.000 | -40.203 | -0.190 | -32.843 |
| 3FZS | Quercetin | -6.953 | 0.000 | -55.423 | -0.316 | -38.505 |
| | Kaempferol | -7.544 | -0.335 | -57.252 | -0.359 | -40.407 |
| | Myricetin | -6.895 | 0.000 | -59.023 | -0.300 | -40.893 |
| 3GRS | Quercetin | -8.206 | -0.320 | -79.799 | -0.373 | -54.510 |
| | Kaempferol | -8.206 | -0.320 | -79.799 | -0.373 | -54.510 |
| | Myricetin | -8.652 | -0.297 | -71.841 | -0.376 | -48.929 |

ADME Analysis

The ADME analysis reveals critical insights into the pharmacokinetic profiles of Myricetin, Quercetin, and Kaempferol, which are essential for assessing their suitability as therapeutic agents. Kaempferol stands out with the most favorable pharmacokinetic properties, exhibiting the highest human oral absorption rate at 63.637% and full compliance with Lipinski's Rule of Five, indicating excellent potential for effective absorption and utilization in the human body. Quercetin shows moderate absorption at 51.649%, making it a viable candidate for further development. In contrast, Myricetin, despite its superior binding affinity in molecular docking studies, has the lowest human oral absorption rate at 26.816%, highlighting potential limitations in its bioavailability and necessitating optimization to harness its therapeutic potential fully.

Moreover, the analysis of molecular properties such as dipole moment, hydrogen bond donors and acceptors, and partition coefficient (QPlogPo/w) underscores the

nuanced pharmacokinetic behaviors of these flavonoids. Myricetin's high dipole moment (6.539) and maximum hydrogen bond interactions contribute to its robust binding properties but may also influence its lower lipophilicity and absorption. Kaempferol achieves a superior pharmacokinetic profile with its favorable QPlogPo/w (1.041) and optimal balance of hydrogen bonding properties. Quercetin's moderate properties align with its satisfactory absorption rates. The antioxidant properties of flavonoids derived from natural sources, such as dandelion and hibiscus, demonstrate their broad applicability in functional foods and pharmaceuticals [35]. Their bioavailability and pharmacokinetic behaviors warrant further investigation to enhance their therapeutic efficacy. The comprehensive integration of docking and ADME results underscore the importance of Kaempferol and Quercetin as promising therapeutic candidates while indicating that enhancing Myricetin's pharmacokinetic characteristics could significantly improve its therapeutic efficacy.

Table 3. Pharmacokinetic and ADME Parameters of the Investigated Ligands

| Ligand | Dipole | Mol MW | Donor HB | Accpt HB | QPlogPo/w | No. of meta | QPlogKhsa | Rule of Five | % Human oral absorption |
|-------------|--------|---------|----------|----------|-----------|-------------|-----------|--------------|--------------------------------|
| Quercetin | 4.721 | 302.240 | 4.000 | 5.250 | 0.367 | 5 | -0.343 | 0 | 51.649 |
| Kaempferol | 4.456 | 286.240 | 3.000 | 4.500 | 1.041 | 4 | -0.191 | 0 | 63.637 |
| Myricetin | 6.539 | 318.239 | 5.000 | 6.000 | -0.299 | 6 | -0.489 | 1 | 26.816 |
| Recommended | 1–12.5 | 130–725 | 0–6 | 2–20 | -2–6.5 | 1–8 | -1.5–1.5 | Max 4 | > 80% is high < 25% is poor |

This study investigates the antioxidant properties of Myricetin, Quercetin, and Kaempferol, offering a comprehensive evaluation through molecular docking and ADME (Absorption, Distribution, Metabolism, and Excretion) analyses. In the molecular docking analysis, the interactions of these flavonoids with four specific proteins—1HD2, 1R4U, 3FZS, and 3GRS—were

meticulously examined. The results indicate that Myricetin possesses the highest binding affinity among all the tested proteins. Notably, Myricetin achieved shallow docking scores of -5.330 with 1HD2 and an impressive -8.652 with 3GRS, underscoring its robust binding properties. These findings strongly suggest that Myricetin forms stable interactions with the active sites of these

proteins, demonstrating a high potential for antioxidant activity.

Further emphasizing the robustness of these interactions, Figure 2 presents a detailed 2D visualization of the interactions between the receptor regions of the target proteins and the studied compounds. This figure highlights the specific binding modes and critical interactions, such as hydrogen bonds and hydrophobic contacts that these flavonoids form with the proteins. The visualization clearly shows the interaction sites and the nature of the interactions, providing insights into the molecular mechanisms underpinning the high docking scores. Such detailed interaction maps are crucial for understanding how these compounds stabilize within the protein active sites, facilitating their antioxidant action. These findings underline the importance of exploring structural modifications to flavonoids, such as optimizing the hydroxyl group arrangement and enhancing pharmacokinetic properties, to develop more potent and bioavailable antioxidant therapies.

The ADME analyses provide crucial insights into the pharmacokinetic profiles of Myricetin, Quercetin, and Kaempferol. Kaempferol is a promising candidate for therapeutic applications, given its favorable pharmacokinetic properties, including a high percentage of human oral absorption at 63.637% and full compliance with Lipinski's Rule of Five. In contrast, Myricetin, despite its superior binding affinity, shows a lower human oral absorption rate of 26.816%, indicating potential limitations in its bioavailability. Quercetin, with a human oral absorption rate of 51.649%, displays moderate pharmacokinetic properties. These findings underscore the need to optimize Myricetin's pharmacokinetic properties to harness its therapeutic potential fully. The comprehensive analysis of docking and ADME results underscores the importance of these flavonoids in antioxidant therapy, providing valuable information for future research.

Structural modifications to flavonoids, including optimizing hydroxyl group positions and enhancing phenolic content, are crucial for improving antioxidant efficacy, as evidenced by recent findings [25, 36]. Future research should focus on these aspects to fully harness the therapeutic potential of these compounds.

Conclusion

The comprehensive evaluation of myricetin, quercetin, and kaempferol through DFT calculations and molecular docking has provided valuable insights into their antioxidant properties. Myricetin exhibited the highest binding affinity in docking studies, underscoring its robust interaction with protein active sites and potential as a strong antioxidant agent. Table 1 shows quercetin demonstrated the highest antioxidant activity with its superior thermodynamic stability and significant dipole moment, especially in its diradical form (Qkdr/2). This finding emphasizes the critical role of energy stability and intramolecular hydrogen bonding in enhancing

antioxidant efficacy. While showing substantial antioxidant potential, kaempferol ranked third due to its lower energy stability and dipole moment than quercetin and myricetin. In conclusion, quercetin is the most potent antioxidant among the three flavonoids studied, followed by myricetin and Kaempferol. The study underscores the need for further research to optimize the pharmacokinetic properties of these compounds, particularly myricetin, to exploit their therapeutic potential fully. These findings contribute to the ongoing efforts to design new antioxidant compounds with enhanced stability and efficacy, providing a foundation for future experimental and theoretical investigations in antioxidant research.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Mohamed N., Effect of antioxidants in promoting plant growth under climate change conditions, *Research on Crops*, 21 (2) (2020) 424-434.
- [2] Agrimonti C., Lauro M., Visioli G., Smart agriculture for food quality: Facing climate change in the 21st century, *Crit. Rev. Food Sci. Nutr.*, 61 (6) (2021) 971-981.
- [3] Orona-Tamayo D., Paredes-López O., Chia—The new golden seed for the 21st century: Nutraceutical properties and technological uses, *Sustainable Protein Sources*, Academic Press, (2024) 443-470.
- [4] Bir A., Ghosh A., Chauhan A., Saha S., Saini A. K., Bisaglia, M., Chakrabarti, S., Exosomal Dynamics and Brain Redox Imbalance: Implications in Alzheimer's Disease Pathology and Diagnosis, *Antioxidants*, 13 (3) (2024) 316.
- [5] Tumilaar S. G., Hardianto A., Dohi H., Kurnia D., A Comprehensive Review of Free Radicals, Oxidative Stress, and Antioxidants: Overview, Clinical Applications, Global Perspectives, Future Directions, and Mechanisms of Antioxidant Activity of Flavonoid Compounds, *J. Chem.*, 2024 (1) (2024) 5594386.
- [6] Caiati C., Stanca A., Lepera M. E., Free radicals and obesity-related chronic inflammation contrasted by antioxidants: A new perspective in coronary artery disease, *Metabolites*, 13 (6) (2023) 712.
- [7] Pruteanu L. L., Bailey D. S., Grădinaru A. C., Jäntschi L., Pammi S. S., Suresh B., Giri A., The biochemistry and effectiveness of antioxidants in food, fruits, and marine algae, *Antioxidants*, 12 (4) (2023) 860.
- [8] Sen S., Chakraborty R., Sridhar C., Reddy Y. S. R., De B., Free radicals, antioxidants, diseases and phytochemicals: current status and future prospect, *Int J Pharm Sci Rev Res*, 3 (1) (2010) 91-100.
- [9] Andrés C. M. C., Pérez de la Lastra J. M., Bustamante Munguira E., Juan C. A., Plou F. J., Pérez Lebeña E., Electrophilic Compounds in the Human Diet and Their Role in the Induction of the Transcription Factor NRF2, *Int. J. Mol. Sci.*, 25 (6) (2024) 3521.
- [10] Daniloski D., Petkoska A. T., Lee N. A., Bekhit A. E. D., Carne A., Vaskoska, R., Vasiljevic, T., Active edible packaging based on milk proteins: A route to carry and deliver nutraceuticals, *Trends Food Sci. Technol.*, 111 (2021) 688-705.

- [11] Ayoka T. O., Ezema B. O., Eze C. N., Nnadi C. O., Antioxidants for the Prevention and Treatment of Non-communicable Diseases, *J. Explor. Res. Pharmacol.*, 7 (3) (2022) 179-189.
- [12] Di Meo S., Venditti P., Evolution of the knowledge of free radicals and other oxidants, *Oxid. Med. Cell. Longev.*, 2020 (1) (2020) 9829176.
- [13] Engwa G. A., Nweke F. N., Nkeh-Chungag B. N., Free radicals, oxidative stress-related diseases and antioxidant supplementation, *Altern. Ther. Health Med.*, 28 (1) (2022).
- [14] Wang W., Kang P. M., Oxidative stress and antioxidant treatments in cardiovascular diseases, *Antioxidants*, 9 (12) (2020) 1292.
- [15] Dubois-Deruy E., Peugnet V., Turkieh A., Pinet F., Oxidative stress in cardiovascular diseases, *Antioxidants*, 9 (9) (2020) 864.
- [16] Saini N., Kataria R., Bhardwaj I., Panchal P., Importance of Flavonoids in Agriculture, Flavonoids as Nutraceuticals, *Apple Academic Press*, (2024) 37-55.
- [17] Yan Q., Liu S., Sun Y., Chen C., Yang S., Lin M., Yang Y., Targeting oxidative stress as a preventive and therapeutic approach for cardiovascular disease, *J. Transl. Med.*, 21 (1) (2023) 519.
- [18] Zandavar H., Babazad M. A., Secondary metabolites: Alkaloids and flavonoids in medicinal plants, *Herbs and Spices-New Advances*, *IntechOpen*, (2023).
- [19] Pietta P. G., Flavonoids as antioxidants, *J. Nat. Prod.*, 63 (7) (2000) 1035-1042.
- [20] Ma P., Wang Z., Density Functional Theory Study on Antioxidant Activity of Three Polyphenols, *J. Fluoresc.*, 33 (3) (2023) 933-944.
- [21] Sun C., Cao B., Yin H., Shi Y., Relationship between ES IPT properties and antioxidant activities of 5-hydroxyflavone derivatives, *Chin. Phys. B*, 29(5) (2020) 058202.
- [22] Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Montgomery J. A., Jr. Vreven T., Kudin K. N., Burant J. C., Millam J. M., Iyengar S. S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G. A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J. E., Hratchian H. P., Cross J. B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J.W., Ayala P.Y., Morokuma K., Voth G. A., Salvador P., Dannenberg J.J., Zakrzewski V.G., Dapprich S., Daniels A.D., Strain M.C., Farkas O., Malick D.K., Rabuck A. D., Raghavachari K., Foresman J.B., Ortiz J. V., Cui Q., Baboul A. G., Clifford S., Cioslowski J., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R. L., Fox D. J., Keith T., Al-Laham M.A., Peng C. Y., Nanayakkara A., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Gonzalez C., Pople J.A., Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.
- [23] Dennington II, Keith R., Milliam T., Eppinnett J., Hovell K., Gilliland W. L. R., GaussView, Version 3.09, Semichem, Inc., Shawnee Mission, KS, 2003.
- [24] Schrödinger Release 2024-2: Maestro, Schrödinger, LLC, New York, NY, 2024.
- [25] de Lima, D. P., Júnior, E. D. S. P., de Menezes, A. V., de Souza, D. A., de São José, V. P. B., da Silva, B. P., ... & Carvalho, I. M. M., Chemical composition, minerals concentration, total phenolic compounds, flavonoids content and antioxidant capacity in organic and conventional vegetables, *Food Research International*, 175 (2024) 113684.
- [26] Indiarito, R., Herwanto, J. A., Filianty, F., Lembong, E., Subroto, E., Muhammad, D. R. A., Total phenolic and flavonoid content, antioxidant activity and characteristics of a chocolate beverage incorporated with encapsulated clove bud extract, *CyTA-Journal of Food*, 22(1) (2024) 2329144.
- [27] Schrödinger Release 2024-2: LigPrep, Schrödinger, LLC, New York, NY, 2024.
- [28] Schrödinger Release 2024-2: QikProp, Schrödinger, LLC, New York, NY, 2024.
- [29] Glevitzky I., Dumitrel G. A., Glevitzky M., Pasca B., Otrisal P., Bungau S., Popa, M. J. R. C., Statistical analysis of the relationship between antioxidant activity and the structure of flavonoid compounds, *Rev. Chim*, 70(9) (2019) 3103-3107.
- [30] Wang R., Li W., Fang C., Zheng X., Liu C., Huang Q., Extraction and identification of new flavonoid compounds in dandelion *Taraxacum mongolicum* Hand.-Mazz. with evaluation of antioxidant activities, *Scientific Reports*, 13(1) (2023) 2166.
- [31] Wang Y., Li C., Li Z., Moalin M., Hartog G. J. D., Zhang M., Computational Chemistry Strategies to Investigate the Antioxidant Activity of Flavonoids An Overview, *Molecules*, 29(11) (2024) 2627.
- [32] Miličević A., Flavonoid Oxidation Potentials and Antioxidant Activities-Theoretical Models Based on Oxidation Mechanisms and Related Changes in Electronic Structure, *International Journal of Molecular Sciences*, 25(9) (2024) 5011.
- [33] Monzani E., Koolhaas G.A.A., Spandre A., Leggieri E., Casella L., Gullotti M., Reedijk J., Binding of nitrite and its reductive activation to nitric oxide at biomimetic copper centers, *JBIC Journal of Biological Inorganic Chemistry*, 5 (2000) 251-261.
- [34] Jovanovic S. V., Steenken S., Tosic M., Marjanovic B., Simic M.G., Flavonoids as antioxidants, *J. Am. Chem. Soc.*, 116(11) (1994) 4846-4851.
- [35] Mejía J. J., Sierra L. J., Ceballos J. G., Martínez J. R., Stashenko E.E., Color, antioxidant capacity and flavonoid composition in *Hibiscus rosa-sinensis* cultivars, *Molecules*, 28(4) (2023) 1779.
- [36] Utami Y.P., Yulianty R., Djibir Y.Y., Alam G., Antioxidant Activity, Total Phenolic and Total Flavonoid Contents of *Etlingera elatior* (Jack) RM Smith from North Luwu, Indonesia, *Tropical Journal of Natural Product Research*, 8(1) (2024).

Modeling the Effect of Heat Distribution in Photothermal Therapy by Using Computational Fluid Dynamics (CFD)

Mesude Avci^{1,a,*}

¹ Department of Chemical Engineering, Sivas Cumhuriyet University, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 16/08/2024

Accepted: 14/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Cancer is a mortal disorder around the world, and according to the World Health Organization (WHO), it is a leading cause of death, causing nearly 10 million deaths in 2020. It is commonly treated by chemotherapy, radiotherapy, and surgery. However, the undesirable effects of these treatments encouraged clinicians to find better therapies, such as photothermal therapy (PTT). PTT has been commonly used for being less harmful to the healthy tissues near the cancer cells. However, it is necessary to know that the heat distribution is suitable and that the surrounding tissue is not overheated. This work uses Computational Fluid Dynamics (CFD) to model the cancer cell and the healthy tissue around it as a 3D model using ICEM CFD, a pre-processing program of Ansys Fluent 18.2. It is found that wall shear stress is high, up to 4600 Pa in the top parts of the cell, and lower in others. The highest pressure on the cancer cell goes up to 36000 Pa in the lower parts of the cell. The results of this work could guide researchers in optimizing the photothermal therapy of cancer cells, and the modeling approach could be applied to investigate alternative therapies.

Keywords: Cancer cell, Photothermal therapy, Computational Fluid Dynamics, Heat distribution, Modeling.

mesude@cumhuriyet.edu.tr

<https://orcid.org/0000-0001-8211-7779>

Introduction

Cancer is an uncontrolled failure of abnormal cells that can attack and deteriorate normal bodily tissue. In addition, cancer also expands all over the body in an unpredictable process. According to the World Health Organization (WHO), cancer is a leading cause of death, causing nearly 20 million new cancer cases and 09.7 million deaths in 2022. Moreover, the determined number of people who were alive within 5 years following a cancer diagnosis was 53.5 million. About 1 in 5 people develop cancer in their lifetime [1,2]. However, advancements in traditional cancer therapies, such as photothermal therapy, have provided a continued decrease in cancer death rates since the early 1990s [2], [3]. Cancer is commonly cured by chemotherapy, radiotherapy, and surgery. Nevertheless, these therapies may have undesirable effects on the body, such as damaging healthy cells, injuring the patient's immune system, and causing several harmful side effects. As a result of these inefficiencies, new effective therapies to destroy cancer cells without damaging healthy cells have been commonly developed [4-6]. Among other new therapies, there has been much interest in cancer photothermal therapy because of its minimal adverse effects on the healthy cells around the cancer cells.

Photothermal therapy (PTT) is a well-known method for cancer therapy that uses high temperatures generated from optical energy to activate cancer cell burn death [7-11]. Considering other therapies, such as radiotherapy and chemotherapy, PTT cancer treatment is very selective and minimally invasive because the healing effect of PTT

cancer treatment occurs only at the sick cell site without harming normal tissues [12]. Besides being an effective therapy on its own, hyperthermia has also been well applied and showed a reduction of tumor size with other tumor treatment methods such as chemotherapy and radiation with the [13-16]. Coleman et al. [17] worked on regional hyperthermia where tumor cell temperature is increased above 42°C for a couple of hours.

When heating cancer cells, it is essential to know that the healthy tissue near them is not harmed. Due to hyperthermia, the surrounding tissue and muscles are also exposed to heat. Heat distribution affects the PTT since too much heat can damage cancer cells and human tissue. Therefore, it is essential to examine the suitable heat distribution and the effect of heat in PTT so that hyperthermia will burn the cancer cells without overheating and damaging the surrounding tissues [18].

This study is different from the literature since it addresses the direct effect of temperature both on the cancer cell and also on the healthy tissue around the cancer cell. This study also carefully considers the change of different parameters, such as wall shear stress and pressure in PTT. Moreover, the temperature effect is comprehensively examined by modeling different temperature ranges for the cancer cell and the healthy tissue around it. This study examines the heat effect of PTT on the cancer cell. It evaluates the heat distribution on the cancer cell and surrounding tissue. It also shows the impact of temperature increase, pressure, and wall

shear stress on the cancer cell. Furthermore, the healthy tissue near the cancer cell is also investigated.

Methods

Model Development and Computational Mesh Generation

A cancer cell and healthy tissues around the cancer cell are modeled and meshed as a 3D model using ICEM CFD, a preprocessing program of Ansys Fluent 18.2 (a finite volume approach). A cancer cell is defined as a spherical particle (to reduce the computational time) with a 5 mm diameter, which is not the actual size, and the healthy tissue around the cancer cell is modeled as a cube with a side length of 10 cm. One assumption is that the cancer cell is modeled as a spherical particle. This assumption is made to simplify the simulations. This assumption might not change the results since this study considers the heat effect on the cancer cell, and the cancer cell shape is not very different than the spherical shape. Another assumption, which is using water near the cancer cell, does not change the results since the density and viscosity values of plasma are close to the values of water. The cancer cell size is not the actual size of a human cell, but this work focuses on the heat transfer from the cancer cell and the changes in the healthy tissues around the cancer cell during photothermal therapy. The illustration of the 3D model of the cancer cell and the healthy tissue is represented in Figure 1.

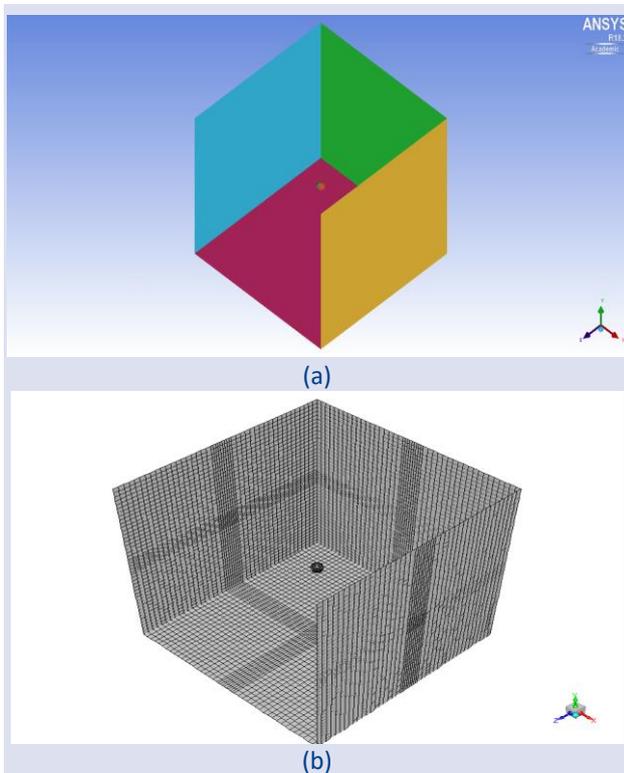


Figure 1. (a) 3D computational model, (b) Meshing of the model. In both figures, the cancer cell is shown as a spherical particle in the middle, and the healthy tissue around the cancer cell is shown as a cube around the sphere.

In Figure 1, the top and one side of the box are not displayed to increase the visibility of the cancer cell inside the box. Once geometric shapes have been produced, meshing proceeded with the formation of hexahedral elements throughout the entire geometry. After the meshes are constructed, the flow geometries are imported into Fluent to solve the incompressible Navier-Stokes equations. Grid independence of the models is tested by refining the temperature at multiple cross-sectional cuts between a more and less refined simulation solution was <3%. The model has approximately 105 524 cells and 112 520 nodes. Grid-independent analysis for temperature is shown in Figure 2.

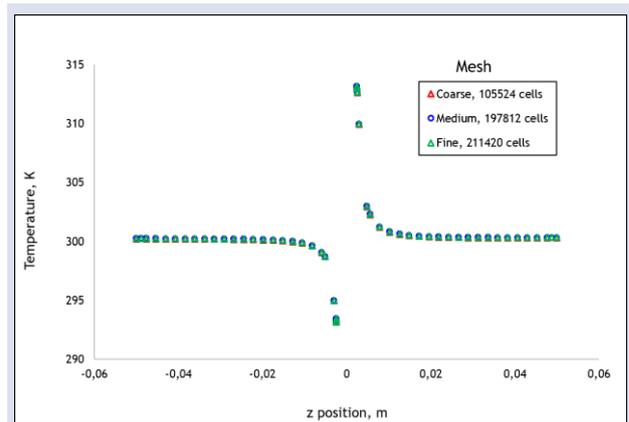


Figure 2. Grid independence analysis for the model.

It can be seen from Figure 2 that temperature values are independent of different mesh sizes. Since different grid sizes are independent of the results, the medium mesh size with 197812 cells is chosen for all simulations for modeling the cancer cell.

Flow Simulation

The finite volume-based Fluent simulator is used to model the cancer cells and healthy tissues and solve the governing equations in the flow field. The specified boundary conditions for the simulations include; the wall boundary conditions for the cancer cell and wall boundary conditions for the sides of the cube, which are representative of the healthy cell, with no-slip boundary conditions on the walls. In Ansys Fluent, the second-order upwind scheme is applied to discretize the momentum and the energy equations, with the SIMPLE algorithm utilized for pressure-velocity coupling. The time parameters for the simulations are specified as the steady state condition. In the case of heat required in photothermal therapy, there is one infrared heat source. Heating the cancer cell is represented by modeling the cell as hot on one side and colder on the other. Different temperatures are specified for the hot side of the cancer cell to define the more prolonged heating and, therefore, to show better the effect of heat transfer to healthy tissues around the sick cell. The current temperatures used in the model are 20°C for the cold part of the cancer cell and 40°C for the hot side. Different variations of the temperatures are also tried, such as the hot part of the cell being increased up to 80°C to have a bigger difference

between the cold and the hot part. Since similar results are obtained, only 20°C for the cold side and 40°C for the hot side of the cell configuration are presented here for simplicity. The simplified model setup schematics are shown in Figure 3.

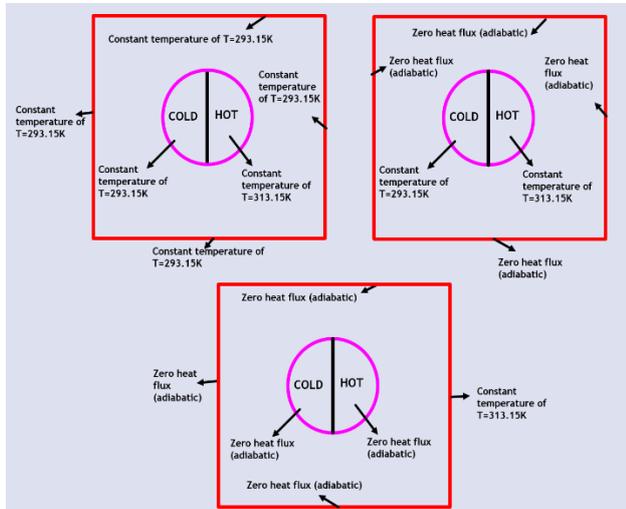


Figure 3. The simplified model setup schematics of the cancer cell and the healthy tissue around it.

The model parameters for the plasma of healthy tissue, such as density, viscosity, specific heat, and

thermal conductivity, are specified depending on temperature. For simplicity, the healthy tissue plasma is considered water, and each particular parameter is defined for a specific equation, which depends on temperature.

Model parameters for the flow simulation

The density, specific heat, thermal conductivity, and dynamic viscosity data of water at different temperatures are taken from the literature [19]–[23]. Regression analysis determines the best fit for these parameters as a temperature function. The resulting formulas are shown in Equations 1 to 4, and the regression plots are represented in Figure 4.

$$\rho = -0.0046T^2 + 2.5183T + 656.86 \tag{1}$$

$$c_p = -0.0005T^3 + 0.5059T^2 - 160.9T + 21204 \tag{2}$$

$$k = -9E - 06T^2 + 0.0068T - 0.6455 \tag{3}$$

$$\mu = 0.3664T^2 - 242.85T + 40733 \tag{4}$$

where ρ is density, c_p is the specific heat, k is the thermal conductivity, μ is the dynamic viscosity, and T is the temperature.

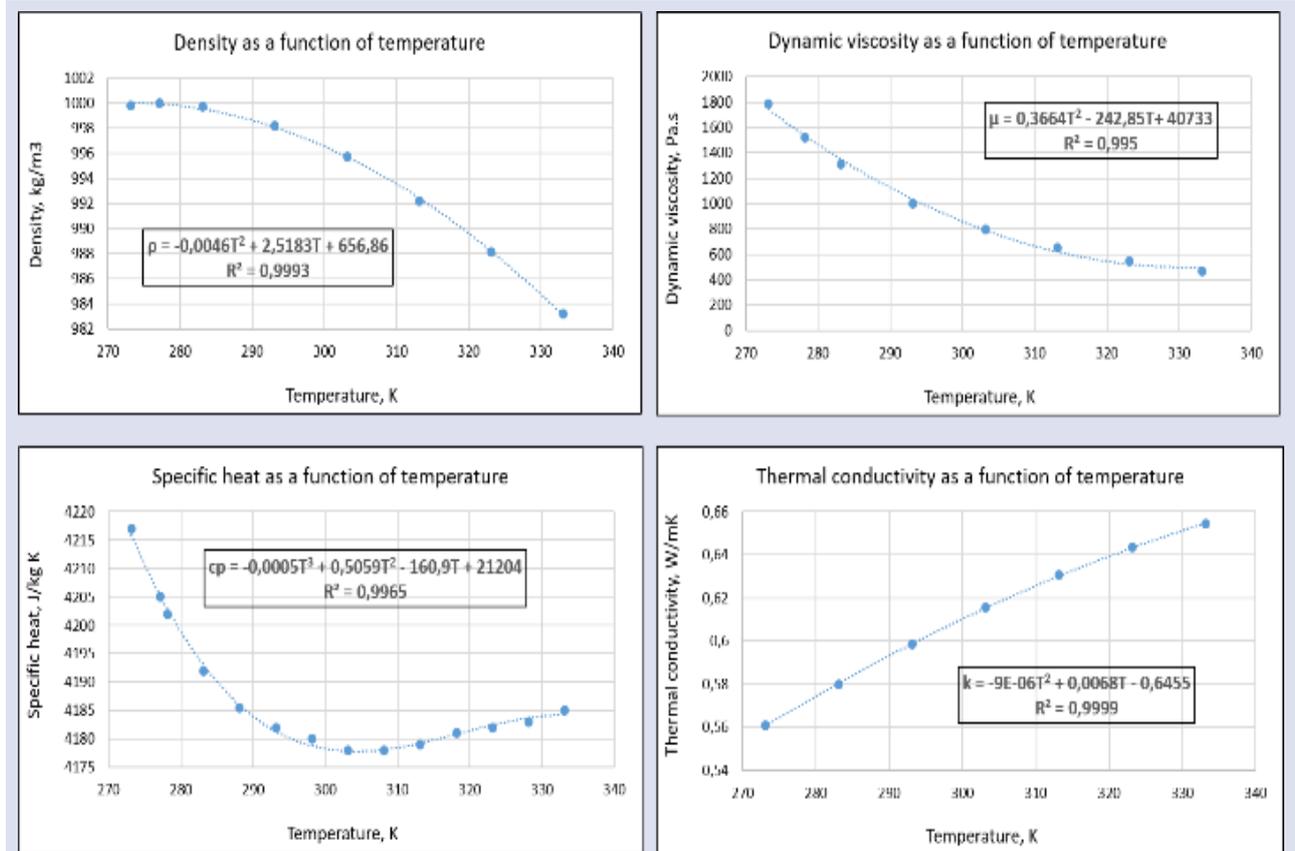


Figure 4. Top left: The density of water as a function of temperature; top right: The dynamic viscosity of water as a function of temperature; bottom left: Specific heat of the water as a function of temperature; and the bottom right: the thermal conductivity of water as a function of temperature.

Results and Discussion

Effect of Temperature on Cancer Cells

Photothermal therapy is represented when the cancer cell is heated on one side and cooler on the other. The highest temperature is 312 K, while the coldest temperature is 294 K. Figure 5 shows the temperature contours of the cancer cell.

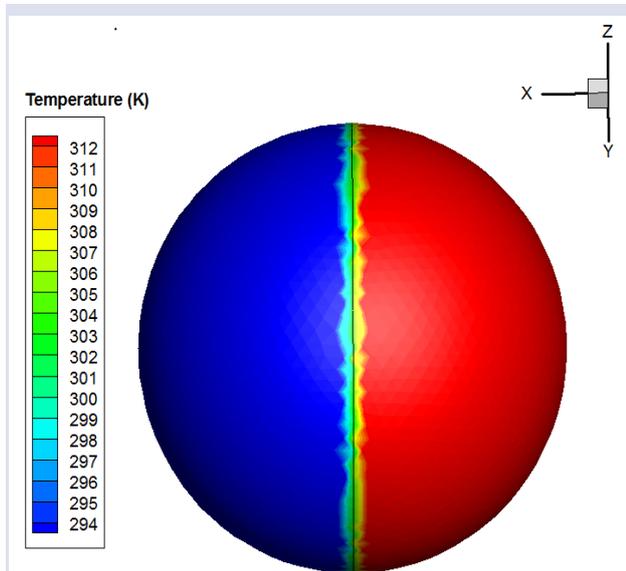


Figure 5. Temperature contours on the cancer cell.

As shown in Figure 5, the side of the cancer cell that is heated by an infrared heat source has a high temperature. After the heating effect on cancer cells, it is important to observe their behavior. The error analysis shows that the model shows adequate correctness, which can also be seen in the R^2 values in Figure 4. The determined R^2 values are 0.9993 for the density equation, 0.995 for the dynamic viscosity equation, 0.9965 for the specific heat distribution, and 0.9999 for the thermal conductivity distribution. Therefore, more parameters of the cancer cell are observed. The change of the wall shear stress on the cancer cell is shown in Figure 6.

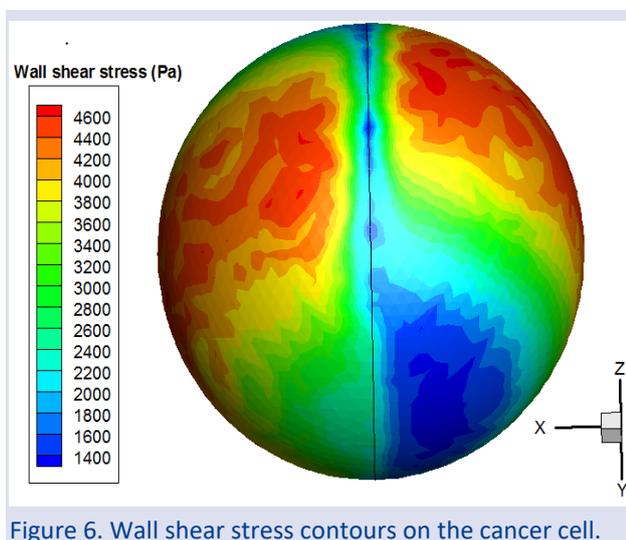


Figure 6. Wall shear stress contours on the cancer cell.

Figure 6 shows the wall shear stress values distribution on the cancer cell. Wall shear stress is high, up to 4600 Pa in the top parts of the cell. The value of the wall shear stress is lower in the lower parts. The lowest value is 1400 Pa. Pressure values are also determined on the cancer cell and are illustrated in Figure 7.

It is essential to analyze the wall shear stress profile on a cancer cell because it will affect the structure and the behavior of cancer cells to the endothelial cells [24]. The wall shear stress has been found to be a key component for a cancer cell and the endothelial cell in the vicinity of the cancer cell in the critical process in the metastatic scenario for affecting the cancer cell stiffness and adhesion of cancer cells to microvasculature [24].

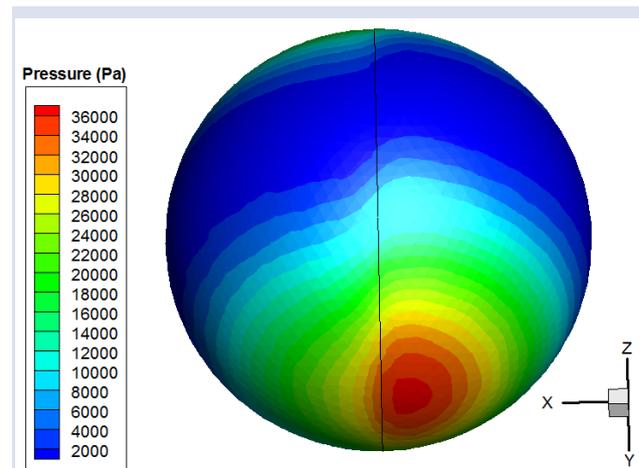


Figure 7. Pressure contours on the cancer cell.

Figure 7 shows the distribution of the pressure values stress on the cancer cell. The highest pressure on the cancer cell goes up to 36000 Pa in the lower parts of the cell. The pressure value is lower on the top parts of the cell. The lowest value is 2000 Pa.

High-pressure values on cancer cells could stimulate their proliferation by affecting their stiffness compared to surrounding cells [25]–[28]. Pressure changes in cancer cells have been found to affect their peripheries, perfusion, and the delivery of chemotherapy. The high-pressure values will also influence the movement of cancer cells, if there are any [25]–[28].

Effect of Temperature on Healthy Tissues in the Vicinity of Cancer Cells

As discussed above, the healthy tissue near the cancer is modeled using a box around the cancer cell. In addition, the heat change around the cancer cell is scrutinized by adding more planes inside the box; some are very close to the cancer cell, and some are attached to the cancer cell. Also, several lines cut through the cancer cell from one side of the box to the other are created for a more detailed analysis. The plane analyzed here cuts the cancer cell at the center, which is presented in Figure 8A. The illustration of the plane attached to the cancer cell and the line are illustrated in Figure 8.

As discussed in part 3.1, the pressure effect in the cancer cell and its vicinity is essential for changing its behavior [25]–[28]. As shown in Figure 11, the pressure values of healthy tissue around the cancer cell increase away from the cancer cell. Providing a quantitative validation for the CFD model is currently not available because of the deficiency of the experimental data.

Conclusion

For a deadly illness, there have been alternative therapies for cancer treatment, such as photothermal therapy. However, it is necessary to know that the heat distribution is suitable and that the surrounding tissue is not overheated. This work models the cancer cells and the surrounding healthy tissue using CFD. The temperature distribution of cancer cells and the healthy tissue is examined. Current findings show that temperature changes occur very near the cancer cell. Moreover, the cancer cell and the healthy tissue pressure are also affected by the PTT of cancer cells. The wall shear stress profile also affects the stiffness and adhesion of cancer cells to the microvasculature. Overall, this study concludes that the cancer cell and the tissue around it can be modeled effectively by using CFD. It has been shown that when photothermal therapy is used, the healthy tissue around the cancer cell is unaffected, demonstrating photothermal therapy's effectiveness. The results of this work could guide researchers in optimizing the photothermal therapy of cancer cells, and the modeling approach could be applied to investigate alternative therapies.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Ferlay J., Ervik M., Lam F., Colombet M., Mery L., Pineros M., Global Cancer Observatory: Cancer Today. Lyon: International Agency for Research on Cancer. <https://gco.iarc.fr/today>. Retrieved July, 2024.
- [2] Gong F., Liu J., Yang J., Qin J., et al., Effective Thermal Transport Properties in Multiphase Biological Systems Containing Carbon Nanomaterials, *RSC Adv.*, 7(22) (2017) 13615–13622.
- [3] Cherukula K., Manickavasagam Lekshmi K., Uthaman S., Cho K., Cho C-S., Park I-K., Multifunctional Inorganic Nanoparticles: Recent Progress in Thermal Therapy and Imaging, *Nanomater*, 6(4) (2016) 76.
- [4] Yang K., Feng L., and Liu Z., Stimuli responsive drug delivery systems based on nano-graphene for cancer therapy, *Adv. Drug Deliv. Rev.*, 105 (2016) 228–241.
- [5] Cherukuri P., Glazer E.S., Curley S.A., Targeted hyperthermia using metal nanoparticles, *Adv. Drug Deliv. Rev.*, 62(3) (2010) 339–345.
- [6] Manthe R. L., Foy S. P., Krishnamurthy N., Sharma B., Labhasetwar V., Tumor Ablation and Nanotechnology, *Mol. Pharm.*, 7(6) (2010) 1880–1898.
- [7] Lal S., Clare S.E., Halas N.J., Nanoshell-Enabled Photothermal Cancer Therapy: Impending Clinical Impact, *Acc. Chem. Res.*, 41(12) (2008) 1842–1851.
- [8] Robinson J.T., Tabakman S.M., Liang Y., Wang H., et al., Ultrasmall Reduced Graphene Oxide with High Near-Infrared Absorbance For Photothermal Therapy, *J. Am. Chem. Soc.*, 133(17) (2011) 6825–6831.
- [9] Liu H., Chen D., Li L., Liu T., et al., Multifunctional Gold Nanoshells On Silica Nanorattles: A Platform for the Combination Of Photothermal Therapy And Chemotherapy With Low Systemic Toxicity, *Angew. Chem. Int. Ed. Engl.*, 50(4) (2011) 891–895.
- [10] Cheng L., Yang K., Chen Q., Liu Z., Organic stealth nanoparticles for highly effective in vivo near-infrared photothermal therapy of cancer, *ACS Nano*, 6(6) (2012) 5605–5613.
- [11] van der Zee J., Heating the Patient: A Promising Approach?, *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, 13(8) (2002) 1173–1184.
- [12] Dahl O., Interaction of Heat and Drugs in Vitro and in Vivo. In: Seegenschmiedt M.H., Fessenden P., Vernon C.C., (Eds.). *Thermoradiotherapy and Thermochemotherapy: Biology, Physiology, Physics*. Springer, Berlin Heidelberg, (1995) 103–121.
- [13] Gonzalez D.G., van Dijk J.D.P., Blank L.E.C.M., Rümke P., Combined Treatment With Radiation And Hyperthermia Inmetastatic Malignant Melanoma, *Radiother. Oncol.*, 6(2) (1986) 105–113.
- [14] Kim J. H., Hahn E. Tokita W., N., Nisce L. Z., Local Tumor Hyperthermia in Combination with Radiation Therapy. 1. Malignant Cutaneous Lesions, *Cancer*, 40(1) (1977) 161–169.
- [15] Coleman A., Augustine C. K., Beasley G., Sanders G., Tyler D., Optimizing Regional Infusion Treatment Strategies for Melanoma of the Extremities, *Expert Rev. Anticancer Ther.*, 9(11) (2009) 1599–1609.
- [16] Kamarudin S., Taib I., Adnan M., Nasir F., et al., Prediction of Heat Distribution on Brain Malignant Tumor Using Hyperthermia Therapy, *AIP Conf. Proc.*, 2955(1) (2023). 20031.
- [17] Haar L., Gallagher J.S., Kell G.S., NBS/NRC steam tables: thermodynamic and transport properties and computer programs for vapor and liquid states of water in SI units. Washington (DC) Hemisphere, (1984).
- [18] Marsh K.N., Recommended reference materials for the realization of physicochemical properties. Blackwell Scientific Publications, (1987).
- [19] Sengers J.V., Watson J.T.R., Improved International Formulations for the Viscosity and Thermal Conductivity of Water Substance, *J. Phys. Chem. Ref. Data*, 15(4) (1986) 1291–1314.
- [20] Archer D.G., Wang P., The Dielectric Constant of Water and Debye-Hückel Limiting Law Slopes, *J. Phys. Chem. Ref. Data*, 19(2) (1990) 371–411.
- [21] Vargaftik N. B., Volkov B. N., Voljak L. D., International Tables of the Surface Tension of Water, *J. Phys. Chem. Ref. Data*, 12(3) (1983) 817–820.
- [22] Dabagh M., Randles A., Role of Deformable Cancer Cells on Wall Shear Stress-Associated-VEGF Secretion by Endothelium in Microvasculature, *PLoS One*, 14(2) (2019) e0211418.
- [23] Basson M.D., Zeng B., Downey C., Sirivelu M.P., Tepe J.J., Increased Extracellular Pressure Stimulates Tumor Proliferation by A Mechanosensitive Calcium Channel and PKC- β , *Mol. Oncol.*, 9(2) (2015) 513–526.
- [24] Ingber D.E., Can Cancer Be Reversed by Engineering the Tumor Microenvironment?, *Semin. Cancer Biol.*, 18(5) (2008) 356–364.
- [25] Gutmann R., Leunig M., Feyh J., Goeat A.E., et al., Interstitial Hypertension in Head and Neck Tumors in Patients: Correlation with Tumor Size, *Cancer Res.*, 52(7) (1992) 1993–1995.
- [26] Raju B., Haug S. R., Ibrahim S. O., Heyeraas K. J., High Interstitial Fluid Pressure in Rat Tongue Cancer is Related to Increased Lymph Vessel Area, Tumor Size, Invasiveness and Decreased Body Weight, *J. Oral Pathol. Med.*, 37(3) (2008) 137–144

In Vitro Tyrosinase and Collagenase Inhibitory and Antioxidant Potential of *Smyrniium rotundifolium* Mill. and *Euphorbia virgata* Waldst.&Kit. from Türkiye

Eda Sönmez Gürer ^{1,a,*}, Ayşe Esra Karadağ ^{2,b}, Rengin Baydar ^{2,c}, Melek Demirel ^{3,d}

¹ Department of Pharmacognosy, Faculty of Pharmacy, Sivas Cumhuriyet University, 58140, Sivas / Türkiye

² Department of Pharmacognosy, School of Pharmacy, Istanbul Medipol University, 34810, Istanbul / Türkiye

³ Department of Horticulture, Faculty of Agriculture, Kocaeli University, 41380, Kocaeli / Türkiye

*Corresponding author

Research Article

History

Received: 19/10/2024

Accepted: 19/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Antioxidants help prevent signs of aging and skin tone inequalities by protecting our skin from free radicals. High tyrosinase inhibition has a whitening effect on the skin, while collagenase inhibition has an anti-sagging effect on the skin. Antioxidant activity and tyrosinase/collagenase enzyme inhibition capacities have mutually supporting effects. The aim of this study was to determine the antioxidant activities, tyrosinase and collagenase inhibitory potentials of ethanol extracts of two medicinal plants from Turkey (*Smyrniium rotundifolium* and *Euphorbia virgata*). In the study, 6 different reference substances and their chemical contents were investigated. Myricetin, quercetin and kaempferol were observed in *S. rotundifolium* extract, and quercetin was observed in *E. virgata* extract. According to the antioxidant capacity results measured by both analyses; *S. rotundifolium* (IC₅₀ DPPH*: 4.9±0.15 µg/mL, IC₅₀ ABTS**: 4.3±0.2 µg/mL) and *E. virgata* (IC₅₀ DPPH*: 4.6±0.11 µg/mL, IC₅₀ ABTS**: 4.1±0.13 µg/mL) extracts were observed to have antioxidant capacities similar to each other. It was determined that *S. rotundifolium* had higher anti-collagenase (27.9±0.13% inhibition) and anti-tyrosinase (11.1±0.14% inhibition) activities compared to *E. virgata* extract. These results showed us that *S. rotundifolium* can be considered as a strong candidate for the management of epidermal hyperpigmentation and skin elasticity and deserves further study.

Keywords: Antioxidant, Collagenase, *Euphorbia virgata*, *Smyrniium rotundifolium*, Tyrosinase.

^a edagurer@cumhuriyet.edu.tr

^b <https://orcid.org/0000-0003-0319-6312>

^c aeguler@medipol.edu.tr

^d <https://orcid.org/0000-0002-3412-0807>

^c rengin.baydar@medipol.edu.tr

^d <https://orcid.org/0009-0008-3754-0293>

^d melek.demirel@kocaeli.edu.tr

^d <https://orcid.org/0000-0002-8477-1122>

Introduction

Enzymes can show their effects outside their natural environment when suitable conditions are provided, such as laboratory or artificial environments. In this way, enzymes and enzyme inhibitors are used in many areas, especially in medicine, food and cosmetics industry, pesticide production and agriculture industry [1]. Tyrosinase enzyme is a copper-containing enzyme that catalyzes the oxidation of monophenols and is also involved in melanin biosynthesis. It is possible to find melanin pigment in many living things (animals, plants, fungi, bacteria) in nature, especially animals [2]. In humans, melanin plays a very important role in both absorbing free radicals and protecting the organism from harmful UV rays [3]. It is known that hereditary and environmental factors are among the main causes of skin color problems. These dermatological events include melasma, solar melanosis, ephelides, senile lentigos and acne scars [4]. Tyrosinase inhibitors are frequently used in the skin care sector, especially as a lightener for blemished skin. Skin products containing synthetic tyrosinase inhibitors have adverse effects such as redness, itching and, in the future, skin cancer [5]. Tyrosinase inhibitors have strong antioxidant activity. Because plants that receive high amounts of sunlight activate oxygen radicals to protect themselves from the sun's harmful rays.

Therefore, plants need to have strong antioxidant effects to protect themselves from damage [6]. Due to health concerns regarding synthetic molecules, there has been a resurgence of interest in natural products in the development of new compounds for the treatment of hyperpigmentation. Discovery of new phytochemical compounds with similar biological effects, enzyme inactivation and antioxidant activity is desirable [3].

Collagenases are proteolytic enzymes involved in extracellular matrix remodeling and degradation, which are responsible for the physiological processes of organ development and tissue regeneration, and play a direct role in skin aging. Collagenases are also involved in the degradation of many matrix and non-matrix proteins, such as growth factors, which can control cell growth and survival [7,8]. Skin aging is an inevitable biological process in the human body and is not only due to increasing age, but also due to both chronological and photo-aging. Examples of photo-aging include exposure to ultraviolet rays, pollution or nicotine. During this aging process, the epidermis becomes thinner and blood vessels become fewer. It is known that this damage occurs as a result of the degradation of elastic fibers, collagen fibers and hyaluronic acid, and significant degenerative changes in the upper dermal connective tissue. It is thought that this

is due to the increased expression of elastase, collagenase and hyaluronidase produced in fibroblasts and various inflammatory cells [9,10].

Oxidative stress occurs as a result of an imbalance between free radicals and antioxidants in the human body. It is known that this imbalance has important consequences on human health. The accumulation of free radicals in the body and their high levels due to uncontrolled production damage cellular components such as lipids, proteins and DNA and can eventually lead to neuronal dysfunction and death. According to the results of experimental studies, it has been observed that oxidative stress accelerates the progression of cancer, cardiovascular system diseases, neurological diseases and many diseases related to aging [7,11–13].

The *Smyrniun* genus, which belongs to the Apiaceae family and has 38 species in the world distribution, is represented by 6 taxa in Turkey, including *S. rotundifolium*. It is a biennial herbaceous plant that grows in stony places, bushes and forest edges [14]. The roots of *Smyrniun* taxa are often used in the treatment of different illnesses as diuretic, depurative and laxative. The fruit has carminative and stomachic effects [15].

The *Euphorbia* genus is the genus with the most species in the Euphorbiaceae family. *Euphorbia* species contain latex. This genus, which has nearly 2000 species, is usually seen in Africa or Madagascar. There are 91 *Euphorbia* species growing in Turkey. It has been known for many years that it has been used externally to heal wounds and skin diseases such as warts and eczema, and internally to cure migraines and intestinal parasites [16,17]. Studies have shown that *Euphorbia* species have cytotoxic, antitumor, antibacterial, anti-inflammatory and anti-HIV activities [18]. Despite the high vitamin C content of *Smyrium* species [14], the lack of evidence for their use in skin diseases among the public and the use of the aboveground parts of *Euphorbia* species in eczema and warts among the public [16,19] led us to examine these plants in terms of dermatological activity.

The purpose of our study was to determine the antioxidant activities and tyrosinase and collagenase enzyme inhibition potentials of ethanol extracts of two medicinal plants from Turkey (*Smyrniun rotundifolium* and *Euphorbia virgata*).

Materials and Methods

Plant Materials

S. rotundifolium and *E. virgata* plants used in the study were collected from the area at coordinates 39°41'40"N, 37°02'25"E, at an altitude of 1400 m from height, within the borders of Sivas İmaret Village between June and July 2023. The taxonomic description of the collected samples was made by Anadolu University Faculty of Pharmacy Faculty Member Professor Yavuz Bülent KÖSE. The collected plant samples were preserved in the Anadolu University Faculty of Pharmacy Herbarium as *S. rotundifolium* (ESSE: 16196), *E. virgata* (ESSE: 16197).

Preparation of Plant Extract

In our study, plant extracts prepared from the aerial parts of *S. rotundifolium* and *E. virgata* were used. The samples were first washed with tap water and then with pure water to prevent possible contamination, then dried on blotting papers, ground in a grinder and 100 grams were taken and 1000 ml of ethanol was added. It was kept at room temperature in a shaker at 150 rpm for 24 hours. At the end of the extraction process, the extract was filtered through filter paper, and then the solvent was removed in a rotary evaporator at 40°C. The obtained extracts was placed in a dark glass bottle and stored at -20 °C to be used in experimental procedures [20].

High Pressure Liquid Chromatography (HPLC)

HPLC analysis was performed using an Agilent 1100 HPLC system equipped with a UV-DAD detector. Extracts of *S. rotundifolium* and *E. virgata* were prepared in ethanol at a concentration of 10 mg/mL and filtered through a 0.22 µm membrane filter. Phenolic reference compounds, including quercetin, myricetin, gallic acid, kaempferol, rosmarinic acid, and apigenin, were dissolved in methanol prior to HPLC analysis. A C18 column (250 x 4.6 mm, 5 µm) was utilized. The mobile phases consisted of Phase A (acetonitrile: distilled water: formic acid (10:89:1, v/v)) and Phase B (acetonitrile: distilled water: formic acid (89:10:1, v/v)), which were applied using a gradient elution. Mobile Phase B was varied from 15% to 100% over a 40-minute analysis period. The flow rate was maintained at 1.0 mL/min, and all samples were injected in triplicate. Detection was carried out at 330 nm, with an injection volume of 20 µL and a column temperature set at 40°C [21].

Antioxidant Capacity

Scavenging of DPPH free radical (2,2-Diphenyl-1-picrylhydrazyl) test

The total antioxidant capacity of the ethanol extracts of *S. rotundifolium* and *E. virgata* were determined using the DPPH free radical (DPPH[•]) scavenging method, as described by Blois and colleagues [22]. In this method, the reaction mixture consisted of 100 µM DPPH[•] prepared in methanol, along with various concentrations of the test formulations. After an incubation period of 30 minutes at a controlled temperature of 25 ± 2 °C, the absorbance of the mixture was measured at a wavelength of 517 nm using a UV spectrophotometer (UV-1800, Shimadzu, Japan) to assess the scavenging activity of the extract [23]. Ascorbic acid, a well-known antioxidant, was used as a positive control to compare the efficacy of the extract in neutralizing free radicals. The results of the assay were expressed as IC₅₀ values (µg/mL), which represent the concentration of the extract required to inhibit 50% of the DPPH[•] in the reaction mixture, indicating the antioxidant potency of the sample.

2,2'-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) (ABTS^{•+}) Antioxidant Activity

The antioxidant capacity of the samples was evaluated using the ABTS radical (ABTS^{•+}) cation decolorization assay, following the protocol described by [24]. In this method, a solution of 7 mM ABTS^{•+} was prepared in water

and then mixed with 2.45 mM potassium persulfate to generate the ABTS radical cation. This mixture was allowed to stand in a dark room at 25°C for 16 hours to ensure complete formation of the radical cation before it was used in the assay. Subsequently, ethanol was added to the reaction mixture, and the absorbance was measured at 734 nm at 25°C to determine the extent of decolorization, which indicates the scavenging activity of the antioxidants in the sample. Each measurement was performed in triplicate to ensure accuracy and reproducibility of the results. In this assay, ethanol was used as the negative control, while Trolox served as the positive control to provide a reference for antioxidant activity [23]. The results were expressed as IC₅₀ values (µg/mL).

Tyrosinase Inhibition

The *in vitro* anti-tyrosinase activity of the ethanol extracts of *S. rotundifolium* and *E. virgata* (20 µg/mL) was assessed following the method outlined by Hearing and Jimenez [25]. This method first involves evaluating the ability of the compounds to inhibit the diphenolase activity of tyrosinase, with L-DOPA serving as the substrate. Tyrosinase from mushroom origin (E.C. 1.14.18.1) (30 U, 28 nM) was dissolved in a sodium phosphate buffer (pH 6.8, 50 mM), and the test compounds were added to the solution, followed by a pre-incubation period of 10 minutes at room temperature. The enzymatic reaction was then initiated by adding 0.5 mM L-DOPA to the mixture, and the change in absorbance at 475 nm was monitored at 37°C. Kojic acid was used as a positive control to benchmark the inhibitory effect of the extract [26].

Collagenase Inhibition

For the evaluation of anti-collagenase activity, the ethanol extracts of *S. rotundifolium* and *E. virgata* (20 µg/mL) was initially dissolved in ethanol. The enzyme

inhibition assay was conducted according to the manufacturer's instructions provided in the "Collagenase Activity Assay Kit (Colorimetric, Abcam 196999)." The inhibition of collagenase by the test substances was measured kinetically using a multi-mode microplate reader (SpectraMax i3) at 345 nm and 37°C. The degree of enzyme inhibition by the extract was calculated by comparing it against the standards included in the kit. The results were expressed as the percentage of inhibition, and the mean inhibition values were calculated for the samples. All assays were conducted in duplicate to ensure reliability and reproducibility of the data [26].

Results and Discussion

High Pressure Liquid Chromatography (HPLC)

Six different reference compounds which includes quercetin, myricetin, gallic acid, kaempferol, rosmarinic acid, and apigenin, were used for phytochemical analysis of *S. rotundifolium* and *E. virgata*. As a result, Table 1 represents the retentions times of the standard compounds based on the related HPLC chromatogram.

Table 1. Retention times of the standard compounds

| Reference compounds | t _R |
|---------------------|----------------|
| Gallic acid | 3.748 |
| Rosmarinic acid | 9.576 |
| Myricetin | 10.425 |
| Quercetin | 14.324 |
| Apigenin | 16.998 |
| Kaempferol | 18.342 |

According to the HPLC chromatogram of *S. rotundifolium* (Figure 1), myricetin, quercetin and kaempferol standards were present in the extract.

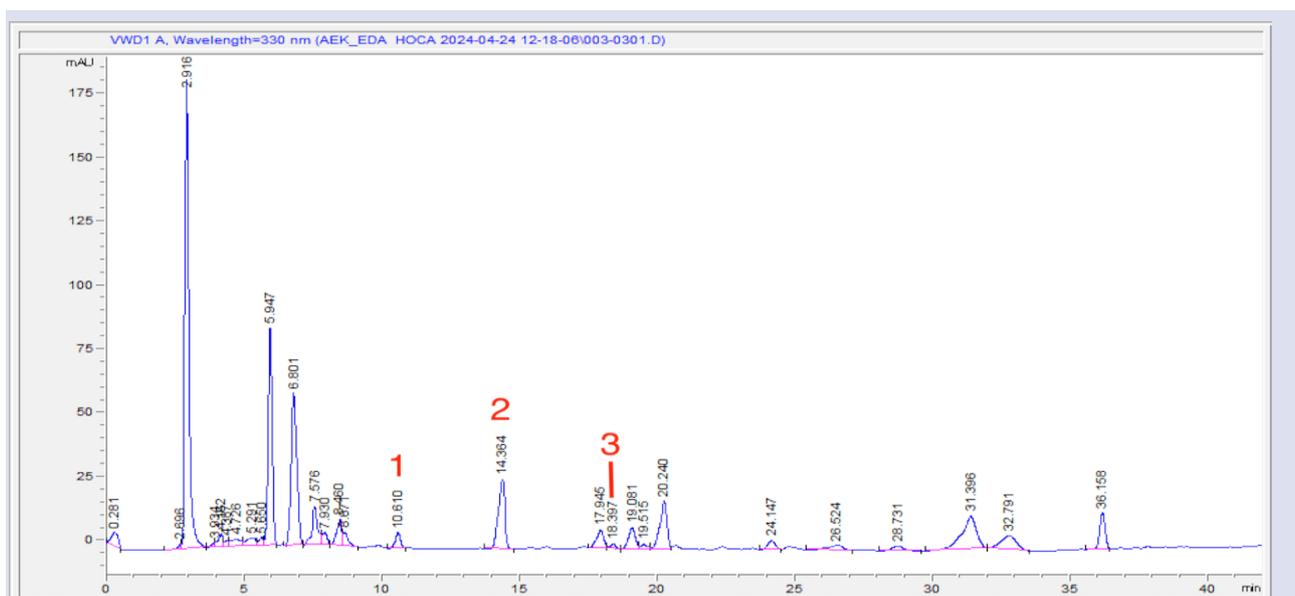


Figure 1. HPLC chromatogram of *S. rotundifolium* ethanol extract (1. myricetin t_R: 10.610, 2. quercetin t_R: 14.364, 3. kaempferol t_R: 18.397)

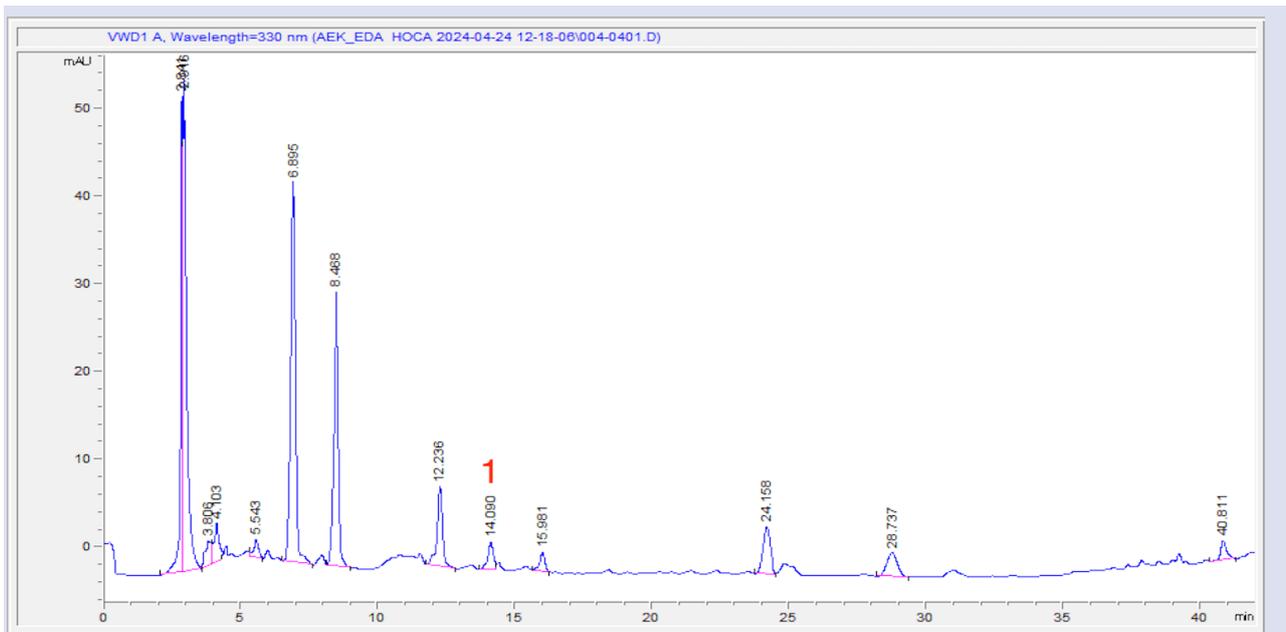


Figure 2. HPLC chromatogram of *E. virgata* ethanol extract (1. quercetin tR: 14.090)

According to the HPLC chromatogram of *E. virgata* (Figure 2), the ethanolic extract was found to contain only quercetin among the standards analyzed.

In 1984 and 1985, Goren et al. examined the phytochemistry of *S. rotundifolium* and isolated sesquiterpene lactones such as α -selinene, germacrone, and furodiene from the fruits. They also isolated oxepine derivatives such as smyrnicordioidide, and isosmyrnicordioidide from the roots of the *S. rotundifolium* [27,28].

Previous studies on *Euphorbia* genus showed that they contain phenolic compounds and flavonoids such as myricetin, rutin, kaempferol, and quercetin [29]. Our HPLC results supported the literature.

Antioxidant Capacity

ABTS^{•+} antioxidant activity tests

The antioxidant capacity of the ethanolic extracts of *S. rotundifolium* and *E. virgata* was investigated using DPPH[•] and ABTS^{•+} antioxidant assays. Based on the results of these analyses, *S. rotundifolium* and *E. virgata* extracts demonstrated similar antioxidant capacities as measured by both assays. The *S. rotundifolium* extract exhibited antioxidant capacity with the IC₅₀ values of 4.9±0.15 µg/mL and 4.3±0.2 µg/mL for the DPPH[•] and ABTS^{•+} assays, respectively. Similarly, the *E. virgata* extract showed antioxidant capacity with the IC₅₀ values of 4.6±0.11 µg/mL and 4.1±0.13 µg/mL for the DPPH[•] and ABTS^{•+} assays. Detailed results were given in Table 2.

Table 2. DPPH[•] and ABTS^{•+} assay results of the *S. rotundifolium* and *E. virgata* ethanolic extracts

| | <i>S. rotundifolium</i> | <i>E. virgata</i> | Positive control |
|--------------------|--------------------------|-------------------|---------------------------------|
| | IC ₅₀ (µg/mL) | | |
| DPPH [•] | 4.9±0.15 | 4.6±0.11 | 0.002±0.0001 (Ascorbic acid) |
| ABTS ^{•+} | 4.3±0.2 | 4.1±0.13 | 0.01±0.0001 (Trolox) |

The antioxidant capacity of both extracts may be related to the phenolic compounds they had according to the HPLC analysis results.

A previous study reported the antioxidant activities of 47 different Mediterranean plant species extracts, which one of them was the leaves of *S. rotundifolium*. According to the results, *S. rotundifolium* extract exhibited a low antioxidant capacity compared to other extracts with the value of 152.4 µg ascorbic acid equivalent per 1 g of extract (total water-soluble antioxidant activity) [30].

Tyrosinase and Collagenase Inhibition

Tyrosinase and collagenase enzyme inhibition capacities of the *S. rotundifolium* and *E. virgata* ethanolic extracts were investigated. According to the results, *S. rotundifolium* extract showed 11.1±0.14% inhibition on tyrosinase enzyme, and 27.9±0.13% inhibition on collagenase enzyme. On the other hand, *E. virgata* extract exhibited 8.9±0.18% inhibition on tyrosinase enzyme, and 8.1±0.1% inhibition on collagenase enzyme. Based on the results, *S. rotundifolium* was found to have higher anti-collagenase and anti-tyrosinase activity compared to *E. virgata* extract. Kojic acid and quercetin were used as positive standards for anti-tyrosinase activity and anti-collagenase activity, respectively (Table 3).

Table 3. Anti-tyrosinase and anti-collagenase results of the *S. rotundifolium* and *E. virgata* ethanolic extracts

| | <i>S. rotundifolium</i> | <i>E. virgata</i> | Positive control |
|-------------------|-------------------------|-------------------|----------------------------|
| | % inhibition | | |
| ATA [*] | 11.1±0.14 | 8.9±0.18 | 98.9±0.002 (Kojic acid) |
| ACA ^{**} | 27.9±0.13 | 8.1±0.1 | 91.4±0.007 (Quercetin) |

*Anti-tyrosinase activity

** Anti-collagenase activity

Although the medicinal potential of plants in Turkey is important, the knowledge in this area and the studies on these plants are limited. Since antioxidant and enzyme inhibition studies on *S. rotundifolium* and *E. virgata* species are quite limited, studies on different species are also included.

Flavonoids are secondary metabolites that exhibit pharmacological activities in many areas including antioxidant, antimicrobial, spasmolytic, diuretic, capillary protection, cytostatic, etc. The source of most of these activities is quercetin, the most important flavonol known. In addition to quercetin, kaempferol and myricetin are also common flavonoids in plants responsible for the activities. If we look at our HPLC results, we can see that our plants contain myricetin, rutin, kaempferol and quercetin, which are responsible for antioxidant activity [31].

In a study, antioxidant and antimicrobial activities of 6 different *Smyrniium* were investigated. The antioxidant properties of methanol extracts were examined by DPPH[•] scavenging method. *S. olusatrum* indicated the strongest radical scavenging activity at a concentration of 1.08 mg/mL (96.75 ± 0.47%). The DPPH[•] scavenging activity of the extract belonging to the *S. rotundifolium* species was observed as 96.15 ± 0.00% at a concentration of 0.99 mg/mL [32]. In a study conducted on several *Smyrniium* species, it was determined that *S. olusatrum* had low inhibitory potential against acetylcholinesterase, butyrylcholinesterase and tyrosinase enzymes [33]. In a different study, it was determined that the methanol extract of *S. cordifolium* had good tyrosinase inhibition potential (137.54 mg kojic acid equivalent/g extract) [3].

In a recent study conducted with the species *S. connatum*, acetylcholinesterase, butylcholinesterase, tyrosinase inhibition and antioxidant activities of the prepared plant extracts were investigated. It was observed that the aerial parts of *S. connatum* were generally more active in terms of antioxidant activity determination. The aerial parts and roots of *S. connatum* showed high inhibitory activity against both cholinesterase enzymes. All extracts showed moderate inhibitory activity against tyrosinase (54.87% and 44.31%). As a result of the findings, it was determined that the biological activity of the aerial parts of *S. connatum* was generally more active than the roots of *S. connatum* [15]. In a study conducted with leaf extracts of the species *S. rotundifolium*, the total phenol content was found to be 157.3 mg GAE/g extract and exhibited low antioxidant activity [30]. In a different study investigating its chemical content, kaempferol, kaempferol 3-β-D-galactoside, kaempferol 3-methyl ether 7-β-D-glucoside and kaempferol 3-diglucoside were found in its content [34].

Since no research on antioxidant and enzyme activity was found in our literature search for the *E. virgata* species, our study is the first in this sense. In a previous study; quercetin-3-O-glucoside, kaempferol, kaempferol-3-O-glucoside, kaempferol-3-rutinoside and rutin flavonoids were isolated from *E. virgata* [35]. In another study conducted to investigate the wound healing effect

of aerial parts of some *Euphorbia* species, hexane, ethyl acetate and methanol solvents were used. The results showed that aerial parts of *E. characias* have wound healing and anti-inflammatory activities in different models [16]. Antimicrobial effects of methanol extracts and latex of some *Euphorbia* species used for medical purposes in Turkey were investigated. The results showed that the extracts of *Euphorbia* species inhibited the growth of the tested microorganisms at different rates. In addition, the MIC (Minimum Inhibitor Concentration) values of the extracts were determined as 31.2-1000 µg [36].

It has been reported that quercetin, kaempferol and ellagic acid isolated from the distillation water formed during the extraction of essential oil from *Rosa damascena* flowers inhibited tyrosinase enzyme 10 times more strongly than the control group kojic acid (56.1 µM) [37]. In the study investigating the tyrosinase enzyme inhibition of quercetin, a flavonoid compound with potential as a skin-lightening agent; it was concluded that quercetin compounds obtained by maceration from *Moringa oleifera* L. leaves could be potential skin-lightening agents [38]. Among the plants in our study, the presence of myricetin, quercetin and kaempferol in *S. rotundifolium* and quercetin in *E. virgata* support our tyrosinase inhibition results.

Conclusion

The cosmetics sector is showing significant growth, especially in skin aging. For this reason, it is known that studies on natural products that prevent aging and pigmentation are popular research areas. Our work; it was targeted to define whether it can be used in dermatological diseases caused by the differentiation of melanin pigment in the skin, in neurological diseases and in the food/cosmetics sector with the anti-tyrosinase activity test; and in the content of topical formulations prepared for anti-wrinkle and skin aging with the anti-collagenase activity test. Since it is known that oxidative stress plays a role in degenerative processes associated with aging, the antioxidant activity of the extracts was also examined in our study.

According to the results, *S. rotundifolium* was found to have higher anti-collagenase and anti-tyrosinase activity compared to *E. virgata* extract. *S. rotundifolium* can be considered as a strong candidate for the development of plant-derived products for the management of hyperpigmentation and skin aging, which deserves further study. To the best of our knowledge, this study can be considered as the first report from Turkey investigating the antioxidant, anti-tyrosinase and anti-collagenase inhibitory effects of *S. rotundifolium* and *E. virgata*.

Conflicts of interest

No conflicts of interest have been declared by the authors.

Acknowledgment

We would like to thank Prof. Dr. Yavuz Bülent KÖSE for identifying the plant.

References

- [1] Yan W., J Jiang L., Zhixiang W.S., Suying T., Comparative Study of Anti-inflammatory Effect of *Rosa laevigata* Root and Stem., *Modern Chinese Medicine*, 3 (2010).
- [2] Bilgin Sökmen B., Yılmazoğlu B., Tirozinaz Enziminin Giresun Yöresinde Yetişen Yenilebilir Kanlıca Mantarından (*Lactarius salmonicolor*) Saflaştırılması ve Karakterizasyonu, *Karadeniz Fen Bilimleri Dergisi*, 8 (2018) 10–23.
- [3] Zengin G., Mahomoodally M.F., Picot-Allain C.M.N., Cakmak Y.S., Uysal S., Aktumsek A., *In Vitro* Tyrosinase Inhibitory and Antioxidant Potential of *Consolida orientalis*, *Onosma isauricum* and *Spartium junceum* from Turkey, *South African Journal of Botany*, 120 (2019) 119–123.
- [4] Trivedi M.K., Yang F.C., Cho B.K., A Review of Laser and Light Therapy in Melasma, *Int. J. of Women's Dermatology*, 3 (2017) 11–20.
- [5] Pillaiyar T., Manickam M., Namasivayam V., Skin Whitening Agents: Medicinal Chemistry Perspective of Tyrosinase Inhibitors, *Journal of Enzyme Inhibition and Medicinal Chemistry*, 32 (2017) 403–425.
- [6] Masuda T., Yamashita D., Takeda Y., Yonemori S., Screening for Tyrosinase Inhibitors Among Extracts of Seashore Plants and Identification of Potent Inhibitors from *Garcinia subelliptica*, *Biosci Biotechnology Biochem.*, 69 (2005) 197–201.
- [7] Silva R.M.G. da, Alves C.P., Barbosa F.C., Santos H.H., Adão K.M., Granero F.O., Antioxidant, Antitumoral, Antimetastatic Effect and Inhibition of Collagenase Enzyme Activity of *Eleutherine bulbosa* (Dayak onion) Extract: *In Vitro*, *In Vivo* and *In Silico* Approaches, *Journal of Ethnopharmacol*, 318 (2024) 117005.
- [8] Freitas-Rodríguez S., Folgueras A.R., López-Otín C., The Role of Matrix Metalloproteinases in Aging: Tissue Remodeling and Beyond, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1863(11) (2017) 2015–2025.
- [9] Lee K.K., Cho J.J., Park E.J., Choi J.D., Anti-Elastase and Anti-Hyaluronidase of Phenolic Substance from *Areca catechu* as a New Anti-Ageing Agent, *Int. J. Cosmet. Sci.*, 23 (2001) 341–346.
- [10] Farage M.A., Miller K.W., Elsner P., Maibach H.I., Characteristics of the Aging Skin, *Adv. Wound Care (New Rochelle)*, 2 (2013) 5–10.
- [11] Pilipović K., Jurišić Grubešić R., Dolenc P., Kučić N., Juretić L., Mršić-Pelčić J., Plant-Based Antioxidants for Prevention and Treatment of Neurodegenerative Diseases: Phytotherapeutic Potential of *Laurus nobilis*, *Aronia melanocarpa*, and *Celastrol*, *Antioxidants*, 12 (2023).
- [12] Negreanu-Pirjol B.S., Oprea O.C., Negreanu-Pirjol T., Roncea F.N., Prelipcean A.M., Craciunescu O., Health Benefits of Antioxidant Bioactive Compounds in the Fruits and Leaves of *Lonicera caerulea* L. and *Aronia melanocarpa* (Michx.) Elliot, *Antioxidant*, 12 (2023).
- [13] Turkoglu A., Duru M.E., Mercan N., Kivrak I., Gezer K., Antioxidant and Antimicrobial Activities of *Laetiporus sulphureus* (Bull.) Murrill, *Food Chem.*, 101 (2007) 267–273.
- [14] Mungan F., Yıldız K., Kılıç M., Kuh M., A Morphological Study of *Smyrniun* (Apiaceae) from Turkey, *Biological Diversity and Conservation*, 8 (2015) 54–59.
- [15] Ayaz F., Eryugur N., Dođru T., Bađcı Y., Dođu S., Biological Activities of the Methanol Extracts of *Smyrniun connatum* Boiss. and Kotschy, *Kahramanmaraş Sütçü İmam Üniversitesi Tarım ve Dođa Dergisi*, 25 (2022) 1216–1224.
- [16] Özbilgin S., Kúpeli Akkol E., Süntar İ., Tekin M., Saltan İřcan G., Wound-Healing Activity of Some Species of *Euphorbia* L., *Records of Natural Products*, 13 (2018) 104–113.
- [17] Tang Q., Su Z., Han Z., Ma X., Xu D., Liang Y., LC–MS Method for Detecting Prostratin in Plant Extracts and Identification of a High-Yielding Population of *Euphorbia fischeriana*, *Phytochem. Lett.*, 5 (2012) 214–218.
- [18] Rojas-Jiménez S., Valladares-Cisneros M.G., Salinas-Sánchez D.O., Pérez-Ramos J., Sánchez-Pérez L., Pérez-Gutiérrez S., Anti-Inflammatory and Cytotoxic Compounds Isolated from Plants of *Euphorbia* Genus, *Molecules*, 29 (2024) 1083.
- [19] Sarı A.O., Ođuz B., Bilgiç A., Tort N., Güvensen A., řenol S.G., Ege ve Güney Marmara Bölgelerinde Halk İlacı Olarak Kullanılan Bitkiler, *A.A.R.I.*, 20 (2010) 1–21.
- [20] Mata A.T., Proença C., Ferreira A.R., Serralheiro M.L.M., Nogueira J.M.F., Araújo M.E.M., Antioxidant and Antiacetylcholinesterase Activities of Five Plants used as Portuguese Food Spices, *Food Chem.*, 103 (2007) 778–786.
- [21] Okur M.E., Ayla ř., Karadađ A.E., Çiçek Polat D., Demirci S., řeçkin İ., *Opuntia ficus indica* Fruits Ameliorate Cisplatin-Induced Nephrotoxicity in Mice, *Biol. Pharm. Bull.*, 43 (2020) 831–838.
- [22] Blois M.S., Antioxidant Determinations by the Use of a Stable Free Radical, *Nature*, 181 (1958) 1199–1200.
- [23] Karakaş N., Karadađ A.E., Yılmaz R., Demirci F., Okur M.E., *In Vitro* Cytotoxicity Evaluation of *Marrubium vulgare* L. Methanol Extract, *Journal of Research in Pharmacy*, 23 (2019) 711–718.
- [24] Re R., Pellegrini N., Proteggente A., Pannala A., Yang M., Rice-Evans C., Antioxidant Activity Applying an Improved ABTS Radical Cation Decolorization Assay, *Free Radic. Biol. Med.*, 26 (1999) 1231–1237.
- [25] Hearing V.J., Jiménez M., Mammalian Tyrosinase—The Critical Regulatory Control Point in Melanocyte Pigmentation, *International Journal of Biochemistry*, 19 (1987) 1141–1147.
- [26] řahin D., Çađlar E.ř., Boran T., Karadađ A.E., Özhanel G., Üstündađ Okur N., Development, Characterization of Naringenin-Loaded Promising Microemulsion Formulations, and Demonstration of Anti-Aging Efficacy by *In Vitro* Enzyme Activity and Gene Expression, *J. Drug Deliv. Sci. Technol.*, 84 (2023) 104422.
- [27] Gören N., Ulubelen A., Jakupovic J., Bohlmann F., Grenz M., Sesquiterpene Lactones from the Fruits of *Smyrniun rotundifolium*, *Phytochemistry*, 23 (1984) 2281–2284.
- [28] Gören N., Ulubelen A., Oxepine Derivatives from the Roots of *Smyrniun rotundifolium*, *Phytochemistry*, 24 (1985) 3051–3052.
- [29] Rozimamat R., Kehrimen N., Aisa H.A., New Compound from *Euphorbia alata* Boiss., *Nat. Prod. Res.*, 33 (2019) 380–385.
- [30] Kaçar D., Bayraktar O., Erdem C., Alamri A.S., Galanakis C.M., Antioxidant and Antimicrobial Activities of Plants Grown in the Mediterranean Region, *JFA Reports*, 2 (2022) 452–461.
- [31] Kulevanova S., Stefova M., Panovska T.K., Stafilov T., HPLC Identification and Determination of Myricetin, Quercetin, Kaempferol and Total Flavonoids in Herbal Drugs, *Macedonian Pharmaceutical Bulletin*, 48 (2002) 25–30.
- [32] Erşin M., Ergönül B., Kalyoncu F., Proximate Composition, Antimicrobial and Antioxidant Activities of Six Wild Edible Celeries (*Smyrniun* L.), *Afr. J. Pharm. Pharmacol*, 6 (2012).

- [33] Orhan I., Tosun F., Skalicka-Woźniak K., Cholinesterase, Tyrosinase Inhibitory and Antioxidant Potential of Randomly Selected Umbelliferous Plant Species and Chromatographic Profile of *Heracleum platytaenium* Boiss. and *Angelica sylvestris* L. var. *sylvestris*, *Journal of the Serbian Chemical Society*, 81 (2016) 357–368.
- [34] Ulubelen A., Abdolmaleky H., Mabry T.J., Flavonoid Glycosides from *Smyrniium perfoliatum*, *Smyrniium creticum* and *Smyrniium rotundifolium*, *J. Nat. Products*, 45 (1982) 507.
- [35] Jassbi A.R., Chemistry and Biological Activity of Secondary Metabolites in *Euphorbia* from Iran, *Phytochemistry*, 67 (2006) 1977–1984.
- [36] Kirbag S., Erecevit P., Zengin F., Guvenc A., Antimicrobial Activities of Some *Euphorbia* Species, *African Journal of Traditional Complementary and Alternative Medicines*, 10 (2013).
- [37] Solimine J., Garo E., Wedler J., Rusanov K., Fertig O., Hamburger M., Tyrosinase Inhibitory Constituents from a Polyphenol Enriched Fraction of Rose Oil Distillation Wastewater, *Fitoterapia*, 108 (2016) 13–19.
- [38] Laksmiani N.P.L., Widianara I.W.A., Pawarrangan A.B.S., Potency of Moringa (*Moringa oleifera* L.) Leaves Extract Containing Quercetin as a Depigmentation Agent Inhibiting the Tyrosinase Enzyme Using *In-Silico* and *In-Vitro* Assay, *Pharmacia*, 69 (2022) 85–92.

Sensitive Analysis of Epilepsy Drug, Phenobarbital, Based on Column Type Solid Phase Extraction and HPLC-DAD System

Halil İbrahim Ulusoy^{1,a,*}, Seren Kilit Bülbül^{1,b}, Ümmügülüm Polat^{1,c}, Esra Durgun^{1,d}, Aslıhan Gürbüzler^{2,e}, Elif Tümay Özer^{3,f}, Bilgen Osman^{3,g}

¹ Department of Analytical Chemistry, Faculty of Pharmacy, Sivas Cumhuriyet University, 58140 Sivas, Türkiye

² Department of Plant and Animal Production, Vocational School of Technical Services, Sivas Cumhuriyet University, 58140 Sivas, Türkiye

³ Department of Chemistry, Faculty of Arts and Science, Bursa Uludağ University, 16059 Bursa, Türkiye

*Corresponding author

Research Article

History

Received: 23/10/2024

Accepted: 19/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

An analytical approach has been developed for the sensitive determination of Phenobarbital, which is used as an antiepileptic drug molecule. The analysis of this active molecule, which has a very limited study even for its direct determination in the literature review, is mostly done with highly complex device systems such as LC-MS. In analysis with conventional HPLC systems, the limit of detection cannot be in most cases lower than 1 µg mL⁻¹. With this study, a separation and preconcentration method based on solid phase extraction (SPE) was developed for trace phenobarbital molecules, so that even very low concentrations could be monitored. In the proposed method, the target molecules were enriched with column type SPE method, and then their analysis were carried out by with the HPLC-DAD system. As an SPE sorbent, a polymeric material, poly(ethylene glycol dimethacrylate-N-methacryloyl-L-tryptophan methyl ester) [poly(EGDMA-MATrp)], was used in extraction experiments. Experimental variables such as pH of medium, type and amount of desorption solvent, electrolyte effect have been studied and optimized step by step. The linear working range under the optimized conditions were determined in the range of 10.00-400.00 ng mL⁻¹ with the limit of detection as 3.57 ng mL⁻¹. Quantitative results were obtained in recovery experiments with the help of model solutions including phenobarbital molecule.

Keywords: Phenobarbital, Solid phase extraction, HPLC, Analytical method development.

^a hiulusoy@cumhuriyet.edu.tr

^b <https://orcid.org/0000-0002-8742-5145>

^c ugulsum@cumhuriyet.edu.tr

^d <https://orcid.org/0000-0001-8393-2396>

^e aslihan@cumhuriyet.edu.tr

^f <https://orcid.org/0000-0002-6753-2808>

^g bilgeno@uludag.edu.tr

^h <https://orcid.org/0000-0001-8406-149X>

ⁱ bkiliteren1997@gmail.com

^j <https://orcid.org/0009-0003-4453-4215>

^k desradurgun2017@gmail.com

^l <https://orcid.org/0009-0009-2613-4019>

^m etumay@uludag.edu.tr

ⁿ <https://orcid.org/0000-0002-5225-0146>

Introduction

Phenobarbital (5-ethyl-5phenyl-1,3-diazinan-2,4,6-trione), as a barbiturate, is a hypnotic drug substance derived from barbituric acid. It is the oldest anticonvulsant drug still used today. Although it has hypnotic and sedative properties, it is not used for these effects today, benzodiazepines are preferred for this purpose[1]. Phenobarbital, partial and is a general central nervous system depressant used in the control and treatment of generalized epilepsy disease [2–4]. Molecular formula of phenobarbital was given in Figure 1.

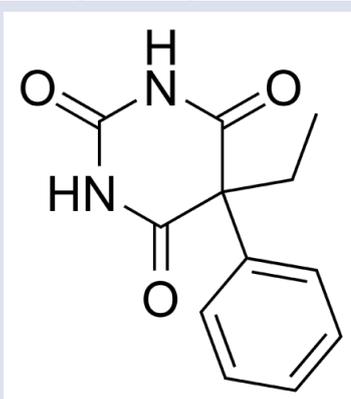


Figure 1. Chemical structure of phenobarbital

The degradation products of phenobarbital at high pHs are diethylmalonic acid, diethylacetylurea, diethyl acetic acid, ammonia and carbonate, respectively. Barbituric acid derivatives are sufficiently acidic due to the nitrogens attached to the carbonyl group in their structures from two directions and the hydrogens in the -CH₂ group, and they form salts with alkalis. If the pH of the aqueous solutions of these salts falls below a certain value, the barbituric acid part precipitates in non-ionized form[5]. Injection solutions are prepared with organic solvents such as dimethylformamide, dimethylacetamide and propylene glycol to increase the solubility and stability of barbituric acid derivatives. Phenobarbital is a drug that stabilizes the work of the brain and nervous system and shows its function[6]. It is used in convulsions in grand mal epilepsy and other types of epilepsy, continuous treatment of eclampsia and spastic conditions, combined treatment of whooping cough, persistent insomnia, chorea minor, angina pectoris, hyperthyroidism and tensions in the climacterium[7,8].

The analysis of food, biological, pharmaceutical and environmental samples is the most important application area of solid phase extraction methods [9-11]. Especially in biological samples, SPE based methods are widely used

for the enrichment of organic molecules and trace heavy metals found mostly in water at $\mu\text{g/L}$ and less [12–14]. Level of target molecules in biological samples is generally lower than detection limit of conventional analysis system. So, easy applicable and effective sample pre-treatment procedures are valuable for trace analysis of drug molecules. Use of new and useful adsorbents in extraction procedures presents new ways for this challenges.

The aim of the presented study is to develop a HPLC based determination method after a sample preparation procedure with a solid phase extraction method for the sensitive and selective determination of phenobarbital molecules that are present at trace levels in real samples and cannot be directly analyzed by classical determination devices.

Material Methods

Chemicals and Reagents

All reagents used during the experiments were of analytical purity and purchased from Sigma or Merck. All solutions used were prepared with ultrapure water with a resistivity of $18.2 \text{ M } \Omega$ obtained from ELGA Pure Lab Flex III device.

For pH 2.0-11.0 Britton Robinson (BR) buffer, a stock BR solution containing H_3BO_3 , H_3PO_4 and CH_3COOH acids was prepared to contain 0.05 M each. 0.1 M NaOH was added dropwise by checking with a pH meter in the appropriate pH ranges according to the acidity constants and adjusted to the desired pH.

Phenobarbital stock solution; 50 mg of phenobarbital supplied by Sigma Aldrich was weighed and put into a volumetric flask and dissolved with some methyl alcohol to make up to 50 ml and transferred to a brown bottle and stored in the refrigerator at $+4^\circ\text{C}$.

Instrumentation

In all chromatographic measurements, Shimadzu (Prominence) HPLC (Kyoto, Japan) device was used. The HPLC device used has LC 20 AD quaternary pump, SPD-M20 A PDA detector, DGU-20A vacuum degasser and CTO-10 AS VP column oven equipment. All separations and determinations were performed on reverse phase C18 column (C18, 250 mm \times 4.6 mm, 5 μm). Evaluation of chromatograms was done via LC Solution 2.0 software. The following devices were used during the optimization experiments; pH meter (pH-2005, JP Selecta, Barcelona, Spain), Vortex (Jeotech, Korea), Ultrasonic Water Bath (SK521 OHP), Shaker (WhyShake SHO-2D).

HPLC Determination Conditions for the Phenobarbital Analysis

Before proceeding to the SPE experiments, the parameters of direct determination were optimized for Phenobarbital molecule by HPLC-DAD. For this purpose, based on literature information, Luna Omega C-18 column was selected as the stationary phase and the ideal eluent

composition was investigated to enable conditions for determination. Many experiments were conducted in isocratic and gradient elution modes with aqueous solutions containing buffers at different pHs of organic eluent phases such as methanol, ethanol and acetonitrile and different eluent phase compositions. Experiments were continued until good peaks were obtained for phenobarbital. As a result of the experiments, the most ideal mobile phase conditions were obtained with methanol and pH 3.0 phosphate buffer eluents. The ideal HPLC operating conditions obtained after optimization were given in Table 1. The chromatogram obtained for 3 different Phenobarbital concentrations under the optimized working conditions was also shown in Figure 2.

Table 1. HPLC operating conditions

| Parameter Value | Isocratic |
|--------------------|---|
| HPLC Mode | Isocratic |
| Eluent | Methyl Alcohol: pH 3.00 Phosphate Buffer (70:30) |
| Eluent Flow Rate | 1 mL/min |
| Run Time | 12 min |
| Column | C18- Luna Omega (250 mm \times 4.6, 5.0 μm) |
| Column Temperature | 30 $^\circ\text{C}$ |
| Injection Volume | 10 μL |

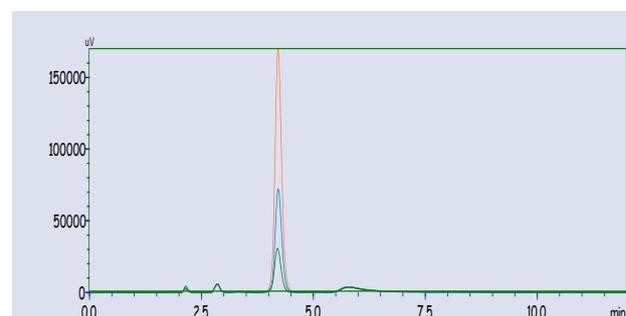


Figure 2. The standard chromatogram of Phenobarbital before SPE

The obtained calibration graph on the HPLC-DAD system with the help of standard solutions prepared at 5 different concentrations (2, 5, 10, 20 and 50 $\mu\text{g mL}^{-1}$) was given in Figure 3.

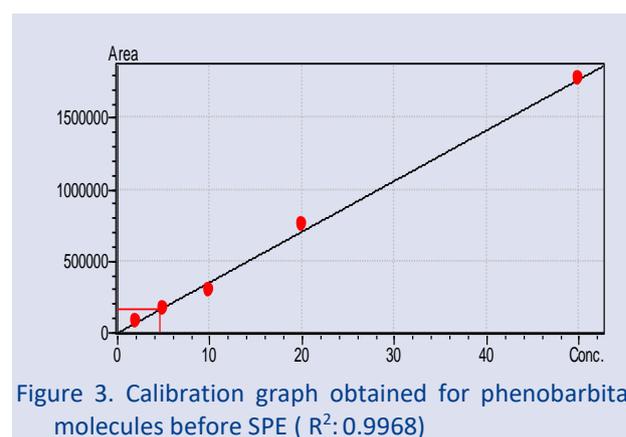


Figure 3. Calibration graph obtained for phenobarbital molecules before SPE ($R^2: 0.9968$)

As seen in Figure 4, the spectrum taken from the DAD detector of the HPLC device for phenobarbital molecules and the maximum absorption wavelength was

determined as 199 nm for further studies. Under optimized HPLC conditions, Table 2 shows direct determination parameters before SPE for phenobarbital molecules.



Figure 4. UV-VIS Spectrum obtained in DAD Detector for Phenobarbital

Table 2. Direct analysis results by HPLC

| Parameter | Value |
|--|--------------------------------|
| Retention Time, min | 4.25 |
| Maximum absorption wavelength, λ | 199 nm |
| Calibration Range | 2.0-50.0 $\mu\text{g mL}^{-1}$ |
| Limit of Detection, LOD | 0.57 $\mu\text{g mL}^{-1}$ |
| RSD % (for 5.0 $\mu\text{g mL}^{-1}$) | 3.47 |
| R^2 | 0.9968 |
| Number of repetitions | 3 |

Synthesis and Characterization of the Used Solid Phase Sorbent [poly(EGDMA-MATrp)]

Synthesis of the used polymeric sorbent was carried out by using N-methacryloyl-L-tryptophan methyl ester (MATrp) and (EGDMA) as a monomer and a cross-linker, respectively. The sorbent microbeads were synthesized by suspension polymerization technique. The details about synthesis and characterization of used adsorbent in SPE are available in our previous studies[15,16]. Briefly, a 200 mg of poly (vinyl alcohol) (PVA) was dissolved in 50 mL of the deionized water and then the dispersion phase was prepared by means of an organic phase including MATrp (4 mL), EGDMA (5 mL), toluene (10 mL) and PVA solution. After that, 100 mg of (AIBN) was transferred to the polymerization mixture. Polymerization was performed at 85 °C for 8 h with a 600 rpm stirring rate. The final product was washed with ethanol and water in order to remove unreacted chemicals and then dried at 50°C in a vacuum oven.

The Proposed Solid Phase Extraction Method

A solid phase extraction based separation and preconcentration method was developed for

Phenobarbital molecules, which are found at trace levels in real samples and are difficult to determine due to the complex matrix of the sample. At the beginning of the experimental studies, preliminary experiments were carried out regarding all the parameters that will provide quantitative transition of the relevant substance to the rich phase. The main goal in solid phase extraction experiments is to create ideal experimental conditions that will provide the transition of analyte species to the solid phase with the highest possible efficiency.

200 mL of poly(EGDMA-MATrp) microbeads were weighed in a 1 mL empty SPE cartridge using an upper frit and a lower frit to prevent the leakage of adsorbent from the SPE cartridge. Prior to extraction, the cartridge packed with poly(EGDMA-MATrp) microbeads was preconditioned with 3 mL of methanol and then with 3 mL of water. The pH of samples was adjusted a value of 4.0 using Britton-Robinson Buffer solutions. The sample solution (50 mL) was passed through the cartridge at a flow rate of 0.75 mL min⁻¹. Then, 500 μL of ACN:MeOH (1:1) mixture was used to elute the analytes retained on the cartridge. The eluent solution was filtered through a 0.45 μm membrane, and 10 μL of the solution were injected into the HPLC-PDA system for analysis.

Results and Discussion

pH effect

The pH of the medium is a very important parameter, as it affects the interaction of the analyte with the solid phase and the reactions between species. Two milliliters (2 mL) of Britton-Robinson (BR) buffer in the range of pH 2–10 was added to all tubes containing phenobarbital, and their volumes were adjusted to 50 mL with ultrapure water. The prepared solutions were then filtered using an

SPE manifold, allowing the analyte molecules to attach to the solid phase, and subsequently analyzed using an HPLC-DAD system. The analytical responses obtained by this optimization were shown in Figure 5. It was observed that the optimum pH value for the enrichment process is 4.00. Consequently, it was understood that phenobarbital becomes positively charged in the acidic region due to its pKa value. According to the literature, the pKa value of phenobarbital is reported as 7.30 [17], which is consistent with our findings.

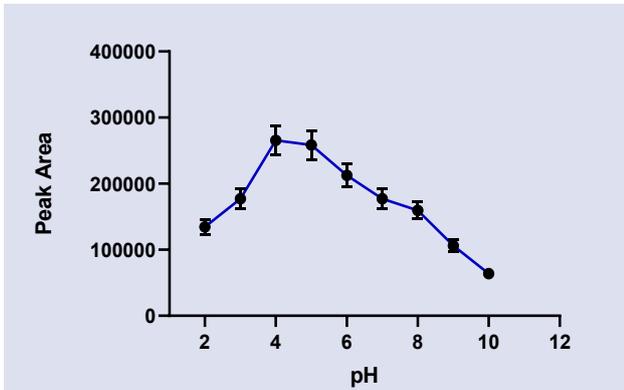


Figure 5. pH effect on the proposed SPE method.

Electrolyte Effect

To observe the effect of electrolyte concentration on the developed method, a series of experiments were performed using NaCl solution. Evaluating how the solid-phase extraction system is influenced by increased electrolyte concentration in the medium and determining whether there are changes in the analyte signals are important parameters for method stability. The ability of the method to perform even in the presence of high electrolyte concentrations is significant, as it confirms its reliability for real sample applications and demonstrates its potential to enhance signals by maintaining charge balance in the medium in certain cases. However, in some instances, analyte signals may be negatively affected as the electrolyte concentration increases. To assess these effects, increasing concentrations of NaCl were added to the model solutions containing analytes, and changes in the peak areas of phenobarbital were monitored. As shown in Figure 6, even at high electrolyte concentrations, no disruptive effect was observed on the analyte signals.

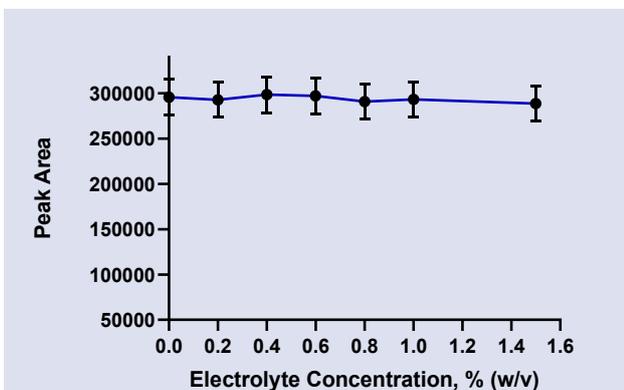


Figure 6. Effect of electrolyte concentration on the proposed SPE method.

Desorption Solvent Selection

After completing the filtration process with the SPE manifold, the next step was to select the most suitable solvent for separating the material bound to the solid phase. The selected solvent needed to completely dissolve the solid phase enriched with phenobarbital while ensuring it would not damage the HPLC system during analysis. When choosing the solvents for this purpose, several candidates were tested based on their compatibility with the HPLC eluent phase and their ability to quantitatively dissolve the SPE. For this process, 1 mL of various solvents, including methanol (MeOH), acetonitrile (ACN), isopropyl alcohol (IPA), ethanol (EtOH), ACN:MeOH (1:1), water, and n-hexane, were tested sequentially. The results obtained with these solvents after the enrichment process were presented in Figure 7.

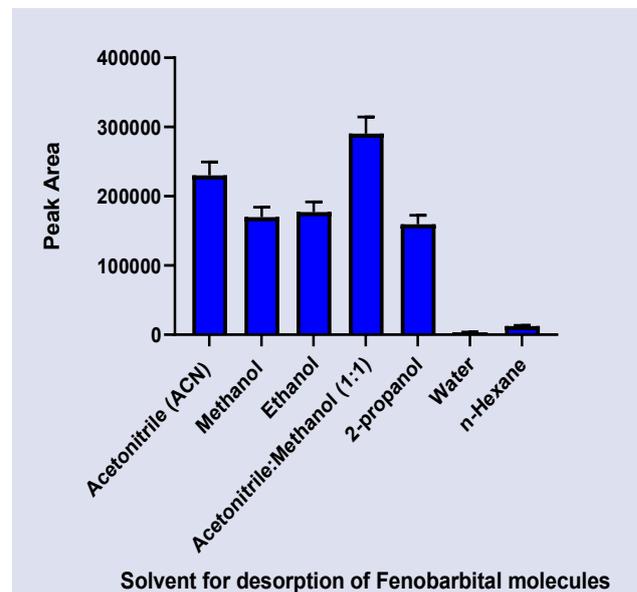


Figure 7. Eluents for desorption process.

As seen in Figure 7, the highest signals were obtained with Acetonitril:Methanol (1:1) solution. Therefore, in subsequent studies, the Phenobarbital-rich solid phase was dissolved with this solvent. Since the amount of solvent used to dissolve the phenobarbital-rich solid phase will directly affect the enrichment factor, it is important how much of the solvent volume will be taken. In order to obtain a high enrichment coefficient, the solvent volume should be at the smallest value. Because the enrichment coefficient decreases as the solvent volume increases. 200 μ L of sample can be placed in HPLC micro vials. Volume optimization was performed by adding 250-2000 μ L of solvent and, as expected and seen in Figure 8, the signals decrease as the solvent volume increases. In order to obtain the best signals, 1000 μ L was determined as the appropriate volume, which is the volume in which the phenobarbital-rich solid phase can be dissolved and filtered, and this value was used in subsequent studies.

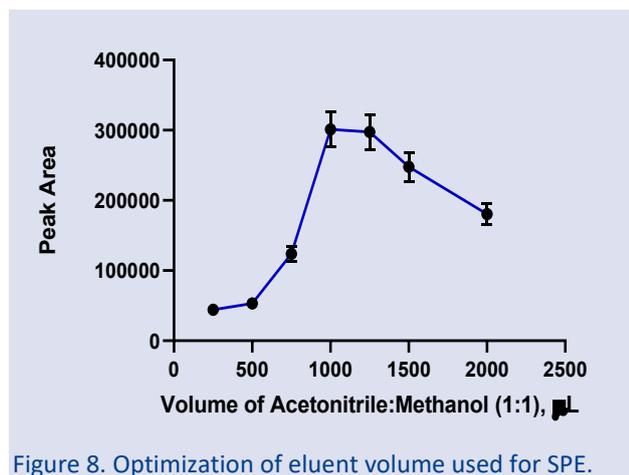


Figure 8. Optimization of eluent volume used for SPE.

Analytical Performance Parameters of the developed SPE-HPLC-DAD based Method

After determining the most suitable experimental conditions for solid phase extraction, enrichment experiments were applied to Phenobarbital solutions at different concentrations in order to determine the linear working range, and as a result, it was determined that the measured signals varied linearly in the range of 10.00-

400.00 ng mL⁻¹. The calibration line obtained by applying the method we developed to these solutions is given in Figure 9. As can be seen, the signals increase proportionally with the concentration. All analytical parameters of the developed method were collectively presented in Table 3.

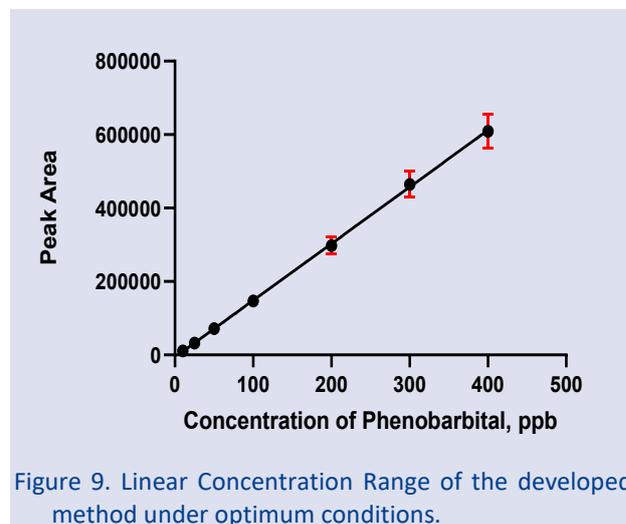


Figure 9. Linear Concentration Range of the developed method under optimum conditions.

Table 3. Analytical parameters of the proposed method

| Parameter | Before SPE | After SPE |
|--|--------------------------------|----------------------------------|
| Linear range | 2.00-50.00 µg mL ⁻¹ | 10.00-400.00 ng mL ⁻¹ |
| Limit of detection(LOD) ^a | 0.56 µg mL ⁻¹ | 3.57 ng mL ⁻¹ |
| Limit of Quantification (LOQ) ^b | 1.91 µg mL ⁻¹ | 9.58 ng mL ⁻¹ |
| RSD (%) (for 250 ng mL ⁻¹) | 4.57 | 3.70 |
| Calibration Sensitivity | 1.12 | 69.44 |
| Correlation Coefficient (R ²) | 0.9968 | 0.9846 |
| Enrichment Factor ^c | - | 50 |
| Enhancement Factor ^d | - | 62 |

^{a,b} The selection limit was calculated using 3 times the standard deviation obtained from blank trial signals with at least 3 replicates, and the quantification limit was calculated using 10 times the standard deviation.

^c Enrichment factor; It was calculated by taking the ratio of the initial aqueous phase volume (50 mL) to the volume obtained after enrichment (1.0 mL)

^d Enhancement Factor; It was computed by taking the ratio of calibration sensitivity(slope) after and before SPE

Conclusion

Solid-phase extraction (SPE) is a method that has been extensively and increasingly used for the separation and enrichment of both organic and inorganic species, especially over the past twenty years. Hundreds of studies are published annually in this field, demonstrating its wide application for various species and sample types. In the last two decades, SPE has been effectively applied for the enrichment of organic species, followed by their determination using appropriate analytical methods.

The key factors contributing to the widespread application of SPE include its simplicity, environmental friendliness, low cost, and easy applicability in almost any laboratory. The primary goal of SPE experiments is to transfer target species from a relatively large volume of aqueous medium into a smaller volume of a phase immiscible with water—like classical liquid-liquid extraction—thereby enriching the analyte to concentrations detectable by analytical instruments.

Despite its advantages, several challenges may arise during the application of SPE. For example, variability in conditioning conditions before extraction, insufficient contact between the inhomogeneous solid-phase support material and the solution, and time-consuming procedures may hinder the process. Additionally, the retaining materials may sometimes lack sufficient selectivity for the analyte of interest. To minimize these drawbacks, optimization studies are carried out to improve method performance.

Like all enrichment methods, SPE aims to concentrate analytes that cannot be directly detected to levels measurable by analytical instruments. In this study, an enrichment and determination method was developed for this purpose. Direct determination of organic analytes, such as phenobarbital, is often challenging in many samples due to the complex matrix structure of real samples and the very low analyte concentrations. For such analytes, which are predominantly analyzed

chromatographically, the minimum concentration detectable using an HPLC system with an absorbance detector is typically no lower than 2 µg mL⁻¹.

To address this limitation, a cost-effective, accurate, precise, and easy-to-apply enrichment method was developed for use in any laboratory. Before initiating the enrichment process, direct determination conditions for phenobarbital were optimized. Following method optimization, analytical parameters were determined. Under optimized conditions, the linear working range for phenobarbital was found to be 10.00–400.00 ng mL⁻¹. The calculated limits of detection (LOD) and quantification (LOQ) were 3.57 ng mL⁻¹ and 9.58 ng mL⁻¹, respectively.

Acknowledgment

This study is the graduation thesis of Seren Kilit Bülbül, who completed her degree at the Faculty of Pharmacy in 2020. The experimental data were collected and organized with the assistance of the other authors. Additionally, the experimental work for this study was conducted using project funds provided by the Scientific Research Projects Commission of Sivas Cumhuriyet University.

Conflict of interests

The authors declare that they have no conflicts of interest.

References

- [1] Vermeij T. A. C., Edelbroek P. M., Robust isocratic high performance liquid chromatographic method for simultaneous determination of seven antiepileptic drugs including lamotrigine, oxcarbazepine and zonisamide in serum after solid-phase extraction, *Journal of Chromatography B*, 857(1) (2007) 40-46.
- [2] Velayati S., Saadati F., Shayani-Jam H., Shekari A., Valipour R., Yaftian M. R., Fabrication and evaluation of a molecularly imprinted polymer electrochemical nanosensor for the sensitive monitoring of phenobarbital in biological samples, *Microchemical Journal*, 174 (2022) 107063.
- [3] Mobed A., Shirafkan M., Charsouei S., Sadeghzadeh J., Ahmadiipour A., Biosensors technology for anti-epileptic drugs, *Clinica Chimica Acta*, 533 (2022) 175-182.
- [4] Lyseng-Williamson K. A., Levetiracetam: a review of its use in epilepsy, *Drugs*, 71 (2011) 489-514.
- [5] A Spectrophotometric Method for the Determination of Nitrite and Nitrate, *Eurasian J. Anal. Chem.*, (2017).
- [6] Jacob S., Nair A. B., An updated overview on therapeutic drug monitoring of recent antiepileptic drugs, *Drugs in R&D*, 16 (2016) 303-316.
- [7] Ventura S., Rodrigues M., Pousinho S., Falcão A., Alves G., Determination of lamotrigine in human plasma and saliva using microextraction by packed sorbent and high performance liquid chromatography–diode array detection: An innovative bioanalytical tool for therapeutic drug monitoring, *Microchemical Journal*, 130 (2017) 221-228.
- [8] D'Urso A., Locatelli M., Tartaglia A., Molteni L., D'Ovidio C., Savini F., de Grazia U., Therapeutic drug monitoring of antiseizure medications using volumetric absorptive microsampling: where are we?, *Pharmaceuticals*, 14(7) (2021) 627.
- [9] Sultan M., Simultaneous HPLC determination and validation of amphetamine, methamphetamine, caffeine, paracetamol and theophylline in illicit seized tablets, *Int. J. Pharm. Pharm. Sci.*, 6(4) (2014) 294-298.
- [10] Pena A., Pelantova N., Lino C. M., Silveira, M. I. N., & Solich, P., Validation of an analytical methodology for determination of oxytetracycline and tetracycline residues in honey by HPLC with fluorescence detection, *Journal of Agricultural and Food Chemistry*, 53(10) (2005) 3784-3788.
- [11] Ulusoy H. İ., Yiğit İ. N., Polat Ü., Durgun E., Gürbüz A., Ulusoy S., Simultaneously HPLC Analysis of B1, B9 and B12 Vitamins at Trace Levels via Cloud Point Extraction, *Cumhuriyet Science Journal*, 44(4) (2023) 716-722.
- [12] Ulusoy H. İ., Polat U., Ulusoy S., Use of newly synthesized magnetic Fe₃O₄ nanoparticles modified with hexadecyl trimethyl ammonium bromide for the sensitive analysis of antidepressant drugs, duloxetine and vilazodone in wastewater and urine samples, *RSC Advances*, 13(29) (2023) 20125-20134.
- [13] Ulusoy S., Erdoğan S., Karaslan M. G., Ateş B., Ulusoy H. İ., Erdemoğlu S., Optimization Of Extraction Parameters For Folic Acid And Antioxidant Compounds From An Edible Plant (*Polygonum Cognatum* Meissn) Using Pressurized Liquid Extraction (PLE) System, *Cumhuriyet Science Journal*, 39(4) (2018) 1069-1080.
- [14] Karaca E., Ulusoy S., Morgül Ü., Ulusoy H. İ., Development of analytical method for sensitive determination of streptozotocin based on solid phase extraction, *Cumhuriyet Science Journal*, 41(4) (2020) 826-831.
- [15] Demir Ö., Ulusoy H. İ., Özer E. T., Osman, B., Development of a new solid phase extraction method for sensitive determination of some carbamate pesticides in water using poly (EGDMA-MATrp) microbeads, *Microchemical Journal*, 158 (2020) 105317.
- [16] Osman B., Özer E. T., Kara A., Yeşilova E., Beşirli N., Properties of magnetic microbeads in removing bisphenol-A from aqueous phase, *Journal of Porous Materials*, 22 (2015) 37-46.
- [17] Methaneethorn J., Leelakanok N., Pharmacokinetic variability of phenobarbital: a systematic review of population pharmacokinetic analysis, *European Journal of Clinical Pharmacology*, 77 (2021)291-309.

Analyzing Diabetic Dynamics with MRK4, and LSTM Techniques with Multiplicative Calculus

Buğçe Eminağa Tatlıcıoğlu^{1,a,*}

¹ Department of Computer Engineering, Faculty of Engineering, Final International University, Kyrenia, Northern Cyprus, via Mersin 10, Türkiye.

*Corresponding author

Research Article

History

Received: 22/02/2024

Accepted: 24/10/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This study compares the use of Long Short-Term Memory (LSTM) networks for predictive modeling with multiplicative calculus. We evaluate and quantitatively analyze both methodologies to determine their prediction performance. While LSTM networks are investigated for their power to learn and generalize patterns, the multiplicative calculus technique is analyzed for its ability to grasp complex connections within the data. This study attempts to shed light on the efficacy of each approach by carefully analyzing error measures including mean squared error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The results aid in the comprehending of the subtleties related to LSTM networks and multiplicative calculus, assisting practitioners and researchers in choosing the best method for tasks involving predictive modeling.

Keywords: Neural network, Runge kutta, Multiplicative calculus, Long Short-Term Memory, Diabetes.

bugce.tatlicioglu@final.edu.tr

<https://orcid.org/0000-0001-8854-4464>

Introduction

The utilization of big data in computational biology research is on the rise due to the quick and practical creation of many data due to improvements in high-throughput computing and biotechnology. The main objective is to analyze the growing corpus of biological data and offer a basis for tackling significant biological and medical problems. Although these techniques can precisely identify patterns and create models from data, they are dependable and efficient. One important use of this study is in the detection and treatment of life-threatening conditions, such as diabetes mellitus (DM). [1-3].

Diabetes is a well-known and serious issue that affects both industrialized and developing nations in the modern world [4]. The hormone responsible for facilitating the body's absorption of glucose from diet is insulin. This deficit, which is often brought on by pancreatic dysfunction, can lead to several significant symptoms, including coma, renal [5], heart disease [6], retinal failure, and cardiovascular problems [7]. Research has shown that the number of adult (18 and older) cases of diabetes increased significantly between 1980 and 2014 [1], and by 2045, more cases are expected. Global estimates of diabetes patients in 2017 were 451 million; by 2045, that number is expected to climb to 693 million [8].

An extra statistical analysis [6] indicates that approximately 500 million individuals globally suffer from diabetes, highlighting the severity of the issue. Furthermore, the study projects a notable rise in diabetes prevalence, with estimations of 25% by 2030 and 51% by 2045. However, there is still a significant need for research focused on improving the health and quality of life (QoL)

of persons with diabetes, as well as reducing the start of disease complications and premature mortality. It's critical to remember that, even in the absence of long-term treatment, effective management and prevention are still achievable, particularly in the case that accurate and early forecasts can be established.

Recent technological developments, particularly the use of artificial intelligence methods, have shown to be extremely advantageous for the healthcare sector. The literature offers a variety of methods and approaches to enhance diabetes accuracy. In recent years, several of diabetes prediction techniques have been developed and published. A framework based on machine learning was introduced in [9]. Numerous classification algorithms were employed by the authors, including Naive Bayes (NB) [10], Support Vector Machine (SVM) [11], AdaBoost (AB) [12], Decision Tree (DT) [13], Random Forest (RF) [14], Logistic Regression (LR) [15], Gaussian Process Classification (GPC) [10] and Artificial Neural Network (ANN) [11]. In addition, a logistic regression model was presented by Qawqzeh et al. [15] for the categorization of type 2 diabetes, and it achieved an impressive 92% accuracy.

Nevertheless, a comparison study with known procedures was absent from their research. Further research that focused on the effectiveness of linear support vector machines (SVMs) also used SVMs to classify individuals with diabetes mellitus. However, this study only included scant information on parameter selection and lacked a comprehensive comparison with state-of-the-art technology [16]. Furthermore, although the study did not specify accuracy standards, the PIMA

Indian Diabetes Study was used as the basis for the investigation of the categorization of diabetes using naïve Bayes (NB) and Support Vector Machines (SVMs). Numerous additional studies have used a variety of machine learning methods and datasets to predict diabetes; some of these studies have produced encouraging results. Some of these research, too, lacked direct comparisons with cutting-edge techniques [17, 18].

Furthermore, a thorough assessment of machine learning algorithms for diabetes prediction from 2010 to 2019 was carried out by Hussain et al. (2021). Based on the Matthews correlation coefficient, they discovered that Random Forests and naïve Bayes performed better overall [19].

Complex-valued function solutions involving complex variables have been addressed by Bashirov and Riza [20, 21] and Uzer [22] through the expansion of geometric (multiplicative) calculus. Differential equation modeling has become a part of this discipline [23]. Due to its extensive applications in fields including applied and pure mathematics, engineering [24], applied mechanics, quantum physics, analytical chemistry, astronomy, and biology, nonlinear equations have attracted a lot of attention [24, 25]. Aniszewska et al. [26] investigated an alternate use of the Runge-Kutta technique in dynamical systems, and [27] created the bigeometric variant of the Runge-Kutta method. For approximating solutions to ordinary differential equations, temporal discretization makes considerable use of this approach, which belongs to an important family of implicit and explicit iterative methods in numerical analysis [28]. The Runge-Kutta family includes the popular RK4 technique, which is well-known in numerical analysis for its efficiency in approximating solutions to ordinary differential equations.

LSTM networks have been increasingly explored in the medical field for predicting and managing various diseases beyond diabetes. Researchers have applied LSTMs in the prediction and diagnosis of complex conditions such as heart disease [29], Alzheimer's [30], and skin diseases [31], where temporal patterns in patient data play a crucial role.

Proposed Method

According to this notion, individuals with diabetes are split into two categories based on their current health status: group C is for individuals with challenges, and group D is for those without issues. The paradigm treats acute and chronic disorders in the same way, without distinguishing between various types of consequences. Initially, individuals without difficulties in the diabetic population can advance to the complications group at the pace indicated by the symbol λ . This transition results in a reduction in group D's population by λ times the current number, while group C's population increases proportionately. For individuals with diabetes-related problems, three probable outcomes are taken into account: disability, death, or recovery. It is assumed that those who have recovered would remain in the diabetic

category and keep their diabetes status. According to the diabetic complications model (DC), people with diabetes who encounter difficulties may be able to overcome them, pass away from them, or become disabled. It's noteworthy to notice that people who recover from diabetes are still considered to be diabetics, even if those who die or become crippled are no longer considered to be part of the community. The variables γ , δ , and ν , in that order, indicate the rates of recovery, disability, and mortality due to complications.

The number of people in the complications section fluctuates as a result. The number of people experiencing difficulties connected to their recovery, the number of fatalities, and the number of people with impairments all drop by γ times, δ times, and ν times, respectively, in the compartment. Patients with uncomplicated diabetes are becoming more prevalent at the same time as twice as many people are recovering from complications.

Natural fatalities that occur in both compartments are also taken into consideration by the model; in this case, ν represents the mortality rates. Natural mortality reduces the population by ν times the current number in the compartment (C) for complications and by ν times the current number in compartment (D) for simple diabetes.

Furthermore, it is predicted that the number of people with simple diabetes will rise along with the total incidence of diabetes, which is indicated by the letter.

A proposed formulation of this DC model is as follows, according to [29]. Give the mathematical equations or expressions that match the formulation of the model as it appears in the cited publication. For a more in-depth explanation, you can insert any specific equations or expressions from the text here.

$$\frac{dD}{dt} = I - (\lambda - \nu)D + \gamma C \quad (1)$$

$$\frac{dC}{dt} = \lambda D - (\gamma + \delta + \nu + \mu)C \quad (2)$$

where > 0 and $\lambda, \gamma, \mu, \delta$, and $\nu > 0$.

Two different approaches were used in the examination of the DC model: long short-term memory (LSTM) technology and multiplicative Runge-Kutta. These techniques were evaluated and contrasted for their efficacy in modeling the dynamics of the comorbid diabetic population.

Multiplicative Runge Kutta

The fact that multiplicative calculus is limited to positive-valued functions of real variables is one of its drawbacks. Complex multiplicative calculus can effectively address this constraint, though. Uzer's groundbreaking work [22] provided an initial introduction to complex multiplicative calculus, which was further developed by Bashirov and Riza's in-depth mathematical analyses in [20] and [32]. The key motivation for expanding to the complex domain is the knowledge that the derivative is a local property. Functions can be transformed into complex-valued

functions of real variables to escape the restriction to positive values. This transformation also eliminates the need for the Cauchy-Riemann criteria, allowing the real and imaginary components to be clearly separated from one another.

$$\lim_{h \rightarrow 0} \frac{f(x+h)^{1/h}}{f(x)} \tag{3}$$

This definition becomes crucial since $f(x)$ is a positive function on the interval A and has a well-defined derivative at x . In these cases, the relationship between the multiplicative and classical derivatives can be expressed as follows:

$$f^*(x) = \exp\left(\frac{f'(x)}{f(x)}\right) \tag{4}$$

where $(\ln \circ f)(x) = \ln f(x)$.

Some basic rules of differentiation are:

$$\begin{aligned} (cf)^*(x) &= f^*(x) \\ (fg)^*(x) &= f^*(x)g^*(x) \\ \left(\frac{f}{g}\right)^*(x) &= \frac{f^*(x)}{g^*(x)} \\ (f^h)^* &= f^*(x)^{h(x)} \cdot f(x)^{h'(x)} \\ (foh)^*(x) &= f^*(h(x))^{h'(x)} \end{aligned} \tag{5}$$

The multiplicative Runge-Kutta techniques are applied in the above context, where c is a positive constant and functions f , g , and h are differentiable, to estimate solutions to the multiplicative differential equations of the following form:

$$y^*(x) = f(x, y), \quad y(x_0) = y_0 \tag{6}$$

An equivalent derivation of the ordinary Runge-Kutta (RK) techniques may be made for the multiplicative Runge-Kutta (MRK) methods. In particular, the fourth-order Runge-Kutta technique (RK4) is frequently used because it strikes a compromise between accuracy and simplicity. The Runge-Multiplicative obtained in [33]

$$\begin{aligned} y(x+h) &= y(x) \cdot f_0^{h/6} \cdot f_1^{h/3} \cdot f_2^{h/3} \cdot f_3^{h/6} \\ f_0 &= f(t, y) \\ f_1 &= f\left(t + \frac{h}{2}, yf_0^{h/2}\right) \end{aligned} \tag{7}$$

$$f_2 = f\left(t + \frac{h}{2}, yf_1^{h/2}\right) \tag{8}$$

$$f_3 = f(t+h, yf_2^h) \tag{9}$$

By using the properties of multiplicative calculus we have to change the form of DC-model to multiplicative by using equation (4). The multiplicative version of DC-Model will be:

$$\left(\frac{dD}{dt}\right)^* = \exp\left(\frac{dD/dt}{D}\right) = \exp\left(\frac{I - (\lambda - \nu)D + \gamma C}{D}\right) \tag{10}$$

$$\left(\frac{dC}{dt}\right)^* = \exp\left(\frac{dC/dt}{C}\right) = \exp\left(\frac{\lambda D - (\gamma + \delta + \nu + \mu)C}{C}\right) \tag{11}$$

Long Short-Term Memory (LSTM)

Long input sequences may be processed efficiently by recurrent neural networks (RNNs) containing feedback connections, as the Long Short-Term Memory (LSTM) architecture created by Hochreiter and Schmidhuber in 1997 [34]. A conventional LSTM system consists of an input gate, an output gate, a forget gate, and a cell. This enables it to manage the information flow via the gates and preserve values for arbitrary lengths of time [35]. This design is particularly well-suited for tasks requiring time series data because of its ability to handle delays in unforeseen events [36]. In many situations, LSTMs are preferable to RNNs, hidden Markov models, and other sequence learning approaches because they were specifically designed to address the vanishing gradient problem that typically happens during RNN training. Moreover, LSTMs are advantageous for handling sequences with varying length gaps [37].

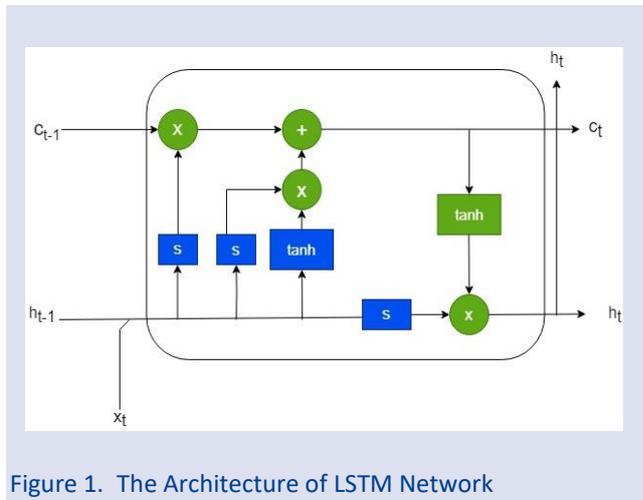


Figure 1. The Architecture of LSTM Network

Tensors of particular forms are essential to the construction of an LSTM, as seen in Figure 1, in order to facilitate effective information flow throughout the network. The dimensions of the cell state tensor (c_t) often show up as (batch size, num units), which indicates the number of LSTM units and the simultaneous processing of samples. Potential values for addition to the cell state are stored in the candidate cell state tensor, which has a form similar to the cell state tensor. Similarly, the output of the LSTM cell, the hidden state (h_t) tensor, is consistent in shape (batch size, num units) and guarantees smooth information transfer across the network. These tensor configurations are pivotal in facilitating the effective processing and retention of information across lengthy sequences, rendering LSTM networks highly suitable for diverse sequential data tasks.

The provided TensorFlow/Keras code describes a Sequential model architecture with two Long Short-Term Memory (LSTM) layers and a Dense layer, designed for sequence prediction tasks. Each LSTM layer contains 64 memory units and handles input sequences of shape (None, 2), offering flexibility in managing sequences of varying lengths while capturing complex temporal relationships. Efficient training and the introduction of non-linearity are crucial for learning intricate patterns. This can be achieved by using the 'relu' activation function, along with Glorot uniform initialization for both the kernel and recurrent weights. The way the returning sequences are configured in both LSTM layers makes it easier to generate output sequences at every time step, which improves the predictive power of the model. The Dense layer with two units then uses a linear activation function to provide predictions based on the sequences that have been analyzed. The model remains stable and consistent throughout thanks to the Glorot uniform initialization of kernel weights and the zero initialization of bias terms. This well designed architecture is a strong answer for a wide range of sequential data processing jobs because it makes use of the advantages of both the Dense layer's flexibility for precise predictions and the capabilities of LSTM layers for capturing temporal correlations. The offered code uses TensorFlow/Keras to assemble and train a sequential model. The mean squared error (MSE) loss function and the Adam optimizer are used to construct the model, and accuracy is included as a metric for assessment during training. Using a batch size of 32, the model is fitted to the training data for 200 epochs during the training phase. Furthermore, validation data is supplied to evaluate the model's performance on data that was not encountered during training. The 'history' variable contains the training history for further examination.

Table 1. Model Summary

| Layer (type) | Output Shape |
|--|------------------|
| LSTM | (None, None, 64) |
| LSTM | (None, None, 64) |
| Dense | (None, None, 2) |
| Total params | 50,306 |
| Trainable params Non-trainable params | 50,306 0 |

Two LSTM layers and a Dense layer make up the model. With 64 units in the output tensor, each LSTM layer's output form of (None, None, 64) denotes variable-length sequences. The output form of the Dense layer is (None, None, 2), which represents predictions in two units. The model has 50,306 total parameters, all of which may be trained. The model does not contain any non-trainable parameters.

Error Analysis and Model Comparison

This section conducts a comprehensive study of the model's performance, including a forecast for 992 time

steps and a full error analysis. Two key error measures are used to assess the efficacy of the models: Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). These metrics quantify the variation between the expected values and the actual values (xi). For MSE, average squared differences are computed, providing a measure of overall prediction accuracy. On the other hand, MAPE gives information about the relative accuracy by computing the percentage difference between the actual and projected values. In order to determine the specific benefits and drawbacks of the models, it is necessary to have a complete understanding of the predictive capability of the models, which is provided by both measurements.

The MSE formula and MAPE formula can be summarized as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x}_i)^2$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{x_i - \tilde{x}_i}{x_i} \right|$$

The predicted values are indicated by \tilde{x}_i , the actual values are denoted by x_i , and the total number of data points in the dataset is indicated by n. The error analysis that follows will offer a detailed comparison of the models and clarify how well each one captures the underlying patterns in the dataset.

Simulation and Comparison Results

The effectiveness of the Long Short-Term Memory (LSTM) networks and the conventional Runge-Kutta method of order 4 (RK4) has been assessed and contrasted with the Multiplicative Runge-Kutta method of order 4 (MRK4) utilizing three important error metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE).

The Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE) for both D(t) and C(t) were computed in order to assess how well the RK4 and MRK4 approaches replicated the dynamics of the diabetic population model.

The dynamics of two variables, D(t) and C(t), with their rates of change controlled by certain parameters, are modeled in this system of differential equations. The following is a definition of the parameters used in this model: With respect to $\lambda=0.1$, $\mu=0.2$, $\gamma=0.05$, $\delta=0.3$, and $\nu=0.1$, $I=1.0$ denotes a constant input. Important processes including growth, decay, and the relationships between the two variables are controlled by these factors. The system describes the evolution over time of D(t), which may correspond to a factor connected to diabetes, and C(t), which may represent a control mechanism. Since both variables begin at the same value, D (0) = 1.0 and C (0) = 1.0 are the beginning conditions. The simulation has a step size of 0.1 and runs from 1 to 200 units. The numerical solution to the issue is obtained using the 4th-order Runge-Kutta (RK4) approach, which computes the

development of $D(t)$ and $C(t)$ at each step in time. This approach helps us study the relationships and changes in the variables over time, giving us important information about the behavior of the system, given these initial conditions and the parameters that have been provided.

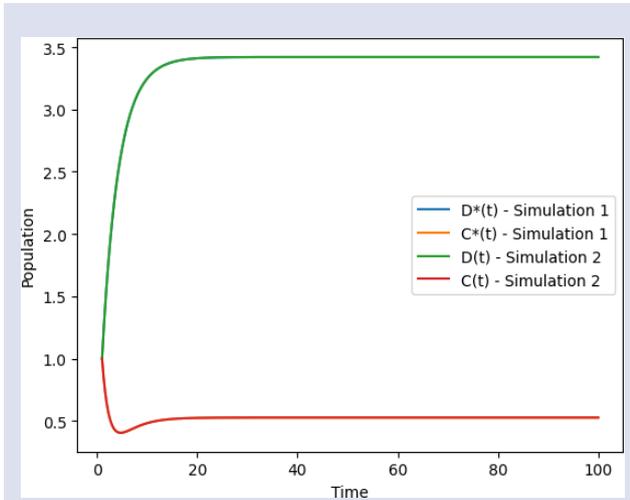


Figure 2. Comparison of MRK4 and RK4

These metrics are crucial indicators of how accurately and consistently each numerical method represents the model's behavior across time. Figure 2 offers critical benchmarks for assessing how accurate the simulations are in comparison to the actual data. This graph demonstrates exactly how MRK4 and RK4 work together and in a similar manner.

Table 2: Error Metrics for $D(t)$ and $C(t)$

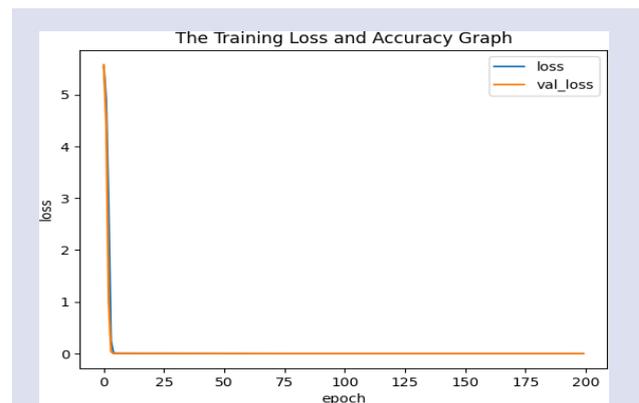
| Metric | $D(t)$ | $R(t)$ |
|--------|---------------------------|--------------------------|
| MSE | 6.9523×10^{-19} | 4.0598×10^{-18} |
| RMSE | 8.3381×10^{-10} | 2.0149×10^{-9} |
| MAE | 1.08997×10^{-10} | 3.5262×10^{-10} |

All of the error metrics (MSE, RMSE, and MAE) in Table 2 for both components D and C are extremely close to zero, suggesting that RK4 and MRK4 are essentially producing results that are almost similar with very little fluctuation. This might imply that MRK4 is a tweak that maintains accuracy on par with traditional RK4.

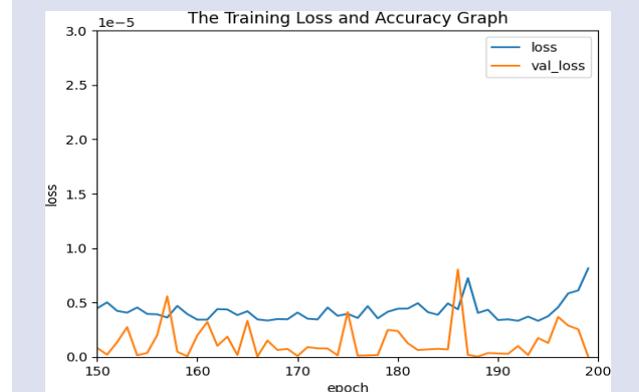
The model is configured using the mean squared error (MSE) as its loss function and leverages the well-known Adam optimizer for neural network optimization.

During a predetermined number of epochs (200 in this example), the training procedure involves regularly presenting the training set and its associated output to the model with a batch size of 32. Throughout this training process, vigilant monitoring is employed, and the progression of the validation loss as well as the training loss is graphically shown. An important tool for analyzing the convergence patterns and spotting any overfitting tendencies in the model is this chart.

Plotting the training loss and validation loss over the duration of epochs, the resulting graphical representation provides important insights into the learning dynamics of the model. Following the training phase, predictions are created on the validation set, allowing for a direct comparison of the projected values with the actual data. To completely evaluate the model's predictive capacity, standard evaluation metrics including mean squared error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE) are produced. When combined, these measures offer a quantitative assessment of the model's accuracy and give academics enlightening guidance on the reliability and efficacy of the predictions.



(a) Validation loss

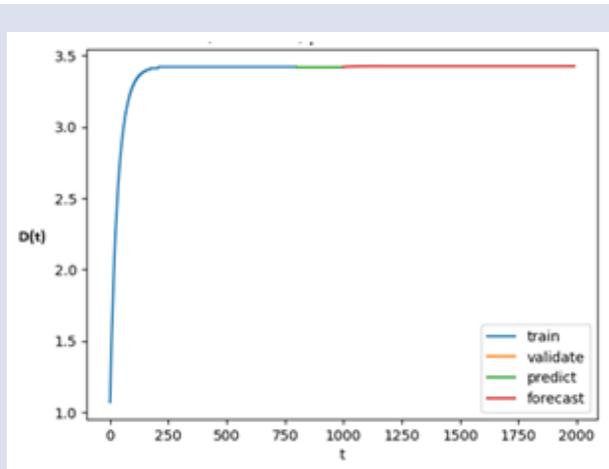


(b) Validation loss with zoom in

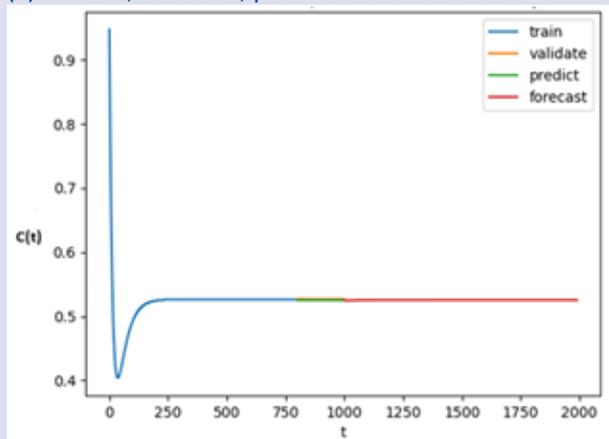
Figure 3. Validation loss during the training process

The validation loss's evolution during the training process is seen in Figure 3. The graph shows how the model's performance changed throughout several epochs on the validation set. The y-axis displays the corresponding loss values, while the x-axis represents the training epochs. Examining the validation loss's convergence over epochs is crucial since it tells us something about how effectively the model adapts to new data. A consistent decrease in the validation loss indicates better performance; oscillations or a plateau, on the other hand, can signal potential issues like overfitting or convergence issues. The model's learning phase is shown by the early decline in validation loss in Figure 3, and the

subsequent stability points to a balanced convergence. Researchers may determine how long to train the model for, what regularization strategies to employ, and whether the model is typically dependable in detecting the underlying patterns in the dataset by closely examining this graph.



(a) Trained, validated, predicted and forecasted for D



(b) Trained, validated, predicted and forecasted for C

Figure 4. Comparison Of LSTM and Multiplicative Runge-Kutta

Figures 4a and 4b provide a comprehensive comparison between the prediction performance of LSTM approaches and the numerical approximation for the system's state functions $D(t)$ and $C(t)$. The figure illustrates how effectively an LSTM model works when predicting system behavior using the MRK4 integration strategy. The blue line, which represents the training data, shows how the system's behavior evolved over the model's training phase and shows how well the LSTM was able to recognize the underlying dynamics. The yellow line represents the validation data, which was not used for training but helps determine how well the model can generalize to new data. The model's ability to predict the behavior of the system based on inputs from the validation or test set is demonstrated by the green line, which represents the predicted data. Last but not least, the LSTM model's estimate of the system's future behavior over a certain

period is shown by the red line, which is the anticipated data.

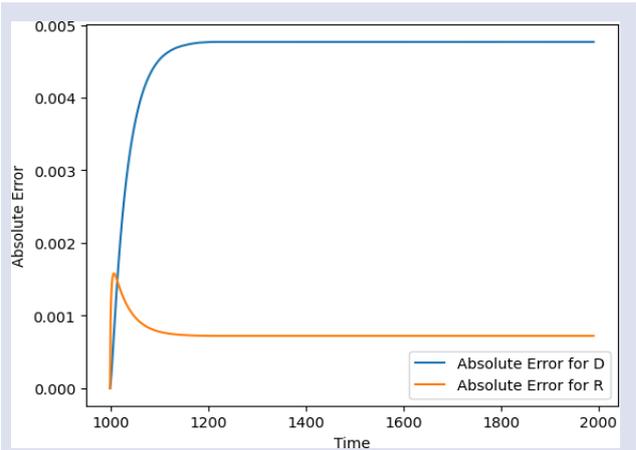


Figure 5. Absolute Error for Each Element

The effectiveness of the LSTM in predicting the behavior of the system when an MRK4 is present is seen in Figure 5. show the absolute error resulting from the LSTM prediction results using the numerical approximation of the state functions $D(t)$ and $C(t)$, respectively. The computation of these errors is based on the absolute difference between the matching data obtained from the LSTM algorithm and the numerical approximation data.

The numerical solution of our RC diabetes model, which serves as validation data, is compared with the absolute errors obtained via the Long Short-Term Memory (LSTM) technique in Figure 5. Particularly in terms of how well the LSTM performs on data that wasn't used for training, these errors serve as quantitative metrics that reveal information about the prediction accuracy of the model. The figures provide useful information on how effectively the LSTM approximates and captures the behavior of our RC diabetes model.

The absolute errors shown in Figure 4c may be used to calculate a quantitative evaluation of the differences between the deep learning approaches (LSTM) predictions and the actual numerical solution obtained with Multiplicative Runge-Kutta (MRK). The results of this comparative research show that the MRK-based numerical approach and the LSTM method accord quite well. The findings show a substantial relationship between the reliability of the LSTM model and its accuracy in capturing and reproducing the dynamics of the RC diabetes model.

Table 3: Performance Metrics

| Metric | Value |
|---------------------------------------|-----------------------|
| Mean Absolute Error (MAE) | 0.004594215327693641 |
| Mean Squared Error (MSE) | 2.151950235700345e-05 |
| Root Mean Squared Error (RMSE) | 0.004638911764304582 |
| Mean Absolute Percentage Error (MAPE) | 0.13433378151151 |

The performance evaluation of the model reveals positive results for several metrics. With mean absolute error (MAE) of 0.0046 and average deviations from true values averaging at a minimal scale, our model demonstrates a high degree of accuracy in forecasting values. The Mean Squared Error (MSE) of around 2.15×10^{-5} , which indicates that the forecasts exhibit minute squared errors, further demonstrates the correctness of the model.

The Root Mean Squared Error (RMSE) of 0.0046, which indicates how frequently projections differ from real data, supports these findings. Furthermore, the Mean Absolute Percentage Error (MAPE) of 0.1343 indicates an average percentage error of around 13.43%, which is rather low given the circumstances of our research but not insignificant. These results corroborate each other and show how well our model predicts outcomes. The final method employed is outstanding; it demonstrates an amazing ability to find intricate patterns in the data, leading to low error projections that are accurate.

Conclusion

In the end, this study thoroughly compared two distinct methodologies: multiplicative calculus and long short-term memory (LSTM) networks. We assessed the prediction performance of both methods using both quantitative analysis and close inspection. It was thoroughly examined if the proposed multiplicative calculus technique could extract complicated connections from the data, and the study looked into the learning and generalization capacities of LSTM networks. The comparison was facilitated by in-depth analyses of error metrics, including mean squared error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE). By comparing the benefits and drawbacks of LSTM networks and multiplicative calculus, we hoped to provide significant insights into the relative effectiveness of each in predictive modeling.

Conflict of interest

There are no conflicts of interest in this work.

Ethical Approval Statement

There is no Ethical Approval needed for this work.

References

- [1] Emerging Risk Factors Collaboration, Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies, *The lancet*, 375(9733) (2010) 2215-222
- [2] Ramsingh J., & Bhuvaneshwari V. (2021). An efficient map reduce-based hybrid NBC-TFIDF algorithm to mine the public sentiment on diabetes mellitus—a big data approach, *Journal of King Saud University-Computer and Information Sciences*, 33(8) 1018-1029.
- [3] Artzi N. S., Shilo, S., Hadar E., Rossman H., Barbash-Hazan S., Ben-Haroush A., ... , Segal E., Prediction of gestational diabetes based on nationwide electronic health records. *Nature medicine*, 26(1) (2020), 71-76.
- [4] Misra A., Gopalan H., Jayawardena R., Hills A. P., Soares M., Reza-Albarrán A. A., Ramaiya K. L., Diabetes in developing countries. *Journal of diabetes*, 11(7) (2019) 522-539.
- [5] Edwards M. S., Wilson D. B., Craven T. E., Stafford J., Fried L. F., Wong T. Y., ... , Hansen K. J., Associations between retinal microvascular abnormalities and declining renal function in the elderly population: the Cardiovascular Health Study. *American journal of kidney diseases*, 46(2) (2005) 214-224.
- [6] Saeedi P., Petersohn I., Salpea P., Malanda B., Karuranga S., Unwin N., et al., IDF Diabetes Atlas Committee Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, *Diabetes research and clinical practice*, 157 (2019) 107843.
- [7] Vaishali R., Sasikala R., Ramasubbareddy S., Remya S., Nalluri S Genetic algorithm based feature selection and MOE Fuzzy classification algorithm on Pima Indians Diabetes dataset. In *2017 international conference on computing networking and informatics (ICCNi)*, (2017) 1-5.
- [8] Cho N. H., Shaw J. E., Karuranga S., Huang Y., da Rocha Fernandes J. D., Ohlogge, A. W., Malanda, B. I. D. F., IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045, *Diabetes research and clinical practice*, 138 (2018) 271-281.
- [9] Maniruzzaman M., Rahman M. J., Al-Mehedi Hasan M., Suri, H. S., Abedin, M. M., El-Baz A., Suri J. S., Accurate diabetes risk stratification using machine learning: role of missing value and outliers. *Journal of medical systems*, 42 (2018) 1-17.
- [10] Brahim-Belhouari S., Bermak A., Gaussian process for nonstationary time series prediction, *Computational Statistics & Data Analysis*, 47(4) (2004) 705-712.
- [11] Cortes C., Vapnik V., Support-vector networks, *Machine learning*, 20 (1995) 273-297.
- [12] Kégl B., The return of AdaBoost. MH: multi-class Hamming trees, (2013) *arXiv preprint arXiv:1312.6086*.
- [13] Tabaei B. P., Herman W. H., A multivariate logistic regression equation to screen for diabetes: development and validation, *Diabetes Care*, 25(11) (2002) 1999-2003.
- [14] Jenhani I., Amor N. B., Elouedi Z., Decision trees as possibilistic classifiers, *International journal of approximate reasoning*, 48(3) (2008) 784-807.
- [15] Qawqzeh Y. K., Bajahzar A. S., Jemmali M., Otoom M. M., Thaljaoui A., Classification of diabetes using photoplethysmogram (PPG) waveform analysis: logistic regression modeling. *BioMed Research International*, (2020).
- [16] Pethunachiyar G. A., Classification of diabetes patients using kernel based support vector machines. In *2020 International Conference on Computer Communication and Informatics (ICCCI)*, (2020) 1-4).
- [17] Abdu-Allah Z. M., Mahmood O. T., AL-Naib A. M. I., Photovoltaic Battery Charging System Based on PIC16F877A Microcontroller, *International Journal of Engineering and Advanced Technology*, 3(4) (2014) 55-59.
- [18] Kumari S., Kumar D., Mittal M., An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *International Journal of Cognitive Computing in Engineering*, 2 (2021) 40-46.

- [19] Hussain A., Naaz S., Prediction of diabetes mellitus: comparative study of various machine learning models, In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020* Springer Singapore, 2 (2021) 103-115.
- [20] Hussain A., Naaz S., Prediction of diabetes mellitus: comparative study of various machine learning models, In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020* Springer Singapore, 2 (2020) 103-115.
- [21] Bashirov A. E., Mustafa R. I. Z. A., On complex multiplicative differentiation, *TWMS Journal of applied and engineering mathematics*, 1(1) (2011) 75-85.
- [22] Bashirov A. E., Norozpour S., On complex multiplicative integration, *TWMS Journal of Applied and Engineering Mathematics*, 7(1) (2017) 82-93.
- [23] Uzer A., Multiplicative type complex calculus as an alternative to the classical calculus, *Computers & Mathematics with Applications*, 60(10) (2010) 2725-2737.
- [24] Babacan Y., Kaçar F., Memristor emulator with spike-timing-dependent-plasticity, *AEU-International Journal of Electronics and Communications*, 73(2017) 16-22.
- [25] Bashirov A. E., Bashirova G., Dynamics of literary texts and diffusion, *Online Journal of Communication and Media Technologies*, 1(3) (2011) 60-82.
- [26] Wang Y., Liao X., Stability analysis of multimode oscillations in three coupled memristor-based circuits, *AEU-International Journal of Electronics and Communications*, 70(12) (2016) 1569-1579.
- [27] Aniszewska D., Multiplicative runge–kutta methods, *Nonlinear Dynamics*, 50(1-2) (2007) 265-272.
- [28] Aniszewska D., Multiplicative runge–kutta methods, *Nonlinear Dynamics*, 50(1-2) (2007) 265–272,.
- [29] Dileep P., Rao K. N., Bodapati P., Gokuruboyina S., Peddi R., Grover A., Sheetal A., An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm, *Neural Computing and Applications*, 35(10) (2023) 7253-7266.
- [30] Srinivasu P. N., SivaSai J. G., Ijaz M. F., Bho, A. K., Kim W., Kang J. J., Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM, *Sensors*, 21(8) (2021) 2852.
- [31] Dua M., Makhija D., Manasa P. Y. L., Mishra P., A CNN–RNN–LSTM based amalgamation for Alzheimer’s disease detection, *Journal of Medical and Biological Engineering*, 40(5) (2020) 688-706.
- [32] Bashirov A. E., Mustafa R. I. Z. A., On complex multiplicative differentiation, *TWMS Journal of applied and engineering mathematics*, 1(1) (2011) 75-85.
- [33] Riza M., Aktöre H., (The Runge–Kutta method in geometric multiplicative calculus, *LMS Journal of Computation and Mathematics*, 18(1) (2015) 539-554.
- [34] Hochreiter S., Schmidhuber J., Long short-term memory, *Neural computation*, 9(8) (1997) 1735-1780.
- [35] Gers F. A., Schmidhuber J., Cummins F., Continual prediction using LSTM with forget gates, In *Neural Nets WIRN Vietri-99: Proceedings of the 11th Italian Workshop on Neural Nets, Vietri Sul Mare, Salerno, Italy, 20–22 May 1999*, Springer London, (1999) 133-138.
- [36] Gers F. A., Schmidhuber J., Cummins F., Learning to forget: Continual prediction with LSTM, *Neural computation*, 12(10) (2000) 2451-2471.
- [37] Gers F. A., Schmidhuber E., LSTM recurrent networks learn simple context-free and context-sensitive languages, *IEEE transactions on neural networks*, 12(6) (2001) 1333-1340.



New accurate conservative finite difference schemes for 1-D and 2-D Schrödinger-Boussinesq Equations

Ayhan Aydın ^{1,a,*}, Taha Mohammed ^{2,3,b}

¹ Department of Mathematics, Faculty of Art and Science, Atılım University, 08830, Incek, Ankara, Türkiye.

² Graduate School of Natural and Applied Sciences, Atılım University, 08830, Incek, Ankara, Türkiye.

³ Al-Farahidi University, Department of Business Management, Faculty of Administration and Economics, Baghdad, Iraq.

*Corresponding author

Research Article

History

Received: 02/03/2024

Accepted: 12/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

In this paper, first-order and second-order accurate structure-preserving finite difference schemes are proposed for solving the Schrödinger- Boussinesq equations. The conservation of the discrete energy and mass of the present schemes are analytically proved. Numerical experiments are given to support the theoretical results. Numerical examples show the efficiency of the proposed scheme and the correction of the theoretical proofs.

Keywords: Conservative numerical methods, Partitioned average vector field method, Soliton solution, Schrödinger- Boussinesq equations.

ayhan.aydin@atilim.edu.tr

<https://orcid.org/0000-0002-0837-9364>

taha.y.mohammed@gmail.com <https://orcid.org/0000-0002-5679-9195>

Introduction

The coupled partial differential equations such as Schrödinger-KDV equations [1], N-coupled nonlinear Schrödinger equations [2], and the coupled Gross-Pitaevskii equations [3] play a crucial role in applied mathematics, engineering, and physics, many of which are the models of various nonlinear phenomena. In this paper, we consider the coupled nonlinear Schrödinger-Boussinesq (CNSB) equations

$$\begin{aligned} iu_t &= -\gamma\Delta u + \xi uv, & i^2 &= -1, \\ v_t &= \Delta\phi, \\ \phi_t &= -\alpha\Delta v + v + f(v) + \omega|u|^2 \end{aligned} \quad (1)$$

where the complex function $u(x, t)$ describes the electrical field of Langmuir oscillations, real valued function $v(x, t)$ represents low frequency density perturbation, $\mathbf{x} = (x_1, \dots, x_d)^T \in \Omega \subset \mathbb{R}^d, t = [0, T]$.

The parameters γ, ξ, ω and $\alpha > 0$ are constant, and $f(x)$ is sufficiently smooth real function with $f(0) = 0$. The equations (1) describe a coupling motion of acoustic and optical wave [4] and the dynamics behavior of Langmuir soliton formation [5]. We study the equations (1) with initial conditions

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad v(\mathbf{x}, 0) = v_0(\mathbf{x}), \quad \phi(\mathbf{x}, 0) = \phi_0(\mathbf{x}), \quad \mathbf{x} \in \Omega \quad (2)$$

$$\begin{aligned} \text{and } (l_1, \dots, l_d)\text{- periodic boundary conditions} \\ u(\mathbf{x}, t) &= u(x_1 + l_1, \dots, x_d, t), \dots, \\ u(\mathbf{x}, t) &= u(x_1, \dots, x_d + l_d, t) \\ v(\mathbf{x}, t) &= v(x_1 + l_1, \dots, x_d, t), \dots, \\ v(\mathbf{x}, t) &= v(x_1, \dots, x_d + l_d, t), \\ \phi(\mathbf{x}, t) &= \phi(x_1 + l_1, \dots, x_d, t), \dots, \\ \phi(\mathbf{x}, t) &= \phi(x_1, \dots, x_d + l_d, t), \end{aligned} \quad (3)$$

which makes long time integration possible. The spatial domain Ω is truncated on a bounded interval in one dimension ($d = 1$), a rectangle in two dimensions ($d = 2$) or a box in three dimensions ($d = 3$). One of the main properties of the periodic- initial- value problem (1)-(3) is conservation of mass (or the Langmuir plasmon number)

$$M(t) := \int_{\Omega} |u(\mathbf{x}, t)|^2 d\mathbf{x} \equiv M(0) \quad (4)$$

and the total energy

$$E(t) := \int_{\Omega} \left(\frac{2\gamma\omega}{\xi} |\nabla u|^2 + \alpha |\nabla v|^2 + |\nabla \phi|^2 + v^2 + 2F(v) + 2\omega v |u|^2 \right) d\mathbf{x} \equiv E(0) \quad (5)$$

where $F(v)$ is the primitive function of $f(v)$. The CNSB equations (1) are conservative systems. Therefore, proper discretization is required to reflect the conservation properties (4)-(5). In general, conservative numerical scheme exhibits better numerical performance on long time integration than a nonconservative one [12].

The CNSB equations (1) have been solved numerically by several authors. In [6], a multi-symplectic Hamiltonian formulation has been presented for the CNSB. Liao et al. proposed two conserved compact finite difference schemes for solving the nonlinear CNSB equations [20]. Bai et al. proposed a quadratic B-spline finite element scheme for the CNSB equations [7]. Zhang et al. studied the implicit conservative difference scheme and obtained an optimal error estimate [8]. Bai et al. studied the CNSB equations by the time-splitting Fourier spectral method [10]. All these numerical methods for the CNSB equations are obtained by constructing the

appropriate discretization. However, these methods for high dimensional CNSB equations ($d \geq 2$) are time-consuming and difficult since the resulting schemes are fully coupled and nonlinearly implicit. Few numerical methods are available in the literature for the numerical solution of multi-dimensional ($d \geq 2$) CNSB equations. Liao et al. proposed a time-splitting exponential wave integrator method based on Gautschi-type quadrature and Fourier pseudo-spectral discretization for solving Boussinesq-like equation and then presented the time-splitting Fourier spectral discretization for Schrödinger-like equation [11]. Two efficient compact finite difference schemes are introduced in [25] and [26]. By a combination of the Crank-Nicolson method and Leap-Frog scheme for temporal discretization and second-order centered difference scheme for spatial approximations, a linear Energy and Mass Preserving Finite Difference Method (EMP-FDM) is devised for 1D CSBEs in [27]. Cai et al. [19] developed second and fourth-order energy-preserving wavelet collocation schemes for CNSB equations based on the Hamiltonian structure and composition technique. The proposed method is energy-preserving but does not preserve mass. In many cases, the conservation of energy is more important than the conservation of mass. Recently, extensive structure-preserving numerical methods have been carried out for the numerical solution of dispersive/dissipative partial differential equations in the literature (see [28] and references therein). For instance, a fourth-order compact and energy-conservative difference scheme is proposed for solving the two-dimensional nonlinear Schrödinger equation, and convergence analysis of the method has been carried out in [29]. The numerical study of fractional differential

equations is also a challenging problem for many authors (see [30] and reference therein). For instance, The Schrodinger equation with the variable-order fractional operator has been solved numerically in [31]. In that study, an implicit fully discrete continuous Galerkin finite element method was developed to tackle this equation while the fractional operator was expressed with a nonsingular Mittag-Leffler kernel. Quispel et al. [13] developed the so-called Average Vector Field (AVF) method for ordinary differential equations (ODEs) which is energy preserving for the Hamiltonian vector field. Under most circumstances, the second-order AVF method yields a fully implicit numerical scheme that requires a nonlinear solver such as Newton’s iteration. Cai et al. [16] developed a more efficient AVF-based method and call this method a partitioned AVF (PAVF) method which can also automatically preserve arbitrary Hamiltonian energy of the Hamiltonian system. In [24] an explicit scheme has been developed for the Zakharov equation using the PAVF method. The main purpose of this study is to develop a new energy-preserving scheme for the CNSB equations (1) based on the PAVF method. Firstly, the equations (1) are written in an infinite Hamiltonian form. Then, the second-order central difference is employed for the spatial discretization to cast the CNSB equations (1) into a finite-dimensional Hamiltonian equation. We use the PAVF method for time integration to develop the energy-preserving scheme. The proposed scheme is semi-implicit which has a significant advantage over the AVF method. In conjunction with the adjoint method, we further present the PAVF composition (PAVF-C) method and the PAVF plus (PAVF-P) method for the CNSB equations (1).

Construction of a finite difference scheme

Many evolution PDEs can be written as an infinite-dimensional Hamiltonian system of the form [23]

$$\frac{dz}{dt} = \mathcal{D} \frac{\delta \mathcal{H}}{\delta z} \tag{6}$$

where $z = z(x, t) \in \mathbb{R}^d \times \mathbb{R}$ and \mathcal{D} is a constant linear differential operator. For example, if $d = 1$ and $z(x, t)$ belongs to the Hilbert space $L^2(\Omega)$, the Hamiltonian \mathcal{H} and the variational derivative is

$$\mathcal{H}(z) = \int_{\Omega} H(x, z, z_x, z_{xx}, \dots) dx, \tag{7}$$

$$\frac{\delta \mathcal{H}}{\delta z} = \frac{\partial H}{\partial z} + \partial_x \left(\frac{\partial H}{\partial z_x} \right) - \partial_{xx} \left(\frac{\partial H}{\partial z_{xx}} \right) - \dots \tag{8}$$

where $\Omega \subset \mathbb{R}$ and $z(x, t)$ satisfies homogenous or periodic boundary conditions. For two dimensions $d = 2$, i.e. $x = (x, y)^T, \Omega = [x_L, x_R] \times [y_L, y_R]$ the Hamiltonian \mathcal{H} and the variational derivative is

$$\mathcal{H}(z) = \iint_{\Omega} H(x, y, z, z_x, z_y, z_{xx}, \dots) dx dy, \tag{9}$$

$$\frac{\delta \mathcal{H}}{\delta z} = \frac{\partial H}{\partial z} + \partial_x \left(\frac{\partial H}{\partial z_x} \right) + \partial_y \left(\frac{\partial H}{\partial z_y} \right) - \partial_{xx} \left(\frac{\partial H}{\partial z_{xx}} \right) - \dots \tag{10}$$

By decomposing $u(x, t) = p(x, y, t) + iq(x, y, t)$ in real and imaginary components, the CNSB equations (1) in two space dimensions can be written as a first order system of equations

$$\begin{aligned} p_t &= -\gamma(q_{xx} + q_{yy}) + \xi q v \\ q_t &= \gamma(p_{xx} + p_{yy}) - \xi p v \\ v_t &= \phi_{xx} + \phi_{yy} \\ \phi_t &= v - \alpha(v_{xx} + v_{yy}) + f(v) + \omega(p^2 + q^2) \end{aligned} \tag{11}$$

which can be written as infinite-dimensional Hamiltonian system (6) with the state variable $z(t) = (p(x, y, t), q(x, y, t), v(x, y, t), \phi(x, y, t))^T$, the differential operator

$$\mathcal{D} = \begin{pmatrix} 0 & \xi/(4\omega) & 0 & 0 \\ -\xi/(4\omega) & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/2 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}$$

and the Hamiltonian

$$H(t) := \iint_{\Omega} \left(\frac{2\gamma\omega}{\xi} |\nabla u|^2 + \alpha |\nabla v|^2 + |\nabla \phi|^2 + v^2 + 2F(v) + 2\omega v |u|^2 \right) dx dy. \tag{12}$$

Remark: We can find that the Hamiltonian $H(t)$ is the total energy (5) of the system (1).

In the following sections, we propose a nonlinear two-level, conservative and second-order accurate finite difference scheme for the problem (1)-(3).

Spatial discretization

For simplicity of notations, we shall consider the CNSB system (1) in one space dimension. Extension to higher dimensions is straightforward. The domain $\{(x, t) | (x, t) \in \Omega \times [0, T]\}$ is discretized into grids by the set (x_j, t_n) of nodes, in which $x = x_L + jh, 0 \leq j \leq N + 1, t_n = n\tau, 0 \leq n \leq M$, where $h = (x_R - x_L)/N, \tau = T/M$, M and N are positive integers. Let p_j^n, q_j^n, v_j^n and ϕ_j^n be the approximations to $p(x, t), q(x, t), v(x, t)$ and $\phi(x, t)$ at the grid (x_j, t_n) , respectively. For convenience, we introduce some notations:

$$\begin{aligned} \delta_t z_j^n &= \frac{z_j^{n+1} - z_j^n}{\tau}, \quad \delta_x^+ z_j^n = \frac{z_{j+1}^n - z_j^n}{h}, \quad \delta_x^- z_j^n = \frac{z_j^n - z_{j-1}^n}{h}, \quad z_j^{n+\frac{1}{2}} = \frac{z_j^{n+1} + z_j^n}{2} \\ \delta_x^{+-} z_j^n &= \delta_x^+ (\delta_x^- z_j^n) = \frac{z_{j+1}^n - 2z_j^n + z_{j-1}^n}{h^2}, \end{aligned} \tag{13}$$

Discrete periodic boundary conditions are treated as

$$z_j = z_{N+j}, \quad j = \dots, -2, -1, 0, 1, 2, \dots$$

Consider the spatial discretization of the equations (11) in one space dimension

$$\begin{aligned} \frac{d}{dt} p_j &= -\gamma(\delta_x^+ q_j) + \xi q_j v_j, \\ \frac{d}{dt} q_j &= \gamma(\delta_x^- p_j) - \xi p_j v_j, \\ \frac{d}{dt} v_j &= \delta_x^{+-} \phi_j, \\ \frac{d}{dt} \phi_j &= -\alpha(\delta_x^{+-} v_j) + v_j + f(v_j) + \omega(p_j^2 + q_j^2) \end{aligned} \tag{14}$$

where $j = 1, \dots, N$. The system of equations (14) can be written as a finite-dimensional canonical Hamiltonian system

$$\frac{d}{dt} Z = \bar{\mathcal{D}} \nabla \bar{H}(Z) \tag{15}$$

where $Z = (p^T, q^T, v^T, \phi^T)^T, p, q, v, \phi \in \mathbb{R}^N$ with entries $p = (p_1, \dots, p_N)^T, q = (q_1, \dots, q_N)^T, v = (v_1, \dots, v_N)^T, \phi = (\phi_1, \dots, \phi_N)^T$, respectively,

$$\bar{\mathcal{D}} = \begin{pmatrix} 0 & \xi I & 0 & 0 \\ -\xi I & 4\omega I & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} I \\ 0 & 0 & \frac{1}{2} I & 0 \end{pmatrix} \tag{16}$$

where $I = N \times N$ identity matrix, 0 is the $N \times N$ zero matrix and

$$\bar{H} = \sum_{i=1}^N \left(\frac{2\gamma\omega}{\xi} ((\delta_x^+ p_j)^2 + (\delta_x^+ q_j)^2) + \alpha (\delta_x^+ v_j)^2 + (\delta_x^+ \phi_j)^2 + v_j^2 + 2F(v_j) + 2\omega v_j |u_j|^2 \right) h. \tag{17}$$

Is the discrete Hamiltonian. The system (15) conserves the discrete Hamiltonian (17) in the sense that

$$\frac{d\bar{H}(Z(t))}{dt} = \nabla \bar{H}(Z(t)) \frac{dZ(t)}{dt} = \nabla \bar{H}(Z(t)) \bar{\mathcal{D}} \nabla \bar{H}(Z(t)) = 0 \tag{18}$$

due to the skew-symmetric property of the matrix $\bar{\mathcal{D}}$. Therefore, the flow of the semi-discrete system (15) preserves the Hamiltonian $\bar{H}(Z)$ exactly. The Hamiltonian system (15) also possesses symplecticity [21, 22]. The conservation of energy is as important as the conservation of symplectic structure in numerical simulations. Therefore, it is natural to

integrate (15) in time with a conservative method. In the following section, energy preserving PAVF methods have been developed for (15).

Temporal discretization

Here we briefly discuss the AVF and PAVF methods and their energy-preserving properties. We consider the Hamiltonian differential equation (15) with $Z = (a, b)^T$. The canonical Hamiltonian system (15) can be written as

$$\frac{d}{dt} \begin{pmatrix} a \\ b \end{pmatrix} = \overline{\mathcal{D}} \begin{pmatrix} \nabla_a \overline{H}(a, b) \\ \nabla_b \overline{H}(a, b) \end{pmatrix} \tag{19}$$

The energy preserving AVF integrator [13] for canonical Hamiltonian equation (19) is given by

$$\frac{1}{\tau} \begin{pmatrix} a^{n+1} + a^n \\ b^{n+1} + b^n \end{pmatrix} = \overline{\mathcal{D}} \begin{pmatrix} \int_0^1 \nabla_a \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, \xi b^{n+1} + (1 - \xi)b^n) d\xi \\ \int_0^1 \nabla_b \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, \xi b^{n+1} + (1 - \xi)b^n) d\xi \end{pmatrix} \tag{20}$$

Theorem 1. [14] The scheme (20) is energy preserving, which possesses the discrete energy i.e.

$$\overline{H}(a^n, b^n) = \overline{H}(a^{n-1}, b^{n-1}) = \dots = \overline{H}(a^0, b^0) \tag{21}$$

Proof. Taking the scalar product on both sides of (20) with

$$\left(\int_0^1 \nabla_a \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, \xi b^{n+1} + (1 - \xi)b^n)^T d\xi, \int_0^1 \nabla_b \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, \xi b^{n+1} + (1 - \xi)b^n)^T d\xi \right)^T$$

using the Fundamental Theorem of Calculus and the skew-symmetry of $\overline{\mathcal{D}}$, we have

$$\frac{1}{\tau} (\overline{H}(a^{n+1}, b^{n+1}) - \overline{H}(a^n, b^n)) = 0$$

that is, the energy \overline{H} is precisely conserved at every time step. This completes the proof.

For polynomial Hamiltonian, the integral can be evaluated exactly, and the implementation is comparable to that of implicit Runge-Kutta method such as the implicit mid-point rule. When Hamiltonian energy is a quadratic function, the resulting AVF scheme is linearly implicit and therefore can be efficiently solved. However, this is not the case to reflect the merit of the AVF method since any symplectic Runge-Kutta method can also achieve the energy conservation of the quadratic Hamiltonian [12]. Under most circumstances, the evaluation of the integration in (20) leads to nonlinear function of Z^{n+1} which further constitutes a fully implicit numerical scheme. The iterative processes are then inevitably required but this leads to an increase in computation complexity, especially for the application of Hamiltonian PDEs. Cai [16] et. al. defined the so-called PAVF method for the Hamiltonian system (19) which has a remarkable advantage over the AVF method (20). The one-step, first-order PAVF method for the Hamiltonian system (19) is written as

$$\frac{1}{\tau} \begin{pmatrix} a^{n+1} + a^n \\ b^{n+1} + b^n \end{pmatrix} = \overline{\mathcal{D}} \begin{pmatrix} \int_0^1 \nabla_a \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, b^n) d\xi \\ \int_0^1 \nabla_b \overline{H}(a^{n+1}, \xi b^{n+1} + (1 - \xi)b^n) d\xi \end{pmatrix} \tag{22}$$

Theorem 2. [14] The scheme (22) is energy preserving in the sense that

$$\overline{H}(a^n, b^n) = \overline{H}(a^{n-1}, b^{n-1}) = \dots = \overline{H}(a^0, b^0) \tag{23}$$

Proof. Taking the scalar product on both sides of (22) with

$$\left(\int_0^1 \nabla_a \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, b^n) d\xi, \int_0^1 \nabla_b \overline{H}(a^{n+1}, \xi b^{n+1} + (1 - \xi)b^n) d\xi \right)^T$$

right-hand side of (22) vanishes by the skew-symmetry of $\overline{\mathcal{D}}$. The left-hand side of (22) can be written as

$$\frac{1}{\tau} \left(\int_0^1 \nabla_a \overline{H}(\xi a^{n+1} + (1 - \xi)a^n, b^n) d\xi \right) (a^{n+1} + a^n) + \frac{1}{\tau} \left(\int_0^1 \nabla_b \overline{H}(a^{n+1}, \xi b^{n+1} + (1 - \xi)b^n) d\xi \right) (b^{n+1} + b^n) = 0$$

which can be written as

$$\frac{1}{\tau} \int_0^1 \frac{d}{d\xi} [\overline{H}(\xi a^{n+1} + (1 - \xi)a^n, b^n) + \overline{H}(a^{n+1}, \xi b^{n+1} + (1 - \xi)b^n)] d\xi = 0.$$

Using the Fundamental Theorem of Calculus, we have

$$\frac{1}{\tau} [\bar{H}(a^{n+1}, b^n) - \bar{H}(a^n, b^n) + \bar{H}(a^{n+1}, b^{n+1}) - \bar{H}(a^{n+1}, b^n)] = 0 \text{ i.e. } \frac{1}{\tau} [\bar{H}(a^{n+1}, b^{n+1}) - \bar{H}(a^n, b^n)] = 0.$$

This completes the proof.

Definition 1. [12] The adjoint Φ_τ^* of the method Φ_τ is the inverse map of the original method with reversed time step $-\tau$.

The adjoint of the explicit Euler method is the implicit Euler method. The adjoint method of the implicit mid-point rule is the implicit mid-point rule itself, that is the implicit mid-point rule is symmetric. Accordingly, reversing the path order in (22), the adjoint method of the PAVF method (22) is written as [16]

$$\frac{1}{\tau} \begin{pmatrix} a^{n+1} + a^n \\ b^{n+1} + b^n \end{pmatrix} = J \begin{pmatrix} \int_0^1 \nabla_a \bar{H}(\xi a^{n+1} + (1 - \xi)a^n, b^{n+1}) d\xi \\ \int_0^1 \nabla_b \bar{H}(a^n, \xi b^{n+1} + (1 - \xi)b^n) d\xi \end{pmatrix} \tag{24}$$

The adjoint method has the same order as the original method. In the following, we consider the composition of the PAVF method (22) and the adjoint method (24) with the same step size. This allows an order increase for the old order method. If Φ_τ is the method (22) and Φ_τ^* is the adjoint method (24) then the composition method [16]

$$\Phi_h = \Phi_{\frac{\tau}{2}} \circ \Phi_{\frac{\tau}{2}}^* \tag{25}$$

is the second order method and conserves the Hamiltonian exactly. Analogously, taking the average of the Φ_τ and Φ_τ^* , we can write the plus method [16]

$$\Psi_h = \frac{1}{2} (\Phi_\tau + \Phi_\tau^*) \tag{26}$$

Average Vector Field for CNSB Equations

We first present the conventional second-order AVF method (20) for CNSB equations (1) [18]. To obtain the AVF integrator, we apply the AVF method to time integration for the semi-discrete system (15) as follows.

$$\begin{aligned} \delta_t (p_j^n) &= -\gamma \delta_x^{+-} \left(q_j^{n+\frac{1}{2}} \right) + \xi \left(\frac{1}{3} q_j^{n+1} v_j^{n+1} + \frac{1}{6} q_j^{n+1} v_j^n + \frac{1}{6} q_j^n v_j^{n+1} + \frac{1}{3} q_j^n v_j^n \right), \\ \delta_t (q_j^n) &= \gamma \delta_x^{+-} (p_j^{n+1/2}) - \xi \left(\frac{1}{3} p_j^{n+1} v_j^{n+1} + \frac{1}{6} p_j^{n+1} v_j^n + \frac{1}{6} p_j^n v_j^{n+1} + \frac{1}{3} p_j^n v_j^n \right), \\ \delta_t (v_j^n) &= \delta_x^{+-} (\phi_j^{n+1/2}) \\ \delta_t (\phi_j^n) &= -\alpha \delta_x^{+-} (v_j^{n+1/2}) + v_j^{n+1/2} + \tilde{f}(v_j^{n+1}, v_j^n) + \frac{\omega}{3} \left(\begin{aligned} &(p_j^{n+1})^2 + p_j^{n+1} p_j^n + (p_j^n)^2 \\ &+ (q_j^{n+1})^2 + q_j^{n+1} q_j^n + (q_j^n)^2 \end{aligned} \right). \end{aligned} \tag{27}$$

where $\tilde{f}(v_j^{n+1}, v_j^n) = \int_0^1 f(\xi v_j^{n+1} + (1 - \xi)v_j^n) d\xi$. We see that the scheme (27) is fully implicit, which requires a time consuming iterative method such as Newton-Raphson method. Now, we propose a more efficient energy-preserving method for CNSB equations (1).

Partitioned Average Vector Field for CNSB Equations

Upon applying the PAVF method (22) to the semi-discrete system (15), we obtain

$$\begin{aligned} \delta_t (p_j^n) &= -\gamma \delta_x^\pm \left(q_j^{n+\frac{1}{2}} \right) + \frac{\xi}{2} (q_j^{n+1} v_j^n + q_j^n v_j^n), \\ \delta_t (q_j^n) &= \gamma \delta_x^{+-} (p_j^{n+1/2}) - \frac{\xi}{2} (p_j^{n+1} v_j^n + p_j^n v_j^n), \\ \delta_t (v_j^n) &= \delta_x^{+-} (\phi_j^{n+1/2}) \\ \delta_t (\phi_j^n) &= -\alpha \delta_x^{+-} \left(v_j^{n+\frac{1}{2}} \right) + v_j^{n+1/2} + \tilde{f}(v_j^{n+1}, v_j^n) + \omega ((p_j^{n+1})^2 + (q_j^{n+1})^2) \end{aligned} \tag{28}$$

The PAVF method (28) is simpler than the AVF method (27). The first two equations of (28) are linearly implicit according to the variables p^{n+1} and q^{n+1} . Once the values p_j^n, q_j^n, v_j^n and ϕ_j^n are known, the values p^{n+1} and q^{n+1} can be solved from the first two equations of (28) and substituted into the last two equations. Then the values ϕ_j^{n+1} and v_j^{n+1} can be solved from the last two equations by using an iterative method such as Newton’s method. Although the PAVF method (28) for the CNSB is semi-implicit, the AVF method (27) is fully implicit which requires more cost per time step than PAVF method (28).

Remark: On the contrary to the semi-implicit structure of the PAVF method (28) for CNSB equations, Cai et. al. [16] have constructed the PAVF method to nonlinear Klien-Gordon-Schrödinger equations that requires to solve two sets of linear algebraic equations.

In addition, to the conservation of energy, the conservation of mass (4), which is quadratic invariant, also plays an important role in physics. Therefore, it is natural to discuss the mass conservation of the method (28).

Theorem 1. The difference scheme (28) is conservative, that is the mass (4) is conserved in the sense

$$M^n = \sum_{k=1}^N ((p_k^n)^2 + (q_k^n)^2)h = M^{n-1} = \dots = M^0, \tag{29}$$

where $M^n = M(p^n, q^n)$.

Proof. Multiplying both sides of the first line of equation (28) with $(p_j^{n+1} + p_j^n)^T$ yields

$$\frac{1}{\tau} (p_j^{n+1} + p_j^n)^T (p_j^{n+1} - p_j^n) = (p_j^{n+1} + p_j^n)^T \left(-\gamma \delta_x^{+-} q_j^n + \frac{1}{2} (q_j^{n+1} v_j^n + q_j^n v_j^n) \right) \tag{30}$$

Multiplying both sides of the second line of equation (28) with $(q_j^{n+1} + q_j^n)^T$ yields

$$\frac{1}{\tau} (q_j^{n+1} + q_j^n)^T (q_j^{n+1} - q_j^n) = (q_j^{n+1} + q_j^n)^T \left(\gamma \delta_x^{+-} p_j^n - \frac{1}{2} (p_j^{n+1} v_j^n + p_j^n v_j^n) \right). \tag{31}$$

Adding (30) to (31), we have

$$\frac{1}{\tau} \sum_{i=1}^N \left((p_j^{n+1})^2 + (q_j^{n+1})^2 \right) = \frac{1}{\tau} \sum_{i=1}^N \left((p_j^n)^2 + (q_j^n)^2 \right)$$

which completes the proof.

The adjoint PAVF method of the scheme (27) is given as

$$\begin{aligned} \delta_t (p_j^n) &= -\gamma \delta_x^{+-} \left(q_j^{n+\frac{1}{2}} \right) + \frac{\xi}{2} (q_j^{n+1} v_j^{n+1} + q_j^n v_j^{n+1}), \\ \delta_t (q_j^n) &= \gamma \delta_x^{+-} (p_j^{n+1/2}) - \frac{\xi}{2} (p_j^{n+1} v_j^{n+1} + p_j^n v_j^{n+1}), \\ \delta_t (v_j^n) &= \delta_x^{+-} (\phi_j^{n+1/2}) \\ \delta_t (\phi_j^n)_t &= -\alpha \delta_x^{+-} (v_j^{n+1/2}) + v_j^{n+1/2} + \tilde{f}(v_j^{n+1}, v_j^n) + \omega \left((p_j^n)^2 + (q_j^n)^2 \right). \end{aligned} \tag{32}$$

We can see that the adjoint scheme (32) is semi-implicit due to the nonlinear term $\tilde{f}(v_j^{n+1}, v_j^n)$. The values v_j^{n+1} and ϕ_j^{n+1} can be solved from the last two nonlinear equations of (32) using an iterative method. Then v_j^{n+1} and ϕ_j^{n+1} can be substituted into the first two equations of (32) and p_j^{n+1} and q_j^{n+1} can be obtained. Using the PAVF method (28) with the adjoint method (32), we can write the energy-preserving composition (PAVF-C) method (25) as follows:

$$\begin{aligned} \frac{2}{\tau} (p_j^* - p_j^n) &= -\gamma \delta_x^{+-} (q_j^* + q_j^n) + \xi (q_j^* v_j^n + q_j^n v_j^n) \\ \frac{2}{\tau} (q_j^* - q_j^n) &= \gamma \delta_x^{+-} (p_j^* + p_j^n) - \xi (p_j^* v_j^n + p_j^n v_j^n) \\ \frac{2}{\tau} (v_j^* - v_j^n) &= \delta_x^{+-} (\phi_j^* + \phi_j^n) \\ \frac{2}{\tau} (\phi_j^* - \phi_j^n) &= -\alpha \delta_x^{+-} (v_j^* + v_j^n) + (v_j^* + v_j^n) + 2\tilde{f}(v_j^*, v_j^n) + 2\omega \left((p_j^*)^2 + (q_j^*)^2 \right) \\ \frac{2}{\tau} (p_j^{n+1} - p_j^*) &= -\gamma \delta_x^{+-} (q_j^{n+1} + q_j^*) + \xi (q_j^{n+1} v_j^{n+1} + q_j^* v_j^{n+1}) \\ \frac{2}{\tau} (q_j^{n+1} - q_j^*) &= \delta_x^{+-} (p_j^{n+1} + p_j^*) - \xi (p_j^{n+1} v_j^{n+1} + p_j^* v_j^{n+1}) \\ \frac{2}{\tau} (v_j^{n+1} - v_j^*) &= \delta_x^{+-} (\phi_j^{n+1} + \phi_j^*) \\ \frac{2}{\tau} (\phi_j^{n+1} - \phi_j^*) &= -\alpha \delta_x^{+-} (v_j^{n+1} + v_j^*) + (v_j^{n+1} + v_j^*) + 2\tilde{f}(v_j^{n+1}, v_j^*) + 2\omega \left((p_j^*)^2 + (q_j^*)^2 \right) \end{aligned} \tag{33}$$

We can see that the PAVF-C method (33) inherits the semi-implicit property. If the values p_j^n, q_j^n, v_j^n and ϕ_j^n are known, the values p_j^* and q_j^* can be solved from the first two equations of (33) and substituted into the third and fourth equations which are nonlinear in terms of the unknowns ϕ_j^* and v_j^* . Thus, ϕ_j^* and v_j^* can be obtained by using an iterative method. Here Newton’s method is used as a nonlinear solver. Then, p_j^*, q_j^*, v_j^* and ϕ_j^* can be substituted into the last two equations in (33) and ϕ_j^{n+1} and v_j^{n+1} can be obtained by using Newton’s method. Finally, ϕ_j^{n+1} and v_j^{n+1} can be substituted into the fifth and sixth equations and p_j^{n+1} and q_j^{n+1} can be obtained by using the semi-implicit property of these two equations.

With the adjoint scheme (32), we can write down the corresponding plus scheme (PAVF-P)

$$\delta_t (p_j^n) = -\gamma \delta_x^{+-} (q_j^n) - \frac{\xi}{4} (q_j^{n+1} v_j^n + q_j^n v_j^n + q_j^{n+1} v_j^{n+1} + q_j^n v_j^{n+1})$$

$$\begin{aligned}
 \delta_t (q_j^n) &= \gamma \delta_x^{+-} (p_j^n) - \frac{\xi}{4} (p_j^{n+1} v_j^n + p_j^n v_j^n + p_j^{n+1} v_j^{n+1} + p_j^n v_j^{n+1}), \\
 \delta_t (v_j^n) &= \delta_x^{+-} (\phi_j^n), \\
 \delta_t (\phi_j^n) &= -\alpha \delta_x^{+-} (v_j^n) + v_j^{n+1/2} + \tilde{f}(v_j^{n+1}, v_j^n) + \frac{\omega}{2} ((p_j^{n+1})^2 + (q_j^{n+1})^2 + (p_j^n)^2 + (q_j^n)^2)
 \end{aligned} \tag{34}$$

The scheme (34) is implicit and nonlinear. To obtain the numerical solution $p_j^{n+1}, q_j^{n+1}, v_j^{n+1}$ and ϕ_j^{n+1} in (34), we need an iterative algorithm which increases the computational time. Here Newton’s method is used as a nonlinear solver.

Numerical Results

In this section, we present some numerical experiments to verify the theoretical results and to demonstrate the effectiveness of the conservative schemes (28), (33), (34). $L_\infty^n(\tau)$ and $L_2^n(\tau)$ errors for temporal accuracy is defined by

$$\begin{aligned}
 L_\infty^n &= \max_{1 \leq j \leq N} \{|u(x_j, t_n) - u_j^n| + |v(x_j, t_n) - v_j^n|\}, \\
 L_2^n &= (h \sum_{j=1}^N |u(x_j, t_n) - u_j^n|^2)^{1/2} + (h \sum_{j=1}^N |v(x_j, t_n) - v_j^n|^2)^{1/2}
 \end{aligned} \tag{35}$$

L_∞^j and L_2^j errors for spatial accuracy can be defined analogously. The preservation of the energy and the mass are monitored by the relative errors

$$GE = \frac{|H^n - H^0|}{|H^0|}, \quad GM = \frac{|M^n - M^0|}{|M^0|}$$

where

$$H^n = \sum_{i=1}^N \left(\frac{2\gamma\omega}{\xi} ((\delta_x^+ p_j^n)^2 + (\delta_x^+ q_j^n)^2) + \alpha (\delta_x^+ v_j^n)^2 + (\delta_x^+ \phi_j^n)^2 + (v_j^n)^2 + 2F(v_j^n) + 2\omega v_j^n |u_j^n|^2 \right) h.$$

denotes the discrete Hamiltonian corresponding to (17), and M^n is the mass (29) evaluated at $t = t^n$. The rate of convergence in time discretization is obtained by using

$$order \approx \log(L(\tau_1)/L(\tau_2))/\log(\tau_1/\tau_2)$$

where L represents $L_\infty^n(\tau)$ and $L_2^n(\tau)$ errors at the time steps τ_1 and τ_2 . The rate of convergence in space discretization can be defined analogously. To demonstrate the long-time behavior of the energy-preserving schemes, we take periodic boundary conditions. For simplicity, we present the wave profiles of the PAVF scheme (28) in all computations, since the other schemes AVF (27), PAVF-C (33), and PAVF-P (34) produce the same profile.

One-dimensional CSB equation

In this section, we report some numerical results to exhibit the performance of the schemes (28), (33) and (34) and verify the energy and the mass conservations. Performance at the PAVF schemes (28), (33) and (34) are compared with the fully implicit AVF scheme (27) [18]. An analytical solution has been given in [17].

$$\begin{aligned}
 u(x, t) &= \frac{9}{10} \operatorname{sech}^2 \left(\frac{\sqrt{15}}{10} \left(x - \frac{\sqrt{10}}{5} t \right) \right) \exp \left(i \left(\frac{\sqrt{10}}{10} x + \frac{1}{2} t \right) \right), \\
 v(x, t) &= \frac{9}{10} \operatorname{sech}^2 \left(\frac{\sqrt{15}}{10} \left(x - \frac{\sqrt{10}}{5} t \right) \right), \\
 v_t(x, t) &= \frac{9}{10} \operatorname{sech}^2 \left(\frac{\sqrt{15}}{10} \left(x - \frac{\sqrt{10}}{5} t \right) \right) \tanh \left(\frac{\sqrt{15}}{10} \left(x - \frac{\sqrt{10}}{5} t \right) \right),
 \end{aligned} \tag{36}$$

Accuracy test: First, the correctness of the numerical schemes is examined. The spatial domain has been chosen large enough that solitary wave propagation does not affect the propagation of the wave. The initial conditions are taken from the exact solution (36).

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x), \quad v_t(x, 0) = v_{t,0}(x), \tag{37}$$

Table 1 represents L_∞^j and L_2^j errors and convergence order in space. We note that all methods are both of second order in space.

Table 1. The L_∞^j and L_2^j errors and convergence orders in space of the proposed methods with $M = 4000$, and $-16 \leq x \leq 16$ at $t = 1$.

| | | h | 0.5 | 0.25 | 0.125 |
|----------|--------------|-------|------------------------|------------------------|------------------------|
| AVF [18] | L_∞^j | Error | 1.174×10^{-2} | 3.014×10^{-3} | 7.508×10^{-4} |
| | | Order | / | 1.96 | 2.00 |
| | L_2^j | Error | 1.739×10^{-2} | 4.371×10^{-3} | 1.094×10^{-4} |
| | | order | / | 1.99 | 2.00 |
| PAVF | L_2^j | Error | 1.175×10^{-2} | 3.030×10^{-3} | 7.687×10^{-4} |
| | | Order | / | 1.95 | 2.00 |
| | L_2^j | Error | 1.741×10^{-2} | 4.392×10^{-3} | 1.101×10^{-4} |
| | | order | / | 1.99 | 2.00 |
| PAVF-C | L_∞^j | Error | 1.174×10^{-2} | 3.014×10^{-3} | 7.527×10^{-4} |
| | | Order | / | 1.96 | 2.00 |
| | L_2^j | Error | 1.135×10^{-2} | 4.371×10^{-3} | 1.094×10^{-3} |
| | | order | / | 1.99 | 2.00 |
| PAVF-P | L_∞^j | Error | 7.501×10^{-3} | 1.892×10^{-3} | 4.730×10^{-4} |
| | | Order | / | 1.99 | 2.00 |
| | L_2^j | Error | 4.240×10^{-3} | 1.122×10^{-3} | 2.796×10^{-4} |
| | | order | / | 1.91 | 2.00 |

Table 2 lists the L_∞^n and L_2^n errors and convergence order in time. From the table, we see that all methods reach the second order in time except the PAVF method which is only first-order. From Table 1 and Table 2, we can see that the PAVF-P method has higher accuracy than the PAVF and PAVF-C methods. Table 3 and Figure 1 represent the CPU time of the four methods with different temporal steps. All computations were done on a custom computer with i7-1.80 GHz. To estimate how long a portion of our algorithm takes to run we used the Matlab

stopwatch timer functions, tic and toc. The Matlab (R2024b) built-in functions tic and toc are used to measure the performance of the algorithms. These functions return wall-clock time. Since the PAVF method and the PAVF-C method are semi-implicit, they require a solution of two linear systems of equations as well as the solution of two nonlinear systems of equations. On the other hand, fully implicit PAVF-P and AVF methods require the solution of four nonlinear systems

Table 2. The L_∞^n and L_2^n errors and convergence orders in time of the proposed methods with $h = 0.01$, $-16 \leq x \leq 16$ at $t = 1$.

| | | τ | 0.5 | 0.25 | 0.125 |
|----------|--------------|--------|------------------------|------------------------|------------------------|
| AVF [18] | L_∞^n | Error | 2.889×10^{-3} | 7.274×10^{-4} | 1.845×10^{-4} |
| | | Order | / | 1.99 | 1.98 |
| | L_2^n | Error | 4.902×10^{-3} | 1.243×10^{-3} | 3.210×10^{-4} |
| | | order | / | 1.98 | 1.95 |
| PAVF | L_∞^n | Error | 1.966×10^{-2} | 1.061×10^{-2} | 5.393×10^{-3} |
| | | Order | / | 0.89 | 0.98 |
| | L_2^n | Error | 3.238×10^{-2} | 1.725×10^{-2} | 8.788×10^{-3} |
| | | order | / | 0.91 | 0.97 |
| PAVF-C | L_∞^n | Error | 2.078×10^{-3} | 5.116×10^{-4} | 1.285×10^{-4} |
| | | Order | / | 2.02 | 1.99 |
| | L_2^n | Error | 3.456×10^{-3} | 8.685×10^{-4} | 2.316×10^{-4} |
| | | order | / | 1.99 | 1.91 |
| PAVF-P | L_∞^n | Error | 2.086×10^{-3} | 5.409×10^{-4} | 1.397×10^{-4} |
| | | Order | / | 1.99 | 2.00 |
| | L_2^n | Error | 3.277×10^{-3} | 8.441×10^{-4} | 2.228×10^{-4} |
| | | order | / | 1.96 | 1.92 |

Table 3. Computation time for $-32 \leq x \leq 32$ at $t = 1$.

| τ | PAVF | | PAVF-C | | PAVF-P | | AVF [18] | |
|--------|-----------|------------|-----------|------------|-----------|------------|-----------|------------|
| | $h = 0.1$ | $h = 0.05$ | $h = 0.1$ | $h = 0.05$ | $h = 0.1$ | $h = 0.05$ | $h = 0.1$ | $h = 0.05$ |
| 0.1 | 0.083 | 0.137 | 0.129 | 0.239 | 0.165 | 0.350 | 0.249 | 0.405 |
| 0.05 | 0.121 | 0.240 | 0.217 | 0.263 | 0.220 | 0.479 | 0.240 | 0.509 |
| 0.01 | 0.509 | 1.138 | 0.994 | 2.206 | 0.987 | 2.087 | 1.097 | 2.218 |
| 0.005 | 1.002 | 3.123 | 2.027 | 6.132 | 1.995 | 5.528 | 2.026 | 4.590 |

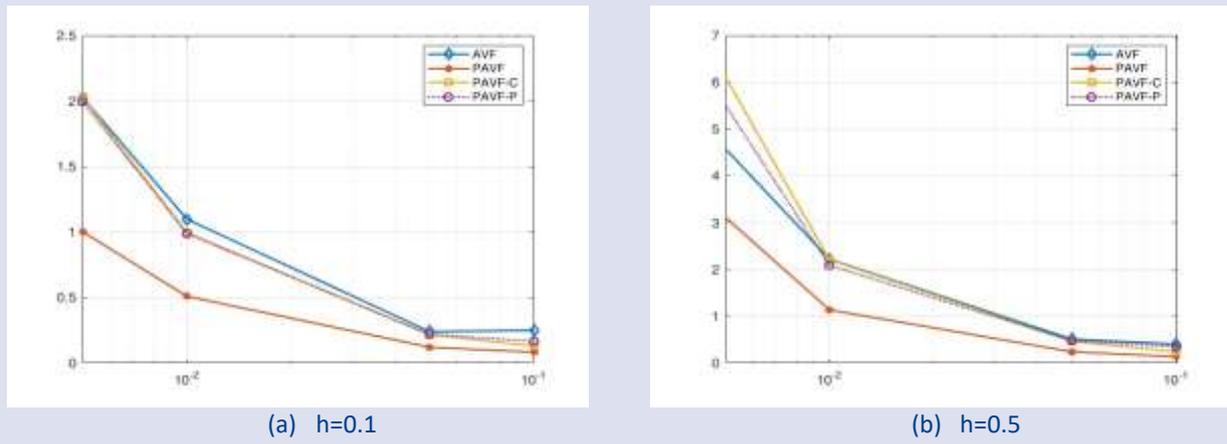


Figure 1. Computation time for $-32 \leq x \leq 32$ at $t=1$.

of equations. Accordingly, the PAVF method is more efficient than the AVF method. Figure 1 verifies this fact. Table 3 also shows that the computational costs of PAVF and PAVF-C are less than the fully implicit PAVF-P and AVF methods. In addition, we notice that the computational times of the PAVF-C and PAVF-P methods are slightly less than twice that of PAVF although the PAVF-C method is the composition of the PAVF method and its adjoint method, and the PAVF-P method is fully implicit. In addition, the PAVF-C and the PAVF-P methods have the same accuracy as the AVF method, but they are more efficient than the AVF method. Finally, we can say that the PAVF-P is the most efficient method.

Next, we test the stability of the scheme concerning the initial data. We consider the perturbed initial data

$u_{noise}(x, 0) = u_0(x)(1 + \mu\theta)$, $v_{noise}(x, 0) = v_0(x)(1 + \mu\theta)$, $v_{t,noise}(x, 0) = v_{t,0}(x)(1 + \mu\theta)$, where μ is the percentage of the noise and θ is the random variable generated from a uniform distribution in the interval $[-0.5, 0.5]$. We denote the perturbed solution as U_p and V_p . Figure 2 represents the solitary wave obtained by the PAVF scheme (28) with noise using the perturbed initial condition with $\mu = 0.2$. Similar results have been obtained for the PAVF-C and PAVF-P methods which are not shown here. From the figure, we see that the small perturbation in the initial data does not yield any significant effect on the wave propagation, which confirms the stability of the methods.

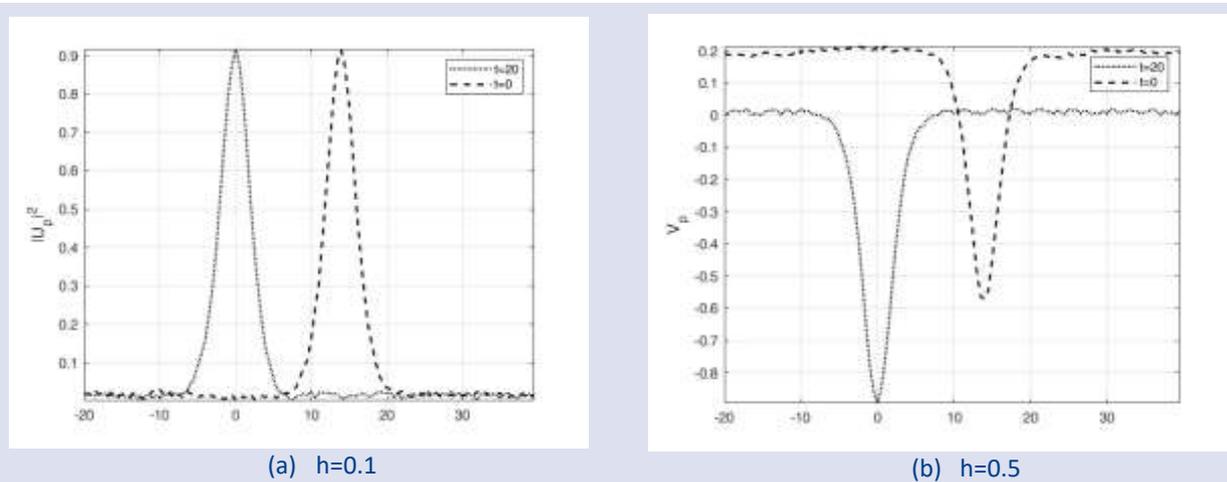


Figure 2. Solitary wave propagation with noise. $\tau=0.001, h=0.5$.

Single solitary wave: In this subsection, we examine the long-time solitary wave simulation. We solve the SB equations (1) in the spatial interval $[-64, 64]$ and temporal interval $[0, 40]$. We choose the spatial interval large enough so that solitary wave propagation and conservation properties are not affected by the boundary conditions. We can see that solitons move forward without any changes in shape as shown in Fig.3. The corresponding errors of total energy H^n and mass M^n are

represented by Fig.4. From the figure, we see that the errors are small during long-time integration which confirms theoretical results. Fig 3(a) represents that all four methods preserve the discrete energy up to round-off errors. From Fig. 3(b) we can see that, the AVF method cannot preserve the mass, but the relative error is bounded. On the other hand, we can see the excellent mass preservation of the other three methods during time integration.

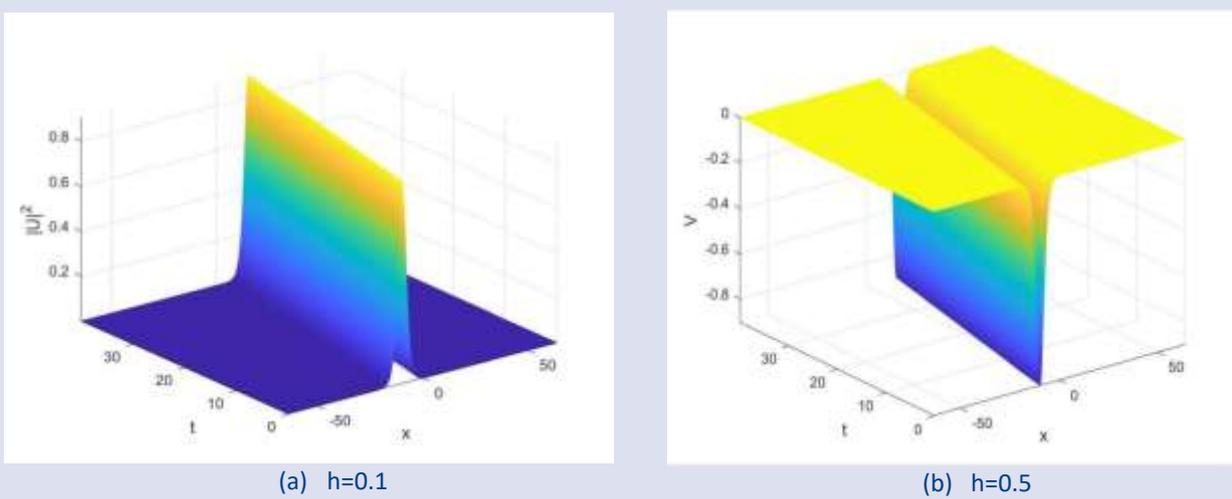


Figure 3. Evolution of PAVF solution for $\tau=0.001, h=0.2$

Numerical results show that the proposed schemes have excellent conservation properties and stability even for long time integration.

Two-dimensional Schrodinger-Boussinesq equation

The advantage of the PAVF method becomes more evident in two-dimensional problems [16]. In this subsection, we consider the two-dimensional SB equations (1) with $d = 2$, and $f(v) = \sin(v)$ [19,20]. We choose the parameters

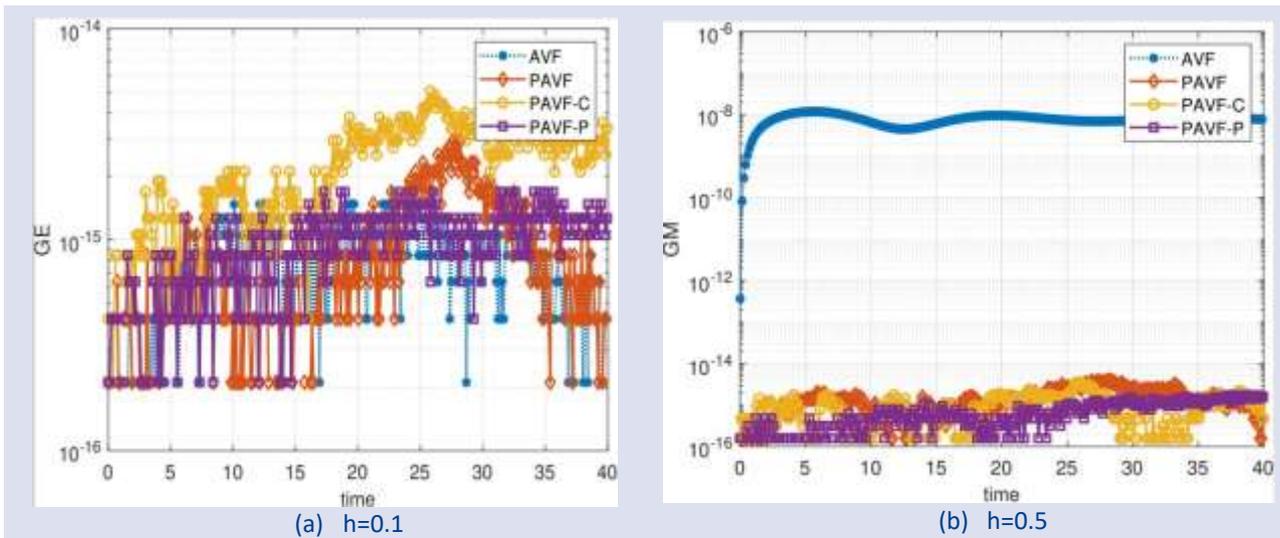


Figure 4. Relative errors in total energy H^n and the mass M^n for $\tau=0.01, h=0.2$.

Table 4. Computation time for $h_x = h_y = 0.25, -10 \leq x, y \leq 10$ at $t = 1$.

| τ | PAVF | PAVF-C | PAVF-P | AVF [18] |
|--------|--------|---------|---------|----------|
| 0.1 | 3.609 | 6.575 | 13.794 | 8.552 |
| 0.05 | 5.160 | 9.670 | 18.047 | 15.798 |
| 0.01 | 28.731 | 57.022 | 97.931 | 66.328 |
| 0.005 | 58.076 | 116.404 | 200.128 | 131.089 |

$\gamma = \alpha = -1$ and $\xi = \omega = -1/10$. The analytical solution of two-dimensional SB equation is not available. The initial data are taken as [19]

$$\begin{aligned}
 u_0(x, y) &= \frac{2}{e^{x^2+2y^2} + e^{-(x^2+2y^2)}} e^{5i/\cosh(\sqrt{4x^2+y^2})} \\
 v_0(x, y) &= e^{-(x^2+2y^2)} \\
 \phi_0(x, y) &= \frac{1}{2} e^{-(x^2+2y^2)}
 \end{aligned}
 \tag{38}$$

Table 4 indicates the computational time obtained by four energy-preserving methods in the spatial domain $\Omega = (-10 \times 10) \times (-10 \times 10)$ with $h_x = h_y = 0.1$ for the different temporal step size. Here, h_x and h_y are step size in x -direction and y -direction, respectively. From the table, we see the PAVF method requires smaller computational time than the AVF method, as expected. Moreover, as in the one-dimensional case, the computational time of PAVF-C method is much less than twice that of PAVF. In addition, it can be seen that the PAVF-P method is a method with the most CPU time.

Fig. 5 shows the surfaces of $|u|, v$ and ϕ at the initial time $t = 0$ and the final time $t = 1$. The spatial interval is set to $\Omega = [-16 \times 16] \times [-16 \times 16]$. The step sizes are taken as $\tau = 0.01$ and $h_x = h_y = 0.25$.

From the left plot of the figure, one can see the initial hump collapses as time evolves many spikes appear at the end. From the middle plot, one can see that the initial hump collapses as time evolves, and a hole appears inside the hump. In addition, there are spikes from under the horizontal plane. The right graph displays the evolution of the plane ϕ . From the right figure, one can see that again the initial hump collapses, but no hole appears. Moreover, spikes from under the horizontal plane. The energy and mass conservations are depicted in Fig.6. From the figure, one can find that PAVF-C preserves the total energy better than the other methods. Moreover, all PAVF methods preserve the mass better than the AVF method.

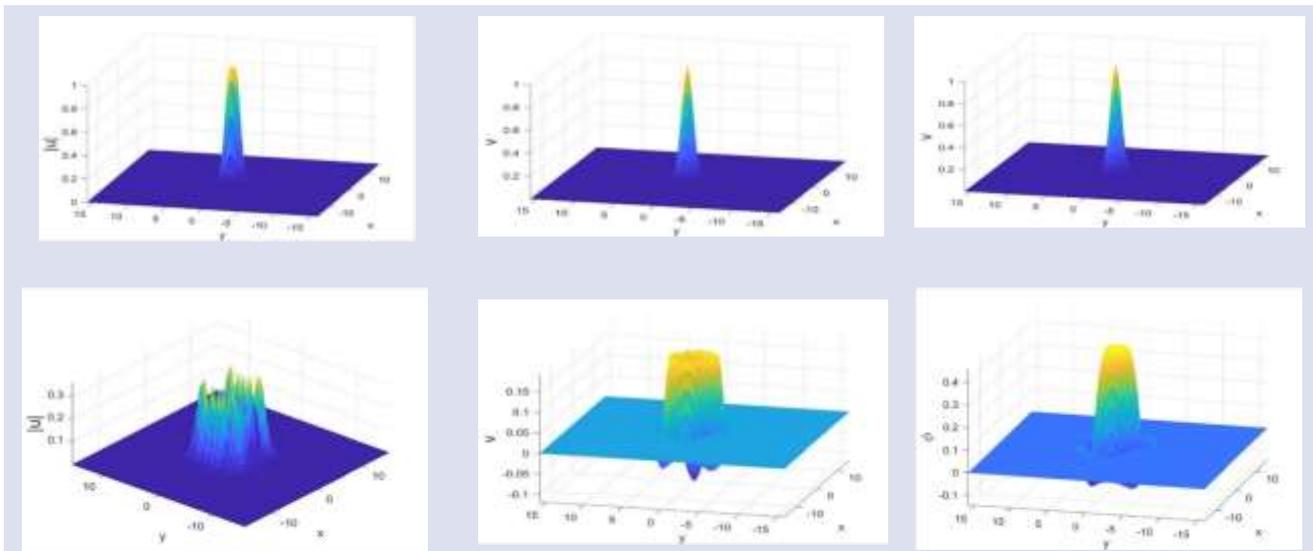


Figure 5. Evolution of PAVF solution for $\tau=0.01, h_x=h_y=0.25$

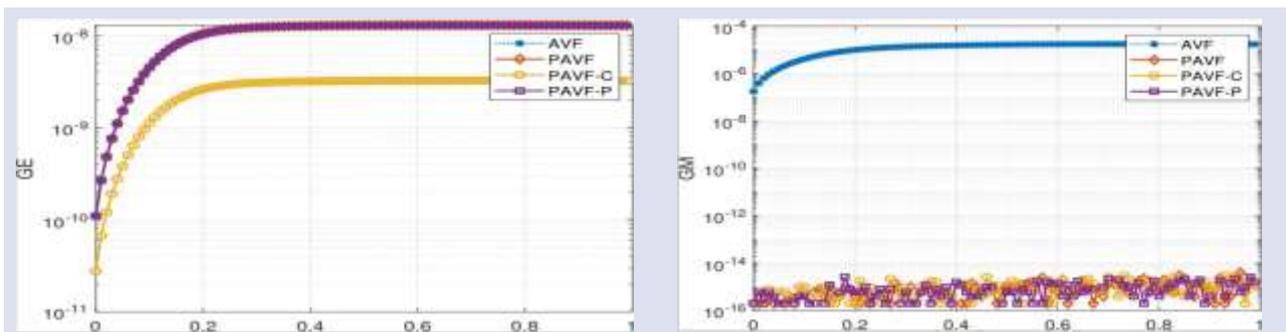


Figure 6. Error in the total energy H^n and the mass L^n for $\tau=0.01, h=0.25$.

Conclusion

In this study, we applied the energy-preserving PAVF method to the one-dimensional and two-dimensional CNSB equations. The method leads to semi-implicit algebraic equations for the CNSB equations. In addition, we find that the new energy-preserving scheme preserves the mass of the equations. In conjunction with the adjoint method of the PAVF method, we further introduce the conservative PAVF composition (PAVF-C) method and (PAVF-P) method. Some numerical results are presented

to demonstrate the accuracy and efficiency of the method for the numerical solution of the CNSB equations. Numerical results confirm the theoretical results. Errors in the methods and computational times are compared with the AVF method. Numerical results verify that the applied schemes simulate both one- and two-dimensional CNSB system well.

Conflict of interest

There are no conflicts of interest in this work.

Acknowledgments

The authors' gratitude to the reviewers for their valuable contributions.

References

- [1] Zhang H., Song S.H., Chen X.D., Zhou W.E., Average vector field methods for the coupled Schrödinger—KdV equations, *Chinese Physics B*, 23(7) (2014) 070208.
- [2] Aydin A., Multisymplectic integration of N-coupled nonlinear Schrödinger equation with destabilized periodic wave solutions, *Chaos, Solitons & Fractals*, 41(2) (2009) 735-751.
- [3] Wang L., Wang Y., Multisymplectic structure-preserving scheme for the coupled Gross–Pitaevskii equations, *International Journal of Computer Mathematics*, 98(4) (2021) 783-806.
- [4] Yajima N., Satsuma J., Soliton solutions in a diatomic lattice system, *Progress of Theoretical Physics*, 62(2) (1979) 370-378.
- [5] Rao N.N., Coupled scalar field equations for nonlinear wave modulations in dispersive media, *Pramana*, 46 (1996) 161-202.
- [6] Huang L.Y., Jiao Y.D., Liang D.M., Multi-symplectic scheme for the coupled Schrödinger–Boussinesq equations, *Chinese Physics B*, 22(7) (2013) 070201.
- [7] Bai D., Zhang L., The quadratic B-spline finite-element method for the coupled Schrödinger–Boussinesq equations, *International Journal of Computer Mathematics*, 88(8) (2011) 1714-1729.
- [8] Zhang L., Bai D., Wang S., Numerical analysis for a conservative difference scheme to solve the Schrödinger–Boussinesq equation, *Journal of computational and applied mathematics*, 235(17) (2011) 4899-4915.
- [9] Fei Z., Pérez-García V.M., Vázquez L., Numerical simulation of nonlinear Schrödinger systems: a new conservative scheme, *Applied Mathematics and Computation*, 71(2-3) (1995) 165-177.
- [10] Bai D., Wang J., The time-splitting Fourier spectral method for the coupled Schrödinger–Boussinesq equations. *Communications in Nonlinear Science and Numerical Simulation*, 17(3) (2012) 1201-1210.
- [11] Liao F., Zhang L., Wang S., Time-splitting combined with exponential wave integrator Fourier pseudospectral method for Schrödinger–Boussinesq system, *Communications in Nonlinear Science and Numerical Simulation*, 55 (2018) 93-104.
- [12] Hairer E., Wanner G., Lubich C., Hairer E., Wanner G., Lubich C., Symmetric Integration and Reversibility, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, (2006) 143-178.
- [13] Quispel G.R.W., McLaren D.I., A new class of energy-preserving numerical integration methods, *Journal of Physics A: Mathematical and Theoretical*, 41(4) (2008) 045206.
- [14] Celledoni E., Grimm V., McLachlan R.I., McLaren D.I., O’Neale D., Owren B. and Quispel G.R.W., Preserving energy resp. dissipation in numerical PDEs using the “Average Vector Field” method, *Journal of Computational Physics*, 231(20) (2012) 6770-6789.
- [15] Wang L., Cai W., Wang Y., An energy-preserving scheme for the coupled Gross-Pitaevskii equations, *Adv. Appl. Math. Mech.*, 13(1) (2021) 203-231.
- [16] Cai W., Li H., Wang Y., Partitioned averaged vector field methods. *Journal of Computational Physics*, 370 (2018) 25-42.
- [17] Wang J., Conservative Fourier spectral scheme for the coupled Schrödinger–Boussinesq equations, *Advances in Difference Equations*, (2018) 1-19.
- [18] Aydin A, Mohammed T.Y., An Energy Preserving Scheme for CSB System, 2nd Int. Grad. Stud.Cong. (IGSCONG-22), Proceeding Book, ISBN: 978-605-73639-2-3, 91-101.
- [19] Cai J., Chen J., Yang B., Efficient energy-preserving wavelet collocation schemes for the coupled nonlinear Schrödinger-Boussinesq system, *Applied Mathematics and Computation*, 357 (2019) 1-11.
- [20] Liao F., Zhang L., Conservative compact finite difference scheme for the coupled Schrödinger–Boussinesq equation, *Numerical Methods for Partial Differential Equations*, 32(6) (2016) 1667-1688.
- [21] Aydin A., Karasözen B., Lobatto IIIA–IIIB discretization of the strongly coupled nonlinear Schrödinger equation, *Journal of computational and applied mathematics*, 235(16) (2011) 4770-4779.
- [22] Aydin A., Karasözen B., Symplectic and multi-symplectic methods for coupled nonlinear Schrödinger equations with periodic solutions, *Computer Physics Communications*, 177(7) (2007) 566-583.
- [23] Leimkuhler B., Reich S., Simulating Hamiltonian dynamics (No. 14). Cambridge University Press, (2004)
- [24] Aydin, A., Sabawe, BAK., New conservative schemes for Zakharov equation, *Turk. J. Math. Comput. Sci.* 15(2) (2023) 227-293.
- [25] Ismail M. S., Ashi H. A., A compact finite difference scheme for solving the coupled nonlinear Schrödinger-Boussinesq equations, *Applied Mathematics*, 7 (2016) 605-615.
- [26] Hu X., Wang S., Zhang L., Maximum error estimates for a compact difference scheme of the coupled nonlinear Schrödinger–Boussinesq equations, *Numerical Methods for Partial Differential Equations*, 35(6) (2019) 1971-1999.
- [27] Deng D., Wu Q., Analysis of the linearly energy-and mass-preserving finite difference methods for the coupled Schrödinger-Boussinesq equations, *Applied Numerical Mathematics*, 170 (2021) 14-38.
- [28] Tingchun W., Boling G., Qiubin X., Fourth-order compact and energy conservative difference schemes for the nonlinear Schrödinger equation in two dimensions, *Journal of Computational Physics*, 243 (2013) 382–399.
- [29] Hong-Lin L., Zhi-Zhong S., Han-Sheng S., Error Estimate of Fourth-Order Compact Scheme For Linear Schrödinger Equations, *SIAM J. Numer. Anal.*, 47(6) (2010) 4381–4401.
- [30] Gholamreza K., Hadi MF., The Matrix Transformation Technique for the Time-Space Fractional Linear Schrödinger Equation, *Iranian Journal of Mathematical Chemistry*, 15(3) (2024) 137-154.
- [31] Gholamreza K., Hadi MF., Unconditionally stable finite element method for the variable-order fractional Schrödinger equation with Mittag-Leffler kernel, *Journal of Mathematical Modeling*, 12(3) (2024) 533–550.



On the Inverse Problems for Conformable Fractional Integro-Dirac Differential System with Parameter Dependent Boundary Conditions

Hediye Dilara Tel^{1,a}, Baki Keskin^{1,b,*}

¹ Department of Mathematics, Faculty of Science, Sivas Cumhuriyet University, Sivas, Türkiye.

*Corresponding author

Research Article

History

Received: 22/01/2024

Accepted: 23/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This study considers a conformable fractional Dirac-type integral differential system, focusing on its mathematical properties and practical implications. Asymptotic formulas have been derived for the solutions, eigenvalues, and nodes of the problem, providing a deeper understanding of the behavior of the system under varying conditions. These asymptotic results form the basis for analyzing the spectral characteristics and node distribution of the system. In addition, an algorithm is developed that effectively solves the inverse nodal problem and reconstructs the system coefficients from the nodal data.

Keywords: Conformable fractional Dirac system, integro-differential operators, inverse nodal problems.

dilara_5820@hotmail.com.tr

<https://orcid.org/0000-0003-1139-6146>

bkeskin@cumhuriyet.edu.tr

<https://orcid.org/0000-0003-1689-8954>

Introduction

The Dirac operator is a mathematical operator that appears in quantum mechanics and quantum field theory. It was introduced by the physicist Paul Dirac in 1928 as a way to describe the behavior of electrons in relativistic quantum mechanics [1]. The first important and comprehensive results regarding these operators were discussed in Levitan's work [2]. Inverse problems for Dirac operators have been addressed and studied in detail by many researchers (see for example [3-5]).

The fractional calculus has gained considerable attention in various scientific disciplines due to its wide range of applications and its effectiveness in dealing with complex systems. Fractional calculus extends the traditional integer-order calculus to include derivatives and integrals of non-integer orders [6-9]. In 2014, Khalil et al. presented a new but easy definition of the fractional derivative, called the compatible fractional derivative [10]. The new definition seems to be a natural extension of the traditional differentiation and seems to agree with known fractional derivatives on polynomials (up to a constant multiple).

This derivative was defined as follows:

Let $f: [0, \infty) \rightarrow \mathbb{R}$ be a given function. The conformable fractional derivative of f of order α is:

$$D_x^\alpha f(x) = \lim_{\epsilon \rightarrow 0} \frac{f(x+\epsilon x^{1-\alpha}) - f(x)}{\epsilon}, D_x^\alpha f(0) = \lim_{x \rightarrow 0^+} D_x^\alpha f(x),$$

for all $x > 0$, $\alpha \in (0, 1]$. If this limit exist and finite at x_0 , we say f is α -differentiable at x_0 . If f is differentiable, then $D_x^\alpha f(x) = x^{1-\alpha} f'(x)$.

In the past few years, fractional calculus has been investigated by several author ([11], [12] and references therein). In recent years, there has been a growing interest

among scholars in exploring fractional generalizations of well-known mathematical problems, including those related to Sturm-Liouville, diffusion and Dirac operators [13-21].

The nodal set consists of points where the eigenfunction vanishes. In 1988, the concept of the inverse nodal problem for the Sturm-Liouville operator was first discussed by McLaughlin [22], and later Hald and McLaughlin showed that it was sufficient to know only the nodal points to determine the potential function with more general boundary conditions [23]. Yang proposed a solution in 1997 to reconstruct the potential and the boundary condition of the Sturm-Liouville operator from its nodes. [24]. Inverse nodal problems continue to be studied by many researchers [25-34].

The inverse nodal problem for Dirac operators involves determining the coefficients of the Dirac operator and other parameters of the problem from the knowledge of the nodal set of the corresponding eigenfunctions. For certain types of Dirac operators with various boundary conditions, it has been demonstrated that a dense subset of the nodal points of the eigenfunctions alone is sufficient to uniquely determine the coefficients of the Dirac operator [35-37].

Eigenvalue problems with eigenvalue-dependent boundary conditions is an important application area in applied sciences. Fulton's [38-39], studies and the references in this study can be cited as examples of studies conducted on this subject until 1980. The most recent examples of its applications in physics can be found in [40]. We refer to [41-42] and references therein regarding studies in this field.

Nowadays, studies on the integro-differential operator have gained significant popularity and interest by many authors and have gained an important place in the literature [43-46]. The inverse nodal problem for Dirac type integro-differential operators was first considered in [47]. It is shown in this study that the coefficients of the operator can be determined by using nodal points. In [48],

the authors have addressed a similar problem where the boundary conditions depend linearly on the spectral parameter.

The conformable fractional derivative was first discussed in [10-11]. Some other definitions and basic properties can be found in these works.

Main Results

Consider the following BVP $L(\theta, \beta, p(x), q(x))$:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} D_x^\alpha Y(x) + \begin{pmatrix} p(x) & 0 \\ 0 & r(x) \end{pmatrix} Y(x) + \int_0^x M(x, t)Y(t)d_\alpha t = \lambda Y, x \in (0, \pi) \tag{1}$$

with

$$(\lambda \cos \theta + a)y_1(0) + (\lambda \sin \theta + b)y_2(0) = 0 \tag{2}$$

$$(\lambda \cos \beta + c)y_1(\pi) + (\lambda \sin \beta + d)y_2(\pi) = 0 \tag{3}$$

where, $M(x, t) = \begin{pmatrix} M_{11}(x, t) & M_{12}(x, t) \\ M_{21}(x, t) & M_{22}(x, t) \end{pmatrix}$, $Y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}$ and $p(x), q(x)$, and $M_{ij}(x, t)$ ($i, j = 1, 2$) are real-valued conformable fractional differentiable functions and $x^{\alpha-1}p(x)$ and $x^{\alpha-1}r(x)$ are continuous on $(0, \pi)$, $0 \leq \theta, \beta < \pi$ are real numbers, λ is the spectral parameter.

Let $\varphi(x, \lambda) = (\varphi_1(x, \lambda), \varphi_2(x, \lambda))^T$ be the solution of (1) satisfying $\varphi(0, \lambda) = (\lambda \sin \theta + b, -\lambda \cos \theta - a)^T$. $\varphi(x, \lambda)$ satisfies the following asymptotic relations for $|\lambda| \rightarrow \infty$,

$$\begin{aligned} \varphi_1(x, \lambda) &= \lambda \sin \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + a \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ b \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{1}{2} \mu(x) \sin \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ \frac{a}{2\lambda} \mu(x) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{b}{2\lambda} \mu(x) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ \frac{1}{2} \mu(0) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) - \theta \right) + \frac{a}{2\lambda} \mu(0) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &- \frac{b}{2\lambda} \mu(0) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - \frac{1}{2} \cos \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t \\ &- \frac{a}{2\lambda} \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t + \frac{b}{2\lambda} \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t \\ &- \frac{1}{2} K(x) \sin \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - \frac{a}{2\lambda} K(x) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &- \frac{b}{2\lambda} K(x) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{1}{2} L(x) \cos \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ \frac{a}{2\lambda} L(x) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - \frac{b}{2\lambda} L(x) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + o \left(\frac{e^{|\tau| \frac{x^\alpha}{\alpha}}}{\lambda} \right) \\ \varphi_2(x, \lambda) &= -\lambda \cos \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - a \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ b \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{1}{2} \mu(x) \cos \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &+ \frac{a}{2\lambda} \mu(x) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - \frac{b}{2\lambda} \mu(x) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &- \frac{1}{2} \mu(0) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) - \theta \right) - \frac{a}{2\lambda} \mu(0) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &- \frac{b}{2\lambda} \mu(0) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) - \frac{1}{2} \sin \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t \\ &- \frac{a}{2\lambda} \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t - \frac{b}{2\lambda} \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \int_0^x \mu^2(t) d_\alpha t \\ &+ \frac{1}{2} K(x) \cos \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{a}{2\lambda} K(x) \cos \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \\ &- \frac{b}{2\lambda} K(x) \sin \left(\lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) + \frac{1}{2} L(x) \sin \left(\theta + \lambda \frac{x^\alpha}{\alpha} - \rho(x) \right) \end{aligned} \tag{4}$$

$$+ \frac{a}{2\lambda} L(x) \sin\left(\lambda \frac{x^\alpha}{\alpha} - \rho(x)\right) + \frac{b}{2\lambda} L(x) \cos\left(\lambda \frac{x^\alpha}{\alpha} - \rho(x)\right) + o\left(\frac{e^{|\tau| \frac{x^\alpha}{\alpha}}}{\lambda}\right)$$

uniformly in $x \in [0, \pi]$, where, $\mu(x) = \frac{1}{2}(p(x) - r(x))$, $\rho(x) = \frac{1}{2} \int_0^x (p(t) + r(t)) d_\alpha t$,
 $K(x) = \int_0^x (M_{11}(t, t) - M_{22}(t, t)) d_\alpha t$, $L(x) = \int_0^x (M_{12}(t, t) - M_{21}(t, t)) d_\alpha t$
 and $\tau = Im\lambda$.

$\Delta(\lambda)$ is called the characteristic function of L and defined by as follows

$$\Delta(\lambda) = \varphi_1(\pi, \lambda)(\lambda \cos \beta + c) + \varphi_2(\pi, \lambda)(\lambda \sin \beta + d), \tag{6}$$

The zeros $\{\lambda_n\}_{n \in \mathbb{Z}}$ of $\Delta(\lambda)$ coincide with the eigenvalues of the problem L . Using (4) and (5), we get

$$\begin{aligned} \Delta(\lambda) &= \lambda^2 \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta + \theta\right) + \lambda a \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta\right) \\ &+ b \lambda \cos\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta\right) + \frac{\mu(x)}{2} \lambda \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) + \beta + \theta\right) \\ &+ \frac{\mu(0)}{2} \lambda \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta - \theta\right) - \frac{1}{2} \lambda \cos\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta + \theta\right) \int_0^\pi \mu^2(t) d_\alpha t \\ &- \frac{1}{2} \lambda K(\pi) \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta + \theta\right) + \frac{1}{2} \lambda L(\pi) \cos\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) - \beta + \theta\right) \\ &+ c \lambda \sin\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) + \theta\right) - d \lambda \cos\left(\lambda \frac{\pi^\alpha}{\alpha} - \rho(\pi) + \theta\right) + o\left(\frac{e^{|\tau| \frac{\pi^\alpha}{\alpha}}}{\lambda^3}\right), \end{aligned} \tag{7}$$

for sufficiently large $|\lambda|$.

By using $\Delta(\lambda_n) = 0$, we get

$$\lambda_n = \frac{(n-1)\pi\alpha}{\pi^\alpha} + \frac{(\rho(\pi)+\beta-\theta)\alpha}{\pi^\alpha} + \frac{D}{(n-1)\pi^\alpha} + o\left(\frac{1}{n}\right) n \geq 2 \tag{8}$$

and

$$\lambda_n = \frac{(n+1)\pi\alpha}{\pi^\alpha} + \frac{(\rho(\pi)+\beta-\theta)\alpha}{\pi^\alpha} + \frac{D}{(n+1)\pi^\alpha} + o\left(\frac{1}{n}\right) n \leq -2$$

for sufficiently large n ,

$$\text{where, } D = a \sin \theta - b \cos \theta - \frac{\mu(x)}{2} \sin 2\beta + \frac{\mu(0)}{2} \sin 2\theta + \frac{1}{2} \int_0^\pi \mu^2(t) d_\alpha t - \frac{L(\pi)}{2} - c \sin \beta + d \cos \beta$$

Lemma 1 For sufficiently large n , the first component $\varphi_1(x, \lambda_n)$ of the eigenfunction $\varphi(x, \lambda_n)$ has exactly $n - 2$ nodes $\{x_n^j: j = 0, 1, \dots, n - 3\}$ in the interval $(0, \pi)$: $0 < x_n^0 < x_n^1 < \dots < x_n^{n-3} < \pi$. The numbers $\{x_n^j\}$ satisfy the following asymptotic formula:

$$\begin{aligned} (x_n^j)^\alpha &= \frac{j\pi^\alpha}{n} - \frac{j\pi^\alpha}{n} \frac{\rho(\pi)+\beta-\theta}{n\pi} + \rho(x_n^j) \frac{\pi^{\alpha-1}}{n} - \theta \frac{\pi^{\alpha-1}}{n} - \rho(x_n^j) \frac{\rho(\pi)+\beta-\theta}{n^2} \pi^{\alpha-2} \\ &+ \theta \frac{\rho(\pi)+\beta-\theta}{n^2} \pi^{\alpha-2} + T \frac{\pi^{2\alpha-2}}{2n^2\alpha} + \frac{\pi^{2\alpha-2}}{2n^2\alpha} \int_0^{x_n^j} \mu^2(t) d_\alpha t - \frac{\pi^{2\alpha-2}}{2n^2\alpha} L(x_n^j) \\ &\div \\ &- T \frac{\rho(\pi)+\beta-\theta}{n^3\alpha} \pi^{2\alpha-3} - \frac{\rho(\pi)+\beta-\theta}{n^3\alpha} \int_0^{x_n^j} \mu^2(t) d_\alpha t - \frac{\rho(\pi)+\beta-\theta}{n^3\alpha} L(x_n^j) + o\left(\frac{1}{n^3}\right). \end{aligned} \tag{9}$$

$$\text{Where, } T = 2a \sin \theta - 2b \cos \theta + \mu(0) \sin 2\theta$$

Proof. From (4), the following equation is valid

$$\begin{aligned} \varphi_1(x, \lambda_n) &= \lambda_n \sin\left(\theta + \lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) + a \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \\ &+ b \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) + \frac{1}{2} \mu(x) \sin\left(\theta + \lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \\ &+ \frac{a}{2\lambda_n} \mu(x) \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) + \frac{b}{2\lambda_n} \mu(x) \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \\ &+ \frac{1}{2} \mu(0) \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x) - \theta\right) + \frac{a}{2\lambda_n} \mu(0) \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \end{aligned}$$

$$\begin{aligned}
 & -\frac{b}{2\lambda_n} \mu(0) \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) - \frac{1}{2} \cos\left(\theta + \lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \int_0^x \mu^2(t) d_\alpha t \\
 & -\frac{a}{2\lambda_n} \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \int_0^x \mu^2(t) d_\alpha t + \frac{b}{2\lambda_n} \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \int_0^x \mu^2(t) d_\alpha t \\
 & -\frac{1}{2} K(x) \sin\left(\theta + \lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) - \frac{a}{2\lambda_n} K(x) \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \\
 & -\frac{b}{2\lambda_n} K(x) \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) + \frac{1}{2} L(x) \cos\left(\theta + \lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) \\
 & + \frac{a}{2\lambda_n} L(x) \cos\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) - \frac{b}{2\lambda_n} L(x) \sin\left(\lambda_n \frac{x^\alpha}{\alpha} - \rho(x)\right) + o\left(\frac{1}{\lambda_n}\right)
 \end{aligned}$$

for sufficiently large n . If we put $\varphi_1((x_n^j)^\alpha, \lambda_n) = 0$, we get

$$\begin{aligned}
 & \lambda_n \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) + a \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta \\
 & - a \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta + b \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta \\
 & + b \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta + \frac{1}{2} \mu(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \\
 & + \frac{a}{2\lambda_n} \mu(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta - \frac{a}{2\lambda_n} \mu(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & + \frac{b}{2\lambda_n} \mu(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta + \frac{b}{2\lambda_n} \mu(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & + \frac{1}{2} \mu(0) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos 2\theta - \frac{1}{2} \mu(0) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin 2\theta \\
 & + \frac{a}{2\lambda_n} \mu(0) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta - \frac{a}{2\lambda_n} \mu(0) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & - \frac{b}{2\lambda_n} \mu(0) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta - \frac{b}{2\lambda_n} \mu(0) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & - \left(\frac{1}{2} \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) + \frac{a}{2\lambda_n} \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta\right) \int_0^{x_n^j x} \mu^2(t) d_\alpha t \\
 & - \left(\frac{a}{2\lambda_n} \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta - \frac{b}{2\lambda_n} \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta\right) \int_0^{x_n^j x} \mu^2(t) d_\alpha t \\
 & - \frac{b}{2\lambda_n} \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \int_0^{x_n^j x} \mu^2(t) d_\alpha t - \frac{1}{2} K(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \\
 & - \frac{a}{2\lambda_n} K(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta + \frac{a}{2\lambda_n} K(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & - \frac{b}{2\lambda_n} K(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta - \frac{b}{2\lambda_n} K(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta \\
 & + \frac{1}{2} L(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) + \frac{a}{2\lambda_n} L(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta \\
 & + \frac{a}{2\lambda_n} L(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta - \frac{b}{2\lambda_n} L(x_n^j) \sin\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \cos\theta \\
 & + \frac{b}{2\lambda_n} L(x_n^j) \cos\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x_n^j) + \theta\right) \sin\theta + o\left(\frac{1}{\lambda_n}\right) = 0
 \end{aligned}$$

$$\lambda_n \tan\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta\right) + a \tan\left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta\right) \cos\theta - a \sin\theta$$

$$\begin{aligned}
 &+ b \cos \theta + b \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta + \frac{1}{2} \mu(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \\
 &+ \frac{a}{2\lambda_n} \mu(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos \theta - \frac{a}{2\lambda_n} \mu(x) \sin \theta + \frac{b}{2\lambda_n} \mu(x) \cos \theta \\
 &+ \frac{b}{2\lambda_n} \mu(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta + \frac{1}{2} \mu(0) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos 2\theta - \frac{1}{2} \mu(0) \sin 2\theta \\
 &+ \frac{a}{2\lambda_n} \mu(0) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos \theta - \frac{a}{2\lambda_n} \mu(0) \sin \theta \\
 &- \frac{b}{2\lambda_n} \mu(0) \cos \theta - \frac{b}{2\lambda_n} \mu(0) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta - \frac{1}{2} \int_0^{x_n^j} \mu^2(t) d_\alpha t - \frac{a}{2\lambda_n} \cos \theta \int_0^{x_n^j} \mu^2(t) d_\alpha t \\
 &- \frac{a}{2\lambda_n} \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta \int_0^{x_n^j} \mu^2(t) d_\alpha t + \frac{b}{2\lambda_n} \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos \theta \int_0^{x_n^j} \mu^2(t) d_\alpha t \\
 &- \frac{b}{2\lambda_n} \sin \theta \int_0^{x_n^j} \mu^2(t) d_\alpha t - \frac{1}{2} K(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) - \frac{a}{2\lambda_n} K(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos \theta \\
 &+ \frac{a}{2\lambda_n} K(x) \sin \theta - \frac{b}{2\lambda_n} K(x) \cos \theta + \frac{1}{2} L(x) + \frac{a}{2\lambda_n} L(x) \cos \theta - \frac{b}{2\lambda_n} K(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta \\
 &+ \frac{b}{2\lambda_n} L(x_n^j) \sin \theta + o\left(\frac{1}{\lambda}\right) + \frac{a}{2\lambda_n} L(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \sin \theta \\
 &- \frac{b}{2\lambda_n} L(x) \tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \cos \theta = 0
 \end{aligned}$$

$$\begin{aligned}
 &\tan \left(\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho(x) + \theta \right) \left\{ 1 + \frac{a}{\lambda_n} \cos \theta + \frac{b}{\lambda_n} \sin \theta \right. \\
 &+ \left. \frac{1}{2\lambda_n} \mu(x_n^j) + \frac{1}{2\lambda_n} \mu(0) - \frac{1}{2\lambda_n} K x_n^j + o\left(\frac{1}{\lambda_n}\right) \right\} = \\
 &\frac{a}{\lambda_n} \sin \theta - \frac{b}{\lambda_n} \cos \theta + \frac{1}{2\lambda_n} \mu(0) \sin 2\theta + \frac{1}{2\lambda_n} \int_0^{x_n^j} \mu^2(t) d_\alpha t - \frac{1}{2\lambda_n} L(x_n^j) + o\left(\frac{1}{\lambda_n}\right)
 \end{aligned}$$

Taylor's expansion formula gives,

$$\begin{aligned}
 &\lambda_n \frac{(x_n^j)^\alpha}{\alpha} - \rho((x_n^j)) + \theta \\
 &= j\pi + \frac{1}{2\lambda_n} \left(2a \sin \theta - 2b \cos \theta + \mu(0) \sin 2\theta + \int_0^{x_n^j} \mu^2(t) d_\alpha t - L(x_n^j) \right) \\
 &+ o\left(\frac{1}{\lambda_n}\right)
 \end{aligned}$$

or

$$\begin{aligned}
 (x_n^j)^\alpha &= \alpha \lambda_n^{-1} \left(\rho((x_n^j)) - \theta + j\pi \right. \\
 &\left. + \frac{1}{2\lambda_n} \left(2a \sin \theta - 2b \cos \theta + \mu(0) \sin 2\theta + \int_0^{x_n^j} \mu^2(t) d_\alpha t - L(x_n^j) \right) \right) + o\left(\frac{1}{\lambda_n}\right)
 \end{aligned}$$

We arrive (9) by using the asymptotic formula

$$\lambda_n^{-1} = \frac{\pi^\alpha}{n\pi\alpha} - \frac{(\rho(\pi) + \beta - \theta)\pi^\alpha}{n^2\pi^2\alpha} + o\left(\frac{1}{n^2}\right)$$

Let X be the set of nodal points. For each fixed $x \in (0, \pi)$ and $\alpha \in (0, 1]$, choose a sequence $(x_n^j) \subset X$ such that x_n^j converges to x . Then the following limits are exist and finite:

$$\lim_{|n| \rightarrow \infty} \left((x_n^j)^\alpha - \frac{j\pi^\alpha}{n} \right) n\pi = -x(\rho(\pi) - \theta + \beta) + \rho(x)\pi^\alpha - \theta\pi^\alpha = f(x)$$

where

$$f(x) = -x(\rho(\pi) - \theta + \beta) + \frac{1}{2}\pi^\alpha \int_0^x [p(t) + r(t)]d_\alpha t - \theta\pi^\alpha \tag{10}$$

and

$$\lim_{n \rightarrow \infty} \left((x_n^j)^\alpha - \frac{j\pi^\alpha}{n} + \frac{j\pi^\alpha}{n} \frac{\rho(\pi) + \beta - \theta}{n\pi} - \frac{\rho(x_n^j)\pi^{\alpha-1}}{n} + \frac{\theta\pi^{\alpha-1}}{n} \right) n^2 = g(x),$$

where

$$g(x) = -\rho(x)(\beta - \theta)\pi^{\alpha-2} + \theta(\beta - \theta)\pi^{\alpha-2} + \frac{\pi^{2\alpha-2}}{2\alpha} + \int_0^x \mu^2(t)d_\alpha t + \frac{\pi^{2\alpha-2}}{2\alpha} L(x) + T \frac{\pi^{2\alpha-2}}{2\alpha} \tag{11}$$

Therefore, proof of the following theorem is clear. Let $\mu(\pi) = 0$, and X be the dense subset of the nodal points.

Theorem 1 Given the set X uniquely determines the coefficients θ and β of the problem L and if $L(x)$ is known, the potential $\Omega(x)$ a.e. on $(0, \pi)$ can be also determined by X . Moreover, $p(x)$ and $r(x)$, θ and β can be reconstructed as follows

Step-1: For each fixed $x \in (0, \pi)$ and $\alpha \in (0, 1]$, choose $(x_n^{j(n)}) \subset X$ such that $(x_n^{j(n)}) \rightarrow x$ as $n \rightarrow \infty$;

Step-2: Find $f(x)$ from (10) and calculate

$$\begin{aligned} \theta &= -f(0)\pi^{-\alpha} \\ \beta &= \frac{f(0) - f(\pi) - f(0)\pi^{1-\alpha}}{\pi} \\ D_x^\alpha \rho(x) &= (D_x^\alpha f(x) - \theta + \beta)\pi^{-\alpha} \end{aligned}$$

Step-3: From (11), find $g(x)$ and calculate

$$\mu^2(x) = (D_x^\alpha g(x) + (D_x^\alpha f(x) - \theta + \beta)(\beta - \theta)\pi^{-2}) \frac{2\alpha}{\pi^{2\alpha-2}} + D_x^\alpha L(x) \tag{12}$$

Step-4: From (10) and (11) calculate

$$\begin{aligned} p(x) &= \frac{D_x^\alpha f(x)}{\pi^\alpha} + \frac{f(0) - f(\pi) - f(0)\pi^{1-\alpha}}{\pi^{1+\alpha}} + f(0) + 2\sqrt{\frac{2\alpha}{\pi^{2\alpha-2}} (D_x^\alpha g(x) + D_x^\alpha \rho(x)(\beta - \theta)\pi^{\alpha-2}) + D_x^\alpha L(x)} \\ r(x) &= \frac{D_x^\alpha f(x)}{\pi^\alpha} + \frac{f(0) - f(\pi) - f(0)\pi^{1-\alpha}}{\pi^{1+\alpha}} + f(0) - 2\sqrt{\frac{2\alpha}{\pi^{2\alpha-2}} (D_x^\alpha g(x) + D_x^\alpha \rho(x)(\beta - \theta)\pi^{\alpha-2}) + D_x^\alpha L(x)} \end{aligned}$$

Acknowledgement

This study is a part of PhD thesis of Hediye Dilara TEL.

[5] Horvath M., On the inverse spectral theory of Schrödinger and Dirac operators, *Trans. Amer. Math. Soc.*, 353 (2001) 4155-4171..

Conflicts of interest

There are no conflicts of interest in this work.

[6] Miller K. S., An Introduction to Fractional Calculus and Fractional Differential Equations, J. Wiley and Sons, New York, NY, USA, 1993.

References

[1] Dirac PAM, The quantum theory of the electron, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117(778) (1928) 610-624.
 [2] Levitan BM., IS. Sargsyan, Sturm Liouville and Dirac operators. Kluwer Academic Publishers: Durdrecht/Boston/London; 1991.
 [3] Albeverio S., R. Hryniv, Mykytyuk Ya., Reconstruction of radial Dirac and Schrödinger operators from two spectra, *J. Math. Anal. Appl.* 339 (2008) 45-57.
 [4] Gasymov MG., Inverse problem of the scattering theory for Dirac system of order $2n$, *Tr. Mosk Mat. Obshch*, 19 (1968) 41-112.

[7] Kilbas A., Srivastava H., and Trujillo J., "Theory and applications of fractional differential equations," in *Math. Studies*, North-Holland, New York, NY, USA, 2006.
 [8] Oldham K., Spanier J., *The Fractional Calculus, Theory and Applications of Differentiation and Integration of Arbitrary Order*, Academic Press, Cambridge, MA, USA, 1974.
 [9] Podlubny I., *Fractional Differential Equations*, Academic Press, Cambridge, MA, USA, 1999.
 [10] Khalil R., M. Horani Al, Yousef A., Sababheh M., A new definition of fractional derivative, *Journal of Computational and Applied Mathematics*, 264 (2014) 65-70.
 [11] Abdeljawad T., On conformable fractional calculus, *Journal of Computational and Applied Mathematics*, 279 (2015) 57-6

- [12] Hammad M. Abu, Khalil R., Abel's formula and Wronskian for conformable fractional differential equations, *Int. J. Differ. Equ. Appl.*, 13(3) (2014) 177–183.
- [13] Al-Refai M., Abdeljawad T., Fundamental results of conformable Sturm–Liouville eigenvalue problems, *Complexity*, (2017) 3720471.
- [14] Çakmak Y., Inverse nodal problem for a conformable fractional diffusion operator, *Inverse Problems in Science and Engineering*, 29-9 (2021) 1308-1322.
- [15] Çakmak Y., Trace Formulae for a Conformable Fractional Diffusion Operator, *Filomat*, 36-14 (2022) 4665-4674.
- [16] Keskin B., Inverse problems for one dimensional conformable fractional Dirac type integro differential system, *Inverse Problems*, 36 (2020) 065001.
- [17] Allahverdiev, B.P., Tuna, H., One-dimensional conformable fractional Dirac system, *Bol. Soc. Mat. Mex.*, (2019).
- [18] Erdal B., Fundamental spectral theory of fractional singular Sturm-Liouville operator, *J Funct Space*, 1 (2013) 113-129.
- [19] Gulsen T., Yilmaz E., Goktas S., Conformable fractional Dirac system on time scales, *J. Inequal. Appl.*, (2017) 161,2017.
- [20] Adalar İ., Özkan A. S., Inverse problems for a conformable fractional Sturm-Liouville operator, *Journal of ill-posed problems*, 28(6) (2020) 775-782.
- [21] Zhaowen Z. , Huixi L. , Jinming C., Yanwei Z., Criteria of limit-point case for conformable fractional Sturm-Liouville operators, *Math Meth Appl Sci.*, 43 (2020) 2548–2557.
- [22] McLaughlin JR., Inverse spectral theory using nodal points as data – a uniqueness result, *J. Diff. Eq.*, 73 (1988) 354–362.
- [23] Hald OH., McLaughlin JR., Solutions of inverse nodal problems, *Inverse Problems*, 5 (1989) 307–347.
- [24] Yang XF., A solution of the nodal problem, *Inverse Problems*, 13 (1997) 203-213.
- [25] Browne PJ., Sleema B.D., Inverse nodal problem for Sturm–Liouville equation with eigenparameter depend boundary conditions, *Inverse Problems*, 12 (1996) 377–381.
- [26] Cheng Y., Law CK. and Tsay J., Remarks on a new inverse nodal problem, *J. Math. Anal. Appl.*, 248 (2000) 145–155.
- [27] Guo Y., Wei Y., Inverse problems: Dense nodal subset on an interior subinterval, *J. Diff. Eq.*, 255 (2002) 2017.
- [28] Law CK., Shen CL., Yang CF., The Inverse Nodal Problem on the Smoothness of the Potential Function, *Inverse Problems*, 15 (1999), no.1, 253-263 (Erratum, *Inverse Problems* 2001; 17: 361-363).
- [29] Ozkan AS., Keskin B., Inverse Nodal Problems for Sturm–Liouville Equation with Eigenparameter Dependent Boundary and Jump Conditions, *Inverse Problems in Science and Engineering*, 23(8) (2015) 1306-1312.
- [30] Wang YP., Yurko V., On the inverse nodal problems for discontinuous Sturm Liouville operators, *J. Differential Equations*, 260 (2016) 4086-4109.
- [31] Wang YP., Lien KY., Shieh CT., Inverse problems for the boundary value problem with the interior nodal subsets, *Applicable Analysis*, 96 (2017) 1229-1239.
- [32] Wei Z., Guo Y., Wei G., Incomplete inverse spectral and nodal problems for Dirac operator, *Adv. Difference Equ.*, 2015 (2015) 88.
- [33] Shieh CT., Yurko VA., Inverse nodal and inverse spectral problems for discontinuous boundary value problems, *J. Math. Anal. Appl.*, 347 (2008) 266-272.
- [34] Yang XF., A new inverse nodal problem, *J. Differential Equations*, 169 (2001) 633-653.
- [35] Guo Y., Wei Y., Inverse Nodal Problem for Dirac Equations with Boundary Conditions Polynomially Dependent on the Spectral Parameter, *Results in Math.*, 67 (2015) 95–110.
- [36] Yang CF., Huang ZY., Reconstruction of the Dirac operator from nodal data, *Integr. Equ. Oper. Theory*, 66 (2010) 539–551.
- [37] Yang CF., Pivovarchik VN., Inverse nodal problem for Dirac system with spectral parameter in boundary conditions, *Complex Anal. Oper. Theory*, 7 (2013) 1211–1230
- [38] Fulton C. T., Two-point boundary value problems with eigenvalue parameter contained in the boundary conditions, *Proc. Roy. Soc. Edinburgh Sect. A* , 77 (3–4) (1977) 293–308.
- [39] Fulton C. T., Singular eigenvalue problems with eigenvalue parameter contained in the boundary conditions, *Proc. Roy. Soc. Edinburgh Sect. A* , 87 (1–2) (1980) 1–34. 10.1017.
- [40] Guliyev N. J., Spectral identities for Schrodinger operators , *Canad. Math. Bull.*, (to appear)
- [41] Guliyev N. J., Essentially isospectral transformations and their applications, *Ann. Mat. Pura Appl.*, 199(4) (2020) 1621–1648.
- [42] Guliyev N. J., Schrödinger operators with distributional potentials and boundary conditions dependent on the eigenvalue parameter, *J. Math. Phys.*, 60(6) (2019) 063501.
- [43] Bondarenko NP., An inverse problem for the integro-differential Dirac system with partial information given on the convolution kernel, *J. Inverse Ill-Posed Probl.*, 27 (2) (2018) 151-157.
- [44] Buterin SA., On an Inverse Spectral Problem for a Convolution Integro-Differential Operator, *Results in Mathematics*, 50 (2007) 173-181.
- [45] Kuryshova YV., Shieh CT., An Inverse Nodal Problem for Integro-Differential Operators, *Journal of Inverse and Ill-posed Problems*, 18 (2010) 357–369.
- [46] Wu B., Yu J., Uniqueness of an Inverse Problem for an Integro-Differential Equation Related to the Basset Problem, *Boundary Value Problems*, 229 (2014).
- [47] Keskin B., Ozkan A. S., Inverse nodal problems for Dirac-type integro-differential operators, *J. Differential Equations*, 263 (2017) 8838–8847
- [48] Keskin B., Tel H. D., Reconstruction of the Dirac-Type Integro-Differential Operator From Nodal Data, *Numerical Functional Analysis and Optimization*, 39-11 (2018) 1208–1220 .

The Unit-transmuted Lindley Distribution with Applications

Emrah Altun ^{1,a,*}, Gholamhossein Hamedani ^{2,b}, Abdullah Fazlı ^{1,c}

¹Department of Mathematics, Bartın University, Bartın 74100, Türkiye.

²Department of Mathematics, Statistics and Computer Science, Marquette University, USA

*Corresponding author

Research Article

History

Received: 14/06/2024

Accepted: 15/10/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

This paper introduces the unit-transmuted Lindley (unit-TL) distribution. The statistical properties of the unit-TL distribution defined between (0,1) are discussed in detail. Several estimation techniques are used to estimate the parameters of the unit-TL distribution. An analysis through simulation is carried out to evaluate the efficacy of the suggested model. Furthermore, a unique regression model is developed for bounded response variables based on the unit-TL distribution. To illustrate the importance of the suggested model in precisely describing restricted datasets, two distinct datasets are examined.

Keywords: Proportion data, Transmuted Lindley, Least squares, Simulation.

^aemrahaltun@bartin.edu.tr

^b<https://orcid.org/0000-0001-5065-2523>

^cgholamhoss.hamedani@marquette.edu ^{id}<https://orcid.org/0000-0002-4435-8595>

^cabdullah.fazli74@gmail.com

^{id}<https://orcid.org/0000-0003-2079-6467>

Introduction

A lot of real-world scenarios require data (such as percentages and proportions) that are restricted to a particular range. This is especially true in economic environments where nondurable consumption is given a percentage of income and industry market shares [1]. Such datasets require the use of flexible distributions to correctly represent them. Despite being widely utilized in many scientific studies, the Beta distribution has certain drawbacks, most notably the inability to describe its functions in an explicit form. Compared to the Beta distribution, the Topp-Leone distribution is simpler and has become more well-known. It was first presented by [2]. Similar to this, the Kumaraswamy distribution was used by many scientist, which was first presented by [3] and made widespread by [4].

Researchers used some mathematical transformations to propose distributions on (0,1) interval. These are $Y = X/(X+1)$ and $Y = \exp(-X)$. Using these transformations, researchers introduced several distributions such as unit-BS by [5], log-XG by [6], log-exponential power by [7], log-Bilal [8], log-WE by [9], a new kind of unit-Lindley by [10], unit-Chen by [11], continuous Bernoulli by [12], unit generalized half-normal by [13] and unit Gompertz by [14].

This study aims to introduce an innovative distribution with closed and accessible statistical features, defined on the unit interval. Compared to popular distributions like Beta, Kumaraswamy, and Topp-Leone, this one has a number of advantages. A closed form of the probability density function (pdf) and moments are among the statistical characteristics of the unit-TL distribution that may be determined. It allows for the introduction of a

regression model of the proposed distribution and shows a better fit than other widely recognized distributions constructed on the unit interval.

The paper is divided as follows: A thorough summary of the mathematical characteristics connected to the suggested distribution is given in Section 2. The methods for estimating parameters, such as weighted least squares, least squares, and maximum likelihood, are covered in Section 3. The simulation study that evaluates the effectiveness of the parameter estimation techniques in finite samples are shown in Section 4. In Section 5, an innovative regression model that uses the generalized linear model approach is introduced as a complement to the Beta regression model. To demonstrate the adaptability of the unit-TL distribution in comparison to popular distributions in (0,1) interval, two real datasets are investigated in Section 6. The paper gets to its conclusion in Section 7.

The Unit-Transmuted Lindley Distribution

The pdf of the Lindley distribution is

$$f(x; \theta) = \frac{\theta^2}{1 + \theta} (1 + x) \exp(-\theta x), x > 0. \quad (1)$$

where $\theta > 0$ is scale parameter. It is possible to represent the pdf provided in (1) as a combination of gamma and exponential distributions. The cumulative distribution function (cdf) is

$$F(x) = 1 - \frac{\theta + 1 + \theta x}{\theta + 1} \exp(-\theta x), x \geq 0. \quad (2)$$

Several authors have examined the Lindley distribution and its statistical characteristics were deduced by [15]. There are several Lindley distribution generalizations in the statistical literature. [16] presented

a novel Lindley distribution, known as the transmuted Lindley (TL) distribution, with the following pdf using the transmutation idea.

$$f(x; \theta, \lambda) = \frac{\theta^2}{\theta + 1} (1 + x) \exp(-\theta x) \left(1 - \lambda + 2\lambda \frac{1 + \theta + \theta x}{\theta + 1} \exp(-\theta x) \right), x > 0. \tag{3}$$

Here, the parameter $|\lambda| \leq 1$ is the transmutation parameter. The cdf of (3) is

$$F(x; \theta, \lambda) = \left(1 - \frac{1 + \theta + \theta x}{\theta + 1} \exp(-\theta x) \right) \left(1 + \lambda \frac{1 + \theta + \theta x}{\theta + 1} \exp(-\theta x) \right), x \geq 0. \tag{4}$$

Now, using $Y = X/(X + 1)$ transformation in (3), we have

$$f(y; \theta, \lambda) = \frac{\theta^2}{\theta + 1} (1 - y)^{-3} \exp\left(-\frac{\theta y}{1 - y}\right) \left(1 - \lambda + 2\lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right), 0 < y < 1. \tag{5}$$

The Equation (5) is symbolized by $Y \sim \text{unit-TL}(\theta, \lambda)$. Its cdf is

$$F(y; \theta, \lambda) = \left(1 - \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) \left(1 + \lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right). \tag{6}$$

Figure 1 shows the pdf plots of the unit-TL distribution for various values of the parameters. Both a left- and a right-skewed unit-TL distribution are possible.

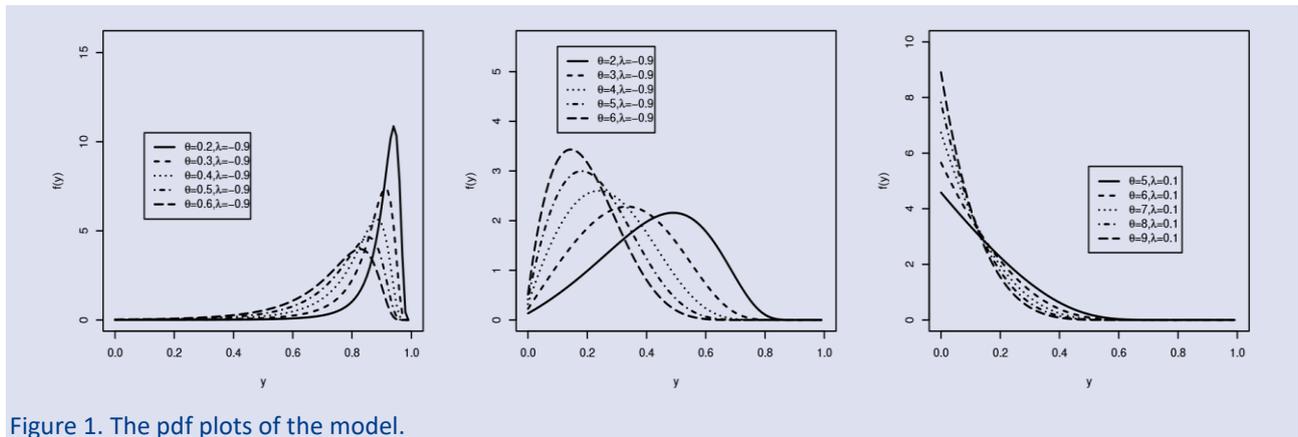


Figure 1. The pdf plots of the model.

The survival function is obtained as $S(y) = 1 - F(y)$, and given by

$$S(y; \theta, \lambda) = 1 - \left(1 - \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) \left(1 + \lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right), 0 \leq y \leq 1. \tag{7}$$

Using the pdf and cdf, the hazard rate function (hrf) is

$$H(y; \theta, \lambda) = \frac{\theta^2}{\theta + 1} (1 - y)^{-3} \exp\left(-\frac{\theta y}{1 - y}\right) \left(1 - \lambda + 2\lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) \times \left[1 - \left(1 - \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) \left(1 + \lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) \right]^{-1}. \tag{8}$$

The hrf forms of the unit-TL model is illustrated (see, Figure 2). The unit-TL has only an increasing form of the hrf.

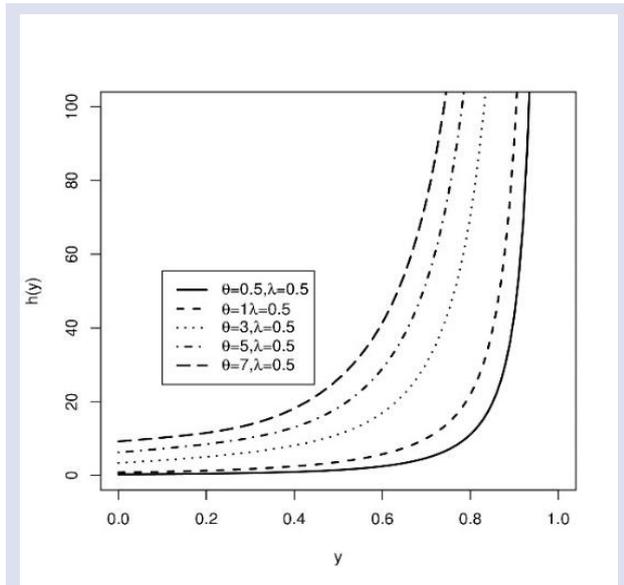


Figure 2. The hazard plots of unit-TL.

Generating random variables from the unit-TL distribution can be done using the properties of the Lindley and transmuted distributions. The cdf of the transmuted distribution is

$$F(x) = (1 + \lambda)G(x) - \lambda G^2(x) \tag{9}$$

where $|u| \leq 1$ and $G(x)$ is the cdf of the Lindley distribution for the transmuted-Lindley distribution. The quantile function of the transmuted distribution can be obtained for the general class of these distributions, as follows

$$Q_G = \left(\frac{\lambda - \sqrt{2\lambda - 4\lambda u + \lambda^2 + 1} + 1}{2\lambda} \right) \tag{10}$$

where $Q_G(\cdot)$ is the quantile function of the Lindley distribution for transmuted-Lindley distribution. The qlindley function defined in the LindleyR package of the R software can be used for that purpose. The below algorithm can be easily implemented in R software to generate random variables from the unit-TL distribution.

1. Define λ and θ parameters.
2. Generate random variables from standard uniform distribution, $u \sim U(0,1)$.
3. Calculate $x_u = \frac{\lambda - \sqrt{2\lambda - 4\lambda u + \lambda^2 + 1} + 1}{2\lambda}$
4. Using the quantile function of the Lindley distribution, calculate $x = \text{qlindley}(x_u, \theta)$
5. Apply the following transformation $y = \frac{x}{(x+1)}$
6. Repeat steps 2-5 n times.

The integration in (11) may be solved to obtain the moments of the unit-TL distribution.

$$E(Y^k) = \frac{\theta^2}{\theta + 1} \int_0^1 y^k (1 - y)^{-3} \exp\left(-\frac{\theta y}{1 - y}\right) \left(1 - \lambda + 2\lambda \frac{1 + \theta + \frac{\theta y}{1 - y}}{\theta + 1} \exp\left(-\frac{\theta y}{1 - y}\right) \right) dy. \tag{11}$$

(11) cannot have an analytical solution. For a given value of k , however, the integration in (11) can be solved. We have the mean of the unit-TL distribution for $k = 1$.

$$E(Y) = \frac{\theta \left(1 - \frac{\lambda}{2}\right) + 1}{(\theta + 1)^2} \tag{12}$$

Estimation

In this section, three parameter estimation methods are described for the unit-TL distribution.

Maximum likelihood estimation

Assume that the unit-TL distribution is the distribution of the random samples $y_1, y_2, y_3, \dots, y_n$. The $\ell(\theta, \lambda)$ function is provided by

$$\begin{aligned} \ell(\theta, \lambda) = & n \log\left(\frac{\theta^2}{\theta + 1}\right) - 3 \sum_{i=1}^n \log(1 - y_i) - \theta \sum_{i=1}^n \frac{y_i}{1 - y_i} \\ & + \sum_{i=1}^n \log\left(\left(1 - \lambda + 2\lambda \frac{1 + \theta + \frac{\theta y_i}{1 - y_i}}{\theta + 1} \exp\left(-\frac{\theta y_i}{1 - y_i}\right) \right)\right) \end{aligned} \tag{13}$$

We obtain the normal equations by calculating partial derivatives of (13) with regard to θ and λ .

$$\frac{\partial \ell}{\partial \theta} = n(\theta+1)\theta^{-2} \left(\frac{2\theta}{\theta+1} - \frac{\theta^2}{(\theta+1)^2} \right) - z_i + \frac{\left(\begin{matrix} -2\lambda y_i \exp(-\theta z_i)(\theta + \theta z_i + 1) \{(\theta+1)(1-y_i)\}^{-1} \\ -2\lambda \exp(-\theta z_i)(\theta + \theta z_i + 1)(\theta+1)^{-2} \\ +2\lambda \exp(-\theta z_i)(z_i + 1)(\theta+1)^{-1} \end{matrix} \right)}{1 - \lambda + 2\lambda y_i (\theta+1)^{-1} \exp(-\theta z_i)(\theta + \theta z_i + 1)},$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{2 \exp(-\theta z_i)(\theta + \theta z_i + 1)(\theta+1)^{-1} - 1}{1 - \lambda + 2\lambda \exp(-\theta z_i)(\theta + \theta z_i + 1)(\theta+1)^{-1}}.$$

Where $z_i = y_i / (1 - y_i)$ and $(\hat{\theta}, \hat{\lambda})$ are the maximum likelihood estimates (MLEs) of the following equations when they are solved simultaneously. These equations incorporate complex functions, which makes the solution impossible to get the MLEs in an explicit form. As such, the solution requires the application of numerical methods. Using the statistical software R, one may determine the MLEs of the parameters.

Least Squares (LS) and Weighted LS (WLS) Estimations

Let $Y_{(i)}$ for $i = 1, 2, 3, \dots, n$ be the random variable denotes the ordered samples from the unit-TL distribution. To obtain the LSEs estimations, the function given in (14) is minimized.

$$\sum_{i=1}^n \left[\left(1 - \frac{1 + \theta + \frac{\theta y_{(i)}}{1 - y_{(i)}}}{\theta + 1} \exp\left(-\frac{\theta y_{(i)}}{1 - y_{(i)}}\right) \right) \left(1 + \lambda \frac{1 + \theta + \frac{\theta y_{(i)}}{1 - y_{(i)}}}{\theta + 1} \exp\left(-\frac{\theta y_{(i)}}{1 - y_{(i)}}\right) \right) - \frac{i}{n + 1} \right]^2 \tag{14}$$

Also, the equation (15) is minimized to get WLSEs of the model parameters.

$$\sum_{i=1}^n \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[\left(1 - \frac{1 + \theta + \theta z_{(i)}}{\theta + 1} \exp(-\theta z_{(i)}) \right) \left(1 + \lambda \frac{1 + \theta + \theta z_{(i)}}{\theta + 1} \exp(-\theta z_{(i)}) \right) - \frac{i}{n+1} \right]^2 \tag{15}$$

Simulation

Using the estimation methods mentioned in previous section, bias, mean squared error (MSE) and mean relative error (MRE) are calculated. The anticipation is to observe that when the sample size increases, the bias and MSE should be near the zero value and MRE should be near the one value. The parameters of the unit-TL distribution are defined as $\theta = 2, \lambda = 0.5$. The size of the sample is raised from 50 to 500 by 5 units.

Figures 3-5 present a graphical summary of the simulation findings. Large sample sizes are associated with biases and MSEs that are, as predicted, close to zero in estimating methods.

The MREs are also close to the one. Nevertheless, compared to other estimate techniques, the MLE method’s biases and MSE values approach the required values more quickly. Furthermore, when the sample size is small, the bias and MSE of the MLE approach are lower than other methods. Therefore, we advise estimating the unit-TL distribution’s parameters using the MLE approach.

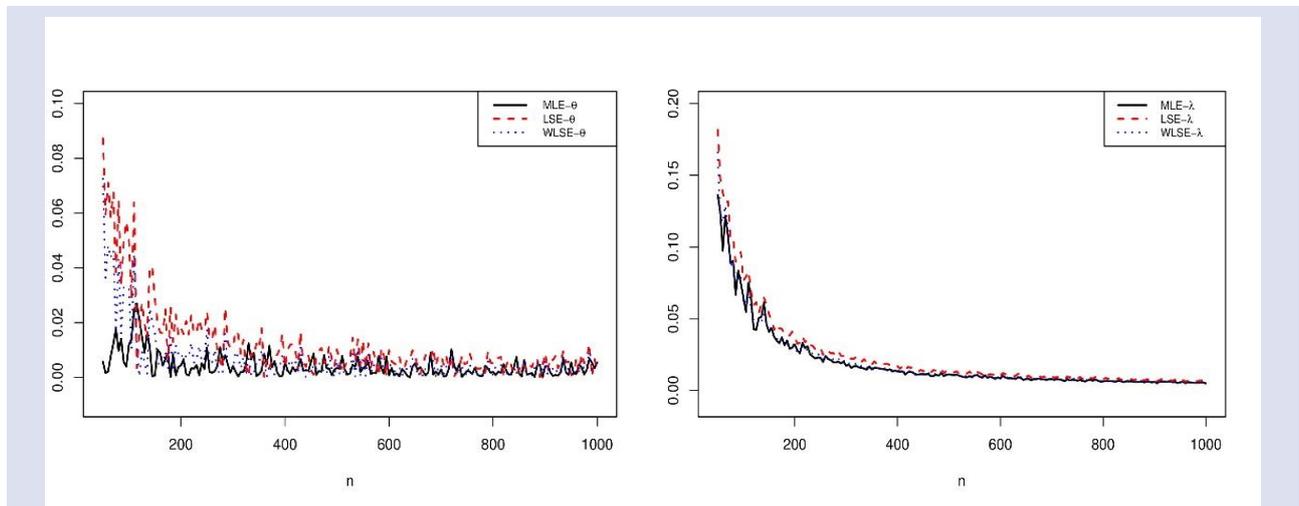


Figure 3. Estimated biases.

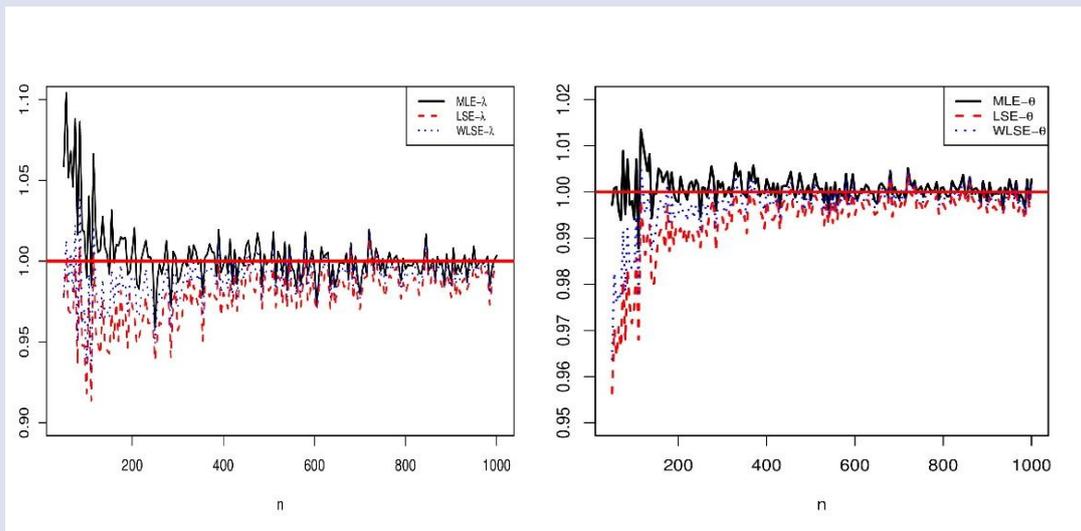


Figure 4. Estimated MREs.

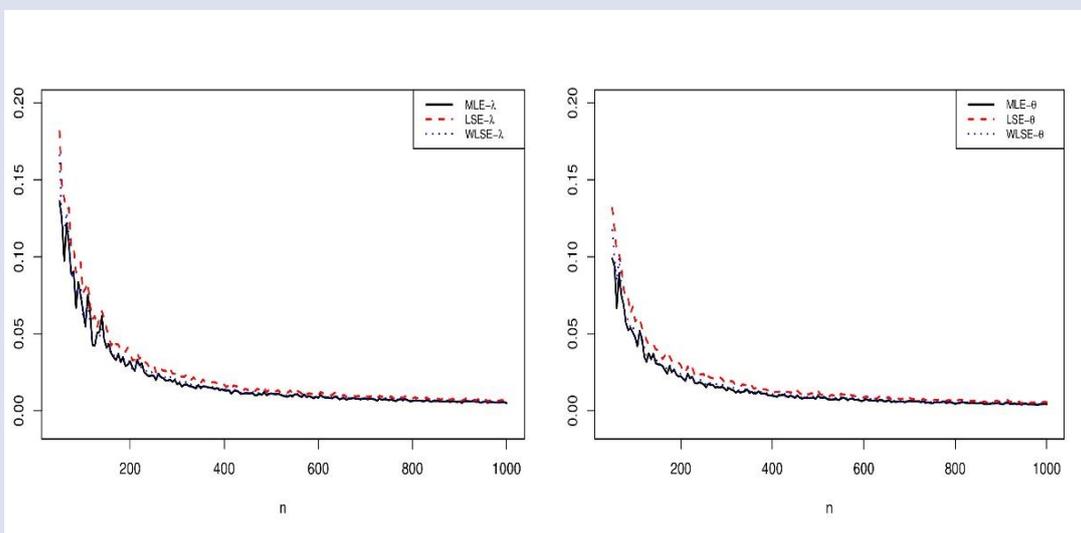


Figure 5. Estimated MSEs.

The Unit-Transmuted Lindley Regression Model

We present an alternative regression model that offers a new approach to modeling of bounded dependent variable with covariates. Let $\theta = (2\mu)^{-1} \left[\left(1 - \frac{\lambda}{2} - 2\mu\right) + \sqrt{4\mu + \left(1 - \frac{\lambda}{2}\right)^2 - 4\mu \left(1 - \frac{\lambda}{2}\right)} \right]$, the pdf is

$$f(y; \mu, \lambda) = \frac{\gamma(\mu, \lambda)^2}{\gamma(\mu, \lambda) + 1} (1 - y)^{-3} \exp\left(-\frac{\gamma(\mu, \lambda)y}{1-y}\right) \times \left(1 - \lambda + 2\lambda \frac{1 + \gamma(\mu, \lambda) + \frac{\gamma(\mu, \lambda)y}{1-y}}{\gamma(\mu, \lambda) + 1} \exp\left(-\frac{\gamma(\mu, \lambda)y}{1-y}\right)\right), \quad (16)$$

Where

$$\gamma(\mu, \lambda) = (2\mu)^{-1} \left[\left(1 - \frac{\lambda}{2} - 2\mu\right) + \sqrt{4\mu + \left(1 - \frac{\lambda}{2}\right)^2 - 4\mu \left(1 - \frac{\lambda}{2}\right)} \right]. \quad (17)$$

In the re-parametrization, we have $E(Y) = \mu$. As in the beta regression method, the independent variables are connected to the dependent variable via link function. Since the dependent variable is defined on (0,1) interval, we use the logit-link function

$$\mu_i = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}, i = 1, \dots, n. \quad (18)$$

Substituting (18) in (16), the log-likelihood function is

$$\begin{aligned} \ell(\beta, \lambda) = & \sum_{i=1}^n \ln \left(\frac{\gamma(\mu_i, \lambda)^2}{\gamma(\mu_i, \lambda) + 1} \right) - 3 \sum_{i=1}^n \ln(1 - y_i) - \sum_{i=1}^n \gamma(\mu_i, \lambda) \frac{y_i}{1 - y_i} \\ & + \sum_{i=1}^n \ln \left(1 - \lambda + 2\lambda \frac{1 + \gamma(\mu_i, \lambda) + \frac{\gamma(\mu_i, \lambda)y_i}{1 - y_i}}{\gamma(\mu_i, \lambda) + 1} \exp \left(-\frac{\gamma(\mu_i, \lambda)y_i}{1 - y_i} \right) \right) \end{aligned} \quad (19)$$

where μ_i is as in (18). The maximization of the log-likelihood function described in Equation (19) is achieved through the optim function in the R software. The asymptotic standard errors are then calculated by help of the inverse of the observed information matrix.

The residuals, as defined by [17] are employed to assess deviations from the assumption of error. These residuals are defined as:

$$\hat{e}_i = -\ln[1 - F(y_i)], i = 1, 2, \dots, n, \quad (20)$$

where $F(y_i)$ is the estimated cdf.

Applications

Water Capacity Data

This section employs a real data application to compare the unit-TL model with Beta, Kumaraswamy, and Topp-Leone distributions. Table 1 presents the MLEs, A, W, AIC and BIC. Here, A and W represent the Anderson-Darling and Cramer-von Mises, respectively. Also, KS represents the Kolmogorov-Smirnov test statistic value.

The dataset consists of monthly water capacity data from the Shasta reservoir in California, USA, spanning from February 1991 to 2010. The reservoir has a maximum capacity of 4552000 AF, and the data was normalized to the interval [0,1] using a normalization equation. [18] previously analyzed this dataset.

The hazard shape information can assist in selecting an appropriate model. [19] developed a useful tool for this purpose, called as TTT plot. A straight diagonal shape in the TTT plot indicates a constant hazard, while a convex shape suggests decreasing hazards and a concave shape indicates increasing hazards.

A bathtub shape in the hazard occurs when it transitions from convex to concave. Examination of Figure 6 reveals that the hrf shape of the data is increasing. Therefore, utilizing a unit-TL distribution appears to be an effective choice for modeling this dataset.

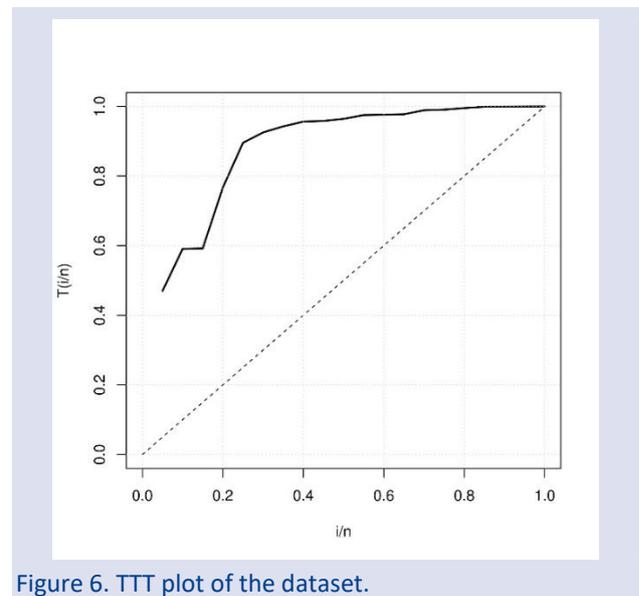


Figure 6. TTT plot of the dataset.

Table 1 displays the estimated parameters, standard errors, and other statistics for the monthly water capacity dataset. The values in Table 1 clearly show that the goodness-of-fit statistics for the unit-TL distribution are at their lowest values. For the monthly water capacity dataset, the proposed distribution may thus be thought to be the best fit model.

Table 1. Estimated parameters of the fitted models.

| Distributions | Parameters | $-\ell$ | AIC | BIC | A | W | K-S | p | |
|---------------|------------|---------|----------|----------|----------|--------|--------|--------|--------|
| Beta | 7.3154 | 2.9098 | -12.5619 | -21.1239 | -19.1324 | 1.6192 | 0.2796 | 0.2359 | 0.1834 |
| | 2.3180 | 0.8754 | | | | | | | |
| Kumaraswamy | 6.3476 | 4.4893 | -13.4747 | -22.9494 | -20.9580 | 1.4245 | 0.2407 | 0.2209 | 0.2447 |
| | 1.5575 | 2.0409 | | | | | | | |
| Topp-Leone | 8.6664 | | -11.5876 | -21.1753 | -20.1795 | 1.7864 | 0.3134 | 0.2549 | 0.1241 |
| | 1.9379 | | | | | | | | |
| unit-TL | 0.6764 | 0.9990 | -16.0715 | -28.1431 | -26.1516 | 0.9005 | 0.1394 | 0.1795 | 0.4847 |
| | 0.1120 | 0.3937 | | | | | | | |

Figure 7 displays the fitted functions of the model on the dataset. As can be seen from the right panel of Figure 7, out of all the models, the unit-TL distribution provides the best match to the monthly water capacity dataset.

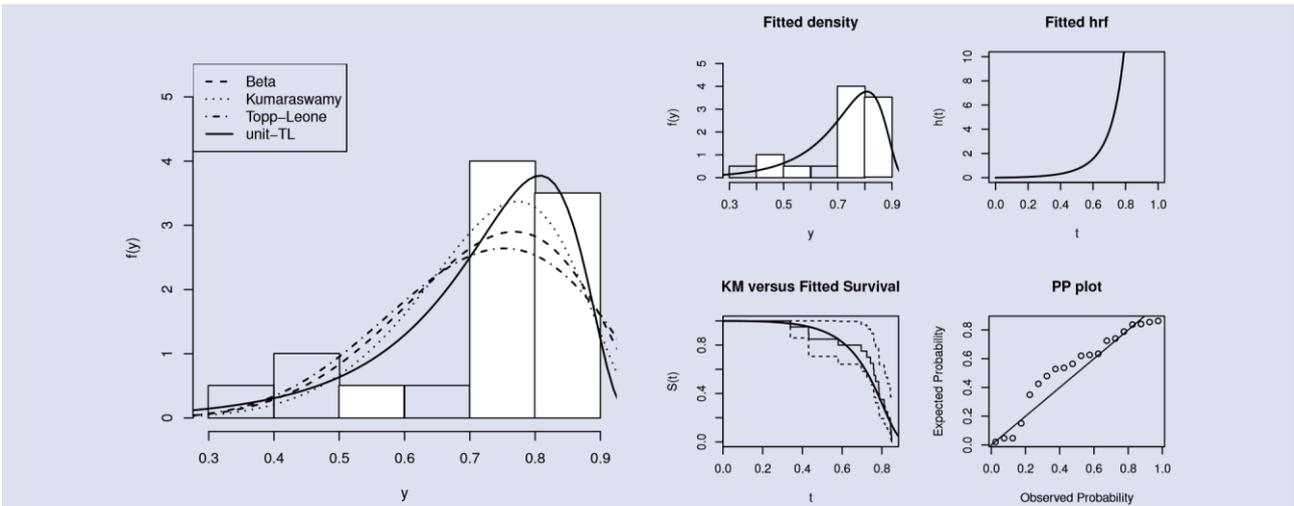


Figure 7: Comparison of the fitted densities (left-panel) and pdf, sf, hrf and PP plot of the unit-TL (right-panel).

Application of regression

In this part, the unit-TL regression model is checked against the beta regression model using the OECD Better Life Index (BLI) dataset. We use the R software’s betareg package to obtain the parameters of the beta regression model.

The purpose of this application is to determine how the variables of water quality (x_1), air pollution (x_2) and

murder rate (x_3) relate to self-reported health (y). One may get the dataset that was utilized in this study at the website of the OECD.

The logit link function is used in both regression models. As a result, the following represents the regression structure for μ_i .

$$\text{logit}(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} \quad (20)$$

Table 2: Results of the fitted regression models.

| Parameters | Beta | | | unit-TL | | |
|------------|----------|----------|--------|----------|----------|--------|
| | Estimate | SE | p | Estimate | SE | p |
| β_0 | 0.9068 | 1.1491 | 0.4300 | 0.9844 | 1.3438 | 0.4638 |
| β_1 | 0.5049 | 1.1506 | 0.6608 | 0.3943 | 1.3481 | 0.7699 |
| β_2 | -0.0424 | 0.0188 | 0.0237 | -0.0450 | 0.0221 | 0.0416 |
| β_3 | -0.6673 | 1.9246 | 0.7288 | -0.9082 | 2.1119 | 0.6672 |
| ϕ | 14.4130 | 3.2140 | <0.001 | - | - | - |
| λ | - | - | - | -0.9990 | 0.7302 | 0.1709 |
| $-\ell$ | - | -28.0600 | - | - | -30.1429 | - |
| AIC | - | -46.1200 | - | - | -50.2858 | - |
| BIC | - | -37.9321 | - | - | -42.0979 | - |

Table 2 provides a summary of the models. Interestingly, in both regression models, the parameter air pollution is found significant that has a detrimental effect on self-reported health.

Utilizing the computed AIC, and BIC values is essential in selecting the most suitable model. With the unit-TL distribution exhibiting the lowest values for these statistics, it is evident that the proposed model outperforms the alternative model in terms of fitting performance for the dataset. Furthermore, Figure 8 illustrates the Cox-Snell residuals for the unit-TL regression model, demonstrating the proximity of the plotted points to the diagonal line.

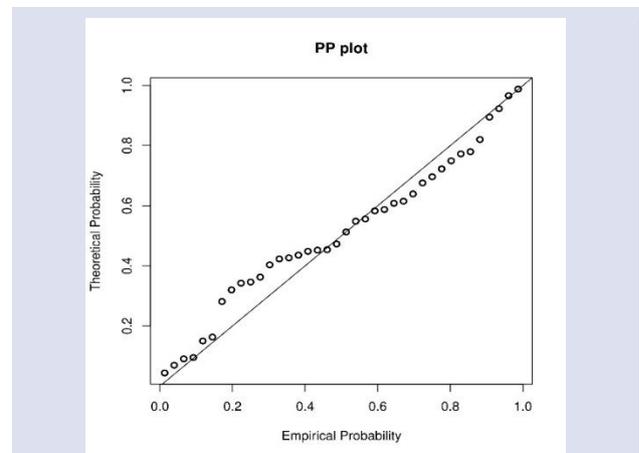


Figure 8. PP plot of the residuals.

Conclusion

We introduce a unique distribution with restricted support and two parameters. A number of statistical properties are obtained. Through a simulation study, the estimate of unknown parameters of the unit-TL distribution is explored using weighted least squares, maximum likelihood, and least squares approaches. To compare the performance and adaptability of the unit-TL distribution with competitive distributions, two datasets are analyzed. Additionally, the unit-TL regression constructed for the modeling of the bounded dependent variable is presented.

Conflict of Interest

There is no conflict of interest.

References

- [1] Papke L. E., Wooldridge J. M., Econometric methods for fractional response variables with an application to 401 (k) plan participation rates, *Journal of applied econometrics*, 11(6) (1996) 619-632.
- [2] Topp C. W., Leone F. C., A family of J-shaped frequency functions, *Journal of the American Statistical Association*, 50(269) (1955) 209-219.
- [3] Kumaraswamy P., A generalized probability density function for double-bounded random processes, *Journal of Hydrology*, 46(1-2) (1980) 79-88.
- [4] Cordeiro G. M, de Castro M., A new family of generalized distributions, *Journal of statistical computation and simulation*, 81(7) (2011) 883-898.
- [5] Mazucheli J., Menezes A. F., Dey S., The unit-Birnbaum-Saunders distribution with applications, *Chilean Journal of Statistics (ChJS)*, 9(1) (2018) 47-57.
- [6] Altun E., Hamedani G. G., The log-xgamma distribution with inference and application, *Journal de la Societe Francaise de Statistique*, 159(3) (2018) 40-55.
- [7] Korkmaz M. C., Altun E., Alizadeh M., El-Morshedy M., The log exponential power distribution: Properties, estimations and quantile regression model, *Mathematics*, 9(21) (2021) 2634.
- [8] Altun E., El-Morshedy, M., Eliwa M. S., A new regression model for bounded response variable: An alternative to the beta and unit-Lindley regression models, *Plos One*, 16(1) (2021) e0245627.
- [9] Altun E., The log-weighted exponential regression model: alternative to the beta regression model, *Communications in Statistics-Theory and Methods*, 50(10) (2021) 2306-2321.
- [10] Altun E., Cordeiro G. M., The unit-improved second-degree Lindley distribution: inference and regression modeling, *Computational Statistics*, 35 (2020) 259-279.
- [11] Korkmaz M. C., Altun E., Chesneau C., Yousof H. M., On the unit-Chen distribution with associated quantile regression and applications, *Mathematica Slovaca*, 72(3) (2022) 765-786.
- [12] Korkmaz M. Ç., Leiva V., Martin-Barreiro C., The continuous Bernoulli distribution: Mathematical characterization, fractile regression, computational simulations, and applications. *Fractal and Fractional*, 7(5) (2023) 386.
- [13] Mazucheli J., Korkmaz M.Ç., Menezes A.F., Leiva V., The unit generalized half-normal quantile regression model: formulation, estimation, diagnostics, and numerical applications. *Soft Computing*, 27(1) (2023) 279-295.
- [14] Mazucheli J., Alves B., Korkmaz M.Ç. (2023). The Unit-Gompertz Quantile Regression Model for the Bounded Responses. *Mathematica Slovaca*, 73(4) (2023) 1039-1054.
- [15] Ghitany, M. E., Atieh, B., Nadarajah, S., Lindley distribution and its application, *Mathematics and computers in simulation*, 78(4) (2008) 493-506.
- [16] Merovci F., Transmuted lindley distribution, *Int. J. Open Problems Compt. Math*, 6(2) (2013) 63-72.
- [17] Cox D. R., Snell E. J., A general definition of residuals, *Journal of the Royal Statistical Society. Series B (Methodological)*, (1968) 248-275.
- [18] Nadar M., Papadopoulos A., Kızılaslan F., Statistical analysis for Kumaraswamy's distribution based on record data, *Statistical Papers*, 54(2) (2013) 355-369.
- [19] Aarset A.S., How to identify a bathtub hazard rate, *IEEE Transactions on Reliability*, 36 (1987) 106-108.

Modified Ridge Estimator for Poisson Regression

Shuaib Mursal Ibrahim ^{1,a,*}, Aydın Karakoca ^{2,b}

¹ SIMAD University, Mogadishu, Somalia

² Necmettin Erbakan University, Konya, Türkiye

*Corresponding author

Research Article

History

Received: 06/10/2023

Accepted: 18/11/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Poisson regression is a statistical model used to model the relationship between a count-valued-dependent variable and one or more independent variables. A frequently encountered problem when modeling such relationships is multicollinearity, which occurs when the independent variables are highly correlated with each other. Multicollinearity can affect the maximum likelihood (ML) estimates of unknown model parameters, making them unstable and inaccurate. In this study, we propose a modified ridge parameter estimator to combat multicollinearity in Poisson regression. We conducted extensive simulations to evaluate the performance of our proposed estimator using the mean squared error (MSE). We also apply our estimator to real data. The results show that our proposed estimator outperforms the ML estimator in both simulations and real data applications.

Keywords: Poisson regression, Multicollinearity, Ridge estimator, Monte Carlo simulations, Maximum likelihood estimation.

 shuaib.mursal@gmail.com

 <https://orcid.org/0000-0002-0258-034X>

 akarakoca@erbakan.edu.tr

 <https://orcid.org/0000-0001-6503-3872>

Introduction

Poisson regression is a statistical technique used to model the relationship between a count-valued response variable and one or more independent variables [1]. It is commonly used in a variety of fields, such as epidemiology, marketing, and finance, to model the incidence of diseases, the number of customers who visit a store, and the number of financial transactions that occur in a given period of time.

One of the key assumptions of Poisson regression is that the independent variables are not correlated. However, in practice, it is common for independent variables to be correlated, which is known as multicollinearity [2]. Multicollinearity can lead to several problems, including unstable coefficient estimates, inflated standard errors, and decreased statistical power [3].

Ridge regression is a regularization technique that can be used to reduce the effects of multicollinearity in Poisson regression models [4]. Ridge regression works by adding a small penalty term to the ML function, which shrinks the coefficient estimates towards zero [5]. This shrinkage can help improve the stability of the coefficient estimates and reduce the impact of multicollinearity in the model.

A number of different ridge estimators have been proposed for Poisson regression models [6-14]. However, the performance of these estimators can vary depending on the specific characteristics of the data.

The purpose of this study is to evaluate the performance of existing ridge estimators and propose new ridge estimators that are more effective in combating multicollinearity in Poisson regression models. Using

Monte Carlo simulations, we compare the performance of the different estimators in terms of their MSEs.

Materials and Methods

Poisson Regression

A random variable Y is said to follow a Poisson probability distribution if the vector of $Y = [y_1, y_2, y_3, \dots, y_n]^T$ are count numbers, and the y_i 's are independent and identically distributed probability mass function with:

$$f(Y = y) = \frac{e^{-\mu} \mu^y}{y!}, \quad \mu > 0 \quad (1)$$

where μ is the mean occurrence of the event y for a specified interval of time and e is the base of the natural logarithm.

The mean and variance of the Poisson distribution are equal:

$$E(Y) = E(Var) = \mu \quad (2)$$

Poisson regression is a statistical model used to model count data. Instead of modeling the linear relationship between the dependent variable and the independent variable(s), Poisson regression models the linear relationship between the log of the expected value of the dependent variable and the independent variable(s) [15]. This ensures that the expected value of the dependent variable is always positive.

The Poisson model can be written as follows:

$$\ln(\mu_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} = x_i \beta \quad (3)$$

where x_i is the i^{th} row of the design matrix X .

The ML method is used to estimate the Poisson regression model coefficients. The likelihood function is given by:

$$L(\beta) = \prod_{i=1}^n f(Y = y) \quad (4)$$

Maximizing the likelihood is the same as maximizing the log-likelihood.

$$l(\beta) = \log\left[\prod_{i=1}^n f(Y = y)\right] \quad (5)$$

From Equation (1), the likelihood is:

$$L(\beta) = \prod_{i=1}^n \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \quad (6)$$

The log likelihood is

$$l(\beta) = \sum_{i=1}^n [-\mu_i + y_i \log(\mu_i) - \log(y_i)] \quad (7)$$

substituting $\mu_i = \exp(x_i \beta)$ from equation (3) with equation (7):

$$l(\beta) = \sum_{i=1}^n [-\exp(x_i \beta) + y_i(x_i \beta) - \log(y_i)] \quad (8)$$

The ML estimates for the coefficients are obtained by taking the first derivative of the log-likelihood function and equating it to zero. This results in the following equations:

$$\frac{\partial l(\beta)}{\partial \beta} = \sum_{i=1}^n [y_i - \exp(x_i \beta)] x_i = 0 \quad (9)$$

These equations are called score equations and can be solved numerically using iterative methods such as iteratively reweighted least squares (IRLS).

The IRLS algorithm produces the following Poisson Maximum Likelihood Estimator (PMLE) for β :

$$\hat{\beta}_{PMLE} = (X^t W X)^{-1} X^t W z \quad (10)$$

where $z = \log(\mu_i) + \frac{y_i - \mu_i}{\mu_i}$ and W is a diagonal matrix with weights $w_{ii} = \exp(x_i \beta)$.

The PMLE is asymptotically normally distributed with a covariance matrix equal to the negative inverse of the Hessian matrix:

$$I_Y(\beta) = -E \left[\sum_{i=1}^n -\exp(x_i \beta) x_i x_i^t \right] = (X^t W X)^{-1} \quad (11)$$

The MSE of the PMLE is given by:

$$MSE(\beta_{MLE}) = E \left[(\hat{\beta}_{PMLE} - \beta)^t (\hat{\beta}_{PMLE} - \beta) \right] = \text{trace}(X^t W X)^{-1} = \sum_{i=1}^{p+1} \frac{1}{\lambda_i} \quad (12)$$

where λ_i is the i^{th} eigenvalue of the matrix $X^t W X$.

Poisson Ridge Regression

When the independent variables in a Poisson regression model are highly correlated, the $X^t W X$ matrix becomes ill-conditioned, meaning that its determinant is close to zero. This can lead to an inflated variance and instability of the ML estimator.

To address this issue, [4] introduced the concept of ridge regression. Ridge regression involves adding a small positive term, known as the ridge parameter, to the diagonal of the $X^t X$ matrix in linear regression.

[16] proposed the Poisson Ridge Regression Estimator (PRRE) to address the problem of multicollinearity in the Poisson regression model. The PRRE is defined as follows:

$$\hat{\beta}_{PRRE} = (X^t W X + kI)^{-1} X^t W X \hat{\beta}_{PMLE} \quad (13)$$

where $k > 0$ and I represents the $p \times p$ identity matrix. When the parameter k is set to 0, the PRRE simplifies the ML estimator. The PRRE helps alleviate the multicollinearity issue and improves the stability of the estimator. The MSE of the PRRE can be calculated as follows:

$$MSE(\beta_{PRRE}) = E \left[(\hat{\beta}_{PRRE} - \beta)^t (\hat{\beta}_{PRRE} - \beta) \right] = \sum_{j=1}^{p+1} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^{p+1} \frac{\alpha_j^2}{(\lambda_j + k)^2} \quad (14)$$

where α_j represents $U \hat{\beta}_{PMLE}$, where U is a matrix, whose columns are the eigenvectors of the $X^t W X$ matrix. The values λ_j correspond to the eigenvalues associated with the eigenvectors of the $X^t W X$ matrix.

For $k > 0$, the MSE of the PRRE is consistently lower than that of the Poisson regression ML estimator. This indicates that the PRRE provides improved accuracy and precision compared with the ML estimator in the Poisson regression model.

Some Existing Ridge Estimators

In this section, we have conducted a systematic review of 30 existing ridge estimators. Inspired by the works of

[9,16,17]. We have proposed modified versions of some of these estimators. Table 1 summarizes the existing ridge estimators we examined.

Table 1. Summary of Existing Ridge Estimators

| No. | Estimator | Reference | No. | Estimator | Reference |
|-----|--|-----------|-----|---|-----------|
| 1 | $k = 0$ | MLE [14] | 17 | $k_{GK} = k_{HK} + \frac{2}{(\lambda_{max} + \lambda_{min})}$ | [18] |
| 2 | $k_{HK1} = \frac{1}{\hat{a}_{max}^2}$ | [4] | 18 | $k_{D1} = \frac{2}{\sum_{i=1}^p \lambda_{max} \hat{a}_i^2}$ | [19] |
| 3 | $k_{HK2} = \frac{1}{\sum_{i=1}^p \hat{a}_{max}^2}$ | [4] | 19 | $k_{D2} = \text{Median} \left(\frac{2}{\lambda_{max} \hat{a}_i^2} \right)$ | [19] |
| 4 | $k_{HKB} = \frac{p}{\sum_{i=1}^p \hat{a}_{max}^2}$ | [4] | 20 | $k_{D3} = \frac{2}{\lambda_{max} (\prod_{i=1}^p \hat{a}_i^2)^{\frac{1}{p}}}$ | [19] |
| 5 | $k_{LW1} = \frac{1}{\lambda_i \hat{a}_i^2}$ | [20] | 21 | $k_{D4} = \frac{2}{\lambda_{max} p} \sum_{i=1}^p \frac{1}{\hat{a}_i^2}$ | [19] |
| 6 | $k_{LW2} = \frac{p}{\sum_{i=1}^p \lambda_i \hat{a}_i^2}$ | [20] | 22 | $k_{Y4} = \text{Max} \left(\sqrt{\frac{1}{\lambda_i \hat{a}_i^2}} \right)$ | [17] |
| 7 | $k_{HSL} = \frac{\sum_{i=1}^p (\lambda_i \hat{a}_i)^2}{(\sum_{i=1}^p \lambda_i \hat{a}_i^2)^2}$ | [21] | 23 | $k_{Y6} = \text{Max} \left(\sqrt{\lambda_i \hat{a}_i^2} \right)$ | [17] |
| 8 | $k_{AM} = \frac{1}{p} \sum_{i=1}^p \frac{1}{\hat{a}_i^2}$ | [6] | 24 | $k_{Y9} = \frac{p}{\sum_{i=1}^p \sqrt{\frac{1}{\lambda_i \hat{a}_i^2}}}$ | [17] |
| 9 | $k_{GM} = \frac{1}{(\prod_{i=1}^p \hat{a}_i^2)^{\frac{1}{p}}}$ | [6] | 25 | $k_{AS1} = \frac{1}{\hat{a}_{max}} + \frac{1}{\lambda_{max}}$ | [22] |
| 10 | $k_{KS} = \frac{\lambda_{max}}{(n-p) + \lambda_{max} \hat{a}_{max}^2}$ | [7] | 26 | $k_{AS2} = \text{Max} \left(\frac{1}{\hat{a}_i} + \frac{1}{\lambda_i} \right)$ | [22] |
| 11 | $k_{A1} = \frac{1}{p} \sum_{i=1}^p \left(\frac{\lambda_{max}}{(n-p) + \lambda_{max} \hat{a}_{max}^2} \right)$ | [8] | 27 | $k_{AS3} = \frac{1}{\text{Min} \left(\frac{1}{\hat{a}_i} + \frac{1}{\lambda_i} \right)}$ | [22] |
| 12 | $k_{A2} = \text{Max} \left(\frac{\lambda_i}{(n-p) + \lambda_i \hat{a}_i^2} \right)$ | [8] | 28 | $k_{AY1} = \frac{\sqrt{5}p}{\lambda_{max} \sum_{i=1}^p \hat{a}_i^2}$ | [23] |
| 13 | $k_{A3} = \text{Median} \left(\frac{\lambda_i}{(n-p) + \lambda_i \hat{a}_i^2} \right)$ | [8] | 29 | $k_{AY2} = \frac{p}{\sqrt{\lambda_{max} \sum_{i=1}^p \hat{a}_i^2}}$ | [23] |
| 14 | $k_{MK4} = \left(\prod_{i=1}^p \sqrt{\hat{a}_i^2} \right)^{\frac{1}{p}}$ | [9] | 30 | $k_{AY3} = \frac{2p}{\sum_{i=1}^p \left(\lambda_i^{\frac{1}{4}} \right) \sum_{i=1}^p \hat{a}_i^2}$ | [23] |
| 15 | $k_{MK5} = \left(\prod_{i=1}^p \frac{1}{\sqrt{\hat{a}_i^2}} \right)^{\frac{1}{p}}$ | [9] | 31 | $k_{AY4} = \frac{2p}{\sqrt{\sum_{i=1}^p \lambda_i \sum_{i=1}^p \hat{a}_i^2}}$ | [23] |
| 16 | $k_{MK6} = \text{Median} \left(\sqrt{\hat{a}_i^2} \right)$ | [9] | | | |

Proposed Ridge Estimators

By incorporating insights from these previous works, we have devised novel modifications to the existing estimators, aiming to enhance their performance. Based on work of Asar and Genç (2017) we applied square root transformation and the absolute value of $\hat{\alpha}$ and proposed the following ridge estimators:

$$\begin{aligned}
 k_{SK1} &= \text{Max} \left(\frac{1}{\sqrt{\text{abs}(\hat{\alpha}_i)}} \right) & k_{SK2} &= \text{Median} \left(\frac{1}{\sqrt{\text{abs}(\hat{\alpha}_i)}} \right) \\
 k_{SK3} &= \frac{1}{p} \sum_{i=1}^p \frac{1}{\sqrt{\text{abs}(\hat{\alpha}_i)}} & k_{SK4} &= \left(\prod_{i=1}^p \frac{1}{\sqrt{\text{abs}(\hat{\alpha}_i)}} \right)^{\frac{1}{p}} \\
 k_{SK5} &= \frac{p}{\sum_{i=1}^p \left(\frac{1}{\sqrt{\text{abs}(\hat{\alpha}_i)}} \right)} & k_{SK7} &= \sqrt{\sum_{i=1}^p \hat{\alpha}_i^2 \sum_{i=1}^p \lambda_i}
 \end{aligned}$$

Simulation Design

To investigate the performance of Poisson regression's MSE under multicollinearity, we generated the dependent variable, Y_i , from a Poisson distribution with parameter μ_i . Here, μ_i is equal to $\exp(x_i\beta)$, where x_i represents the i^{th} row of the design matrix X , $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ is the vector of coefficients, and i ranges from 1 to n .

To control the correlation between the independent variables, we generated them using the following equation:

$$x_{ij} = (1 - \rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{i,p} \tag{15}$$

where i ranges from 1 to n , and j ranges from 1 to p . In this equation, z_{ij} represents independent pseudo-normal random variables, and ρ^2 represents the degree of correlation between any two random variables. In order to observe how the severity of correlation affects the performance of Poisson MSE, we considered four levels of correlation: $\rho^2 = 0.90, 0.95, 0.99$, and 0.999 .

Apart from considering the degree of correlation, we also examined the effects of the Poisson intercept β_0 , the number of independent variables (p) and the number of observations (n) on the performance of the Poisson MSE. Specifically, we chose different values for the intercept β_0 , including $-1, 0$, and 1 . As the intercept decreases, the average value of μ_i decreases as well. Consequently, this decrease in average value results in a higher frequency of zero values for the dependent variable. This can lead to convergence problems in the iteratively reweighted least squares (IRLS) algorithm. The slope coefficients $(\beta_0, \beta_1, \dots, \beta_p)'$ were selected in a way that ensures the sum of their squares $\sum_{j=0}^p \beta_j^2 = 1$.

To analyze the performance of the Poisson MSE, we generated two models, one with 5 independent variables and another with 8 independent variables. Additionally, we varied the sample size, with n set to 50, 100, and 200 for each model.

For conducting these simulations, we utilized MATLAB as the computational tool. The MSE was calculated using the formula:

$$MSE(\hat{\beta}) = \frac{1}{5000} \sum_{j=1}^{5000} (\hat{\beta} - \beta)' (\hat{\beta} - \beta) \tag{16}$$

Results

The results of the simulation study are displayed in Tables 2-7. Four factors that affect the performance of the Poisson MSE estimator—namely, sample size, degree of correlation, the intercept of the Poisson model, and the number of independent variables—were considered and examined through Monte Carlo simulation. The findings are summarized as follows:

The effect of the intercept on the MSE is positive. In other words, as the intercept increases from -1 to 1 , the MSE of all estimators generally decreases.

When considering the effect of the degree of correlation, the results show that as the degree of correlation increases, the MSE also increases. However, this pattern does not hold true for all estimators. For example, in cases where the sample size is small ($n = 50$) or the number of independent variables is large, the MSE of the estimators k_{AM} , k_{GM} , k_{A4} , k_{Y4} , k_{A6} , k_{AS1} , k_{SK1} , and k_{SK6} decrease as the degree of correlation increases, especially if the intercept value is kept at 0 or 1. The MSE of PMLE is the worst performing estimator, while k_{SK6} is the best performing estimator when the intercept is 1 regardless of other factors, k_{SK1} is the best performing estimator when the intercept is -1 , and k_{Y6} is the best performing estimator when the intercept is 0.

The effect of sample size on the MSE is desirable. As the sample size increases, the MSE also decreases. For example, when the degree of correlation is 0.999 and all other factors are kept constant, the MSE decreases significantly, especially for PMLE. On the other hand, as the number of independent variables increases, the MSE also increases. The PMLE estimator has the worst performance in this case.

When considering the estimator with the highest performance across all instances, it is evident that estimator k_{SK6} outperforms all other estimators. Additionally, when analyzing the estimators that secure the top three positions in terms of MSE performance, estimators k_{SK1} and k_{Y6} had consistently ranked second as shown in Figure. 1.

Real Data Application

In this section we apply a real dataset to examine the effectiveness of our proposed estimators. We used the aircraft damage dataset introduced by [15]. This dataset has been used in previous studies by [11], [13] and others. The dataset contains information about 30 strike missions flown by two types of aircraft: the McDonnell Douglas A-4 Skyhawk and the Grumman A-6 Intruder. The explanatory variables in the dataset are defined as follows: x_1 serves as a binary variable indicating the

aircraft type, where A-4 is coded as 0 and A-6 is coded as 1. Additionally, x_2 and x_3 represent the bomb load in tons and the total months of aircrew experience, respectively. The response variable (y) is the number of locations where damage was inflicted on the aircraft. [11] claims that the data suffers from multicollinearity. The eigenvalues of the design matrix of X^tWX is [4.289 789.849 283543.296]. This is supported by the high condition number of the design matrix, which is 257.125. The condition number, CN, is a measure of multicollinearity, and is calculated as follows:

$$CN = \sqrt{\frac{\max(\text{eigenvalue})}{\min(\text{eigenvalue})}} \quad (17)$$

Where $\max(\text{eigenvalue})$ is the largest eigenvalue of the design matrix X^tWX , and $\min(\text{eigenvalue})$ is the smallest eigenvalue of the design matrix X^tWX . A condition number greater than 30 is generally considered to indicate the presence of multicollinearity [11], [13].

The performance of the estimators was assessed using MSE, which was computed using Equation 14. The estimated coefficients are presented in Table 8. The results of the application showed that our proposed ridge estimator k_{SK1} has the smallest MSE value, which shows its superiority in real applications. The worst performing estimator is the Poisson maximum likelihood estimator (PMLE).

Table 2: Estimated MSE when $n = 50$ and $p = 4$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|-----------|-------|-------|--------|---------|-------|-------|--------|--------|-------|-------|--------|--------|
| | 0.900 | | ρ | | 0.900 | | ρ | | 0.900 | | ρ | |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| PMLE | 1.061 | 2.125 | 11.978 | 125.769 | 0.380 | 0.790 | 4.219 | 46.588 | 0.134 | 0.291 | 1.500 | 16.104 |
| k_{HK1} | 0.577 | 0.935 | 3.964 | 38.241 | 0.152 | 0.233 | 0.749 | 8.017 | 0.128 | 0.134 | 0.201 | 1.091 |
| k_{HK2} | 0.825 | 1.504 | 7.280 | 71.456 | 0.310 | 0.585 | 2.611 | 27.180 | 0.126 | 0.259 | 1.087 | 9.501 |
| k_{HKB} | 0.450 | 0.692 | 2.740 | 25.452 | 0.179 | 0.286 | 1.002 | 10.108 | 0.103 | 0.183 | 0.511 | 3.520 |
| k_{LW1} | 1.055 | 2.103 | 11.309 | 87.446 | 0.380 | 0.789 | 4.203 | 44.720 | 0.134 | 0.291 | 1.500 | 16.057 |
| k_{LW2} | 1.004 | 1.943 | 10.306 | 103.921 | 0.371 | 0.757 | 3.818 | 39.231 | 0.133 | 0.286 | 1.421 | 13.988 |
| k_{HSL} | 1.061 | 2.125 | 11.970 | 125.132 | 0.380 | 0.790 | 4.219 | 46.585 | 0.134 | 0.291 | 1.500 | 16.104 |
| k_{AM} | 0.608 | 0.608 | 0.599 | 0.386 | 0.112 | 0.111 | 0.119 | 0.084 | 0.223 | 0.227 | 0.226 | 0.420 |
| k_{GM} | 0.663 | 0.893 | 0.698 | 0.159 | 0.360 | 0.665 | 1.475 | 0.484 | 0.131 | 0.273 | 0.855 | 0.518 |
| k_{KS} | 0.552 | 0.924 | 3.906 | 36.806 | 0.218 | 0.390 | 1.426 | 14.131 | 0.108 | 0.209 | 0.719 | 4.936 |
| k_{A1} | 0.851 | 1.490 | 6.398 | 28.575 | 0.237 | 0.450 | 1.741 | 6.474 | 0.133 | 0.228 | 0.928 | 7.937 |
| k_{A2} | 0.757 | 1.182 | 4.194 | 10.671 | 0.180 | 0.308 | 0.950 | 1.892 | 0.147 | 0.209 | 0.669 | 4.911 |
| k_{A3} | 1.002 | 1.930 | 8.932 | 58.914 | 0.340 | 0.682 | 3.249 | 24.876 | 0.131 | 0.282 | 1.376 | 11.537 |
| k_{MK4} | 0.570 | 0.734 | 0.598 | 0.120 | 0.333 | 0.568 | 1.020 | 0.338 | 0.127 | 0.255 | 0.692 | 0.468 |
| k_{MK5} | 0.369 | 0.459 | 0.662 | 0.641 | 0.162 | 0.231 | 0.388 | 0.470 | 0.096 | 0.160 | 0.359 | 0.659 |
| k_{MK6} | 0.531 | 0.672 | 0.571 | 0.138 | 0.328 | 0.550 | 0.905 | 0.283 | 0.125 | 0.244 | 0.617 | 0.441 |
| k_{GK} | 0.576 | 0.929 | 3.792 | 28.126 | 0.152 | 0.233 | 0.745 | 7.573 | 0.128 | 0.134 | 0.201 | 1.086 |
| k_{D1} | 1.057 | 2.113 | 11.862 | 124.186 | 0.380 | 0.788 | 4.205 | 46.389 | 0.134 | 0.290 | 1.499 | 16.080 |
| k_{D2} | 1.052 | 2.096 | 11.434 | 106.557 | 0.379 | 0.784 | 4.148 | 44.246 | 0.133 | 0.290 | 1.498 | 15.979 |
| k_{D3} | 1.049 | 2.089 | 11.477 | 112.982 | 0.378 | 0.784 | 4.148 | 44.611 | 0.133 | 0.290 | 1.496 | 15.987 |
| k_{D4} | 0.993 | 1.917 | 10.053 | 85.073 | 0.348 | 0.720 | 3.629 | 32.734 | 0.132 | 0.279 | 1.422 | 15.056 |
| k_{Y4} | 0.430 | 0.458 | 0.410 | 0.275 | 0.167 | 0.220 | 0.243 | 0.114 | 0.102 | 0.159 | 0.263 | 0.192 |
| k_{Y6} | 0.678 | 0.661 | 0.634 | 0.615 | 0.040 | 0.036 | 0.031 | 0.030 | 0.164 | 0.160 | 0.149 | 0.143 |
| k_{Y9} | 0.438 | 0.639 | 1.055 | 1.015 | 0.282 | 0.453 | 0.983 | 1.479 | 0.113 | 0.215 | 0.583 | 1.107 |
| k_{AS1} | 0.676 | 1.210 | 6.149 | 54.433 | 0.274 | 0.491 | 2.212 | 23.787 | 0.122 | 0.240 | 0.879 | 8.212 |
| k_{AS2} | 0.827 | 0.789 | 0.587 | 1.067 | 0.166 | 0.140 | 0.095 | 0.057 | 0.361 | 0.353 | 0.245 | 0.093 |
| k_{AS3} | 0.417 | 0.507 | 0.503 | 0.178 | 0.271 | 0.414 | 0.664 | 0.371 | 0.110 | 0.195 | 0.384 | 0.293 |
| k_{AY1} | 1.056 | 2.112 | 11.849 | 124.009 | 0.380 | 0.788 | 4.203 | 46.366 | 0.134 | 0.290 | 1.499 | 16.078 |
| k_{AY2} | 1.015 | 1.993 | 10.797 | 110.642 | 0.372 | 0.763 | 3.967 | 43.154 | 0.133 | 0.288 | 1.466 | 15.396 |
| k_{AY3} | 0.859 | 1.563 | 7.345 | 68.591 | 0.332 | 0.637 | 2.869 | 28.925 | 0.130 | 0.273 | 1.222 | 10.988 |
| k_{AY4} | 0.975 | 1.885 | 9.921 | 100.193 | 0.365 | 0.739 | 3.761 | 40.522 | 0.133 | 0.286 | 1.434 | 14.799 |
| k_{SK1} | 0.328 | 0.369 | 0.374 | 0.159 | 0.148 | 0.198 | 0.241 | 0.105 | 0.089 | 0.138 | 0.265 | 0.296 |
| k_{SK2} | 0.428 | 0.534 | 0.613 | 0.354 | 0.221 | 0.326 | 0.540 | 0.438 | 0.110 | 0.201 | 0.464 | 0.571 |
| k_{SK3} | 0.391 | 0.479 | 0.560 | 0.287 | 0.201 | 0.288 | 0.428 | 0.283 | 0.104 | 0.179 | 0.395 | 0.512 |
| k_{SK4} | 0.409 | 0.509 | 0.610 | 0.341 | 0.217 | 0.315 | 0.491 | 0.380 | 0.108 | 0.191 | 0.430 | 0.568 |
| k_{SK5} | 0.425 | 0.538 | 0.668 | 0.415 | 0.229 | 0.339 | 0.554 | 0.498 | 0.111 | 0.199 | 0.460 | 0.631 |
| k_{SK6} | 0.429 | 0.446 | 0.595 | 0.971 | 0.058 | 0.050 | 0.027 | 0.033 | 0.062 | 0.060 | 0.050 | 0.088 |

Table 3: Estimated MSE when $n = 50$ and $p = 8$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|------------------------|--------|-------|--------|---------|--------|-------|-------|--------|--------|-------|-------|--------|
| | ρ | | | | ρ | | | | ρ | | | |
| | 0.9 | 0.95 | 0.99 | 0.999 | 0.9 | 0.95 | 0.99 | 0.999 | 0.9 | 0.95 | 0.99 | 0.999 |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| <i>PMLE</i> | 2.274 | 4.463 | 25.599 | 256.677 | 0.776 | 1.547 | 8.507 | 91.043 | 0.276 | 0.571 | 2.967 | 32.504 |
| <i>k_{HK1}</i> | 0.994 | 1.653 | 8.725 | 80.008 | 0.240 | 0.388 | 1.563 | 15.921 | 0.127 | 0.160 | 0.315 | 2.569 |
| <i>k_{HK2}</i> | 1.728 | 3.152 | 16.606 | 160.467 | 0.613 | 1.138 | 5.612 | 57.365 | 0.253 | 0.491 | 2.145 | 20.659 |
| <i>k_{HKB}</i> | 0.624 | 0.976 | 4.488 | 41.347 | 0.243 | 0.386 | 1.564 | 15.230 | 0.161 | 0.252 | 0.714 | 5.571 |
| <i>k_{LW1}</i> | 2.271 | 4.444 | 25.029 | 225.854 | 0.776 | 1.547 | 8.494 | 89.959 | 0.276 | 0.571 | 2.966 | 32.465 |
| <i>k_{LW2}</i> | 2.049 | 3.814 | 20.384 | 198.289 | 0.731 | 1.409 | 7.105 | 70.567 | 0.270 | 0.549 | 2.637 | 25.968 |
| <i>k_{HSL}</i> | 2.274 | 4.463 | 25.597 | 252.078 | 0.776 | 1.547 | 8.507 | 91.043 | 0.276 | 0.571 | 2.967 | 32.504 |
| <i>k_{AM}</i> | 0.762 | 0.689 | 0.616 | 0.842 | 0.093 | 0.072 | 0.066 | 0.094 | 0.274 | 0.265 | 0.226 | 0.288 |
| <i>k_{GM}</i> | 1.519 | 2.091 | 2.314 | 0.362 | 0.747 | 1.391 | 3.600 | 1.722 | 0.271 | 0.542 | 1.959 | 2.188 |
| <i>k_{KS}</i> | 1.032 | 1.758 | 9.103 | 84.399 | 0.420 | 0.741 | 3.052 | 30.881 | 0.210 | 0.382 | 1.237 | 11.403 |
| <i>k_{A1}</i> | 0.994 | 1.602 | 5.249 | 31.757 | 0.180 | 0.291 | 1.188 | 5.374 | 0.196 | 0.297 | 0.794 | 5.701 |
| <i>k_{A2}</i> | 0.725 | 0.929 | 2.118 | 10.395 | 0.101 | 0.125 | 0.383 | 1.049 | 0.203 | 0.238 | 0.380 | 2.102 |
| <i>k_{A3}</i> | 1.831 | 3.332 | 13.630 | 87.751 | 0.505 | 0.938 | 4.478 | 35.021 | 0.254 | 0.492 | 2.117 | 16.700 |
| <i>k_{MK4}</i> | 1.165 | 1.489 | 1.642 | 0.423 | 0.674 | 1.134 | 2.216 | 1.222 | 0.258 | 0.491 | 1.425 | 1.568 |
| <i>k_{MK5}</i> | 0.454 | 0.569 | 0.948 | 1.613 | 0.185 | 0.251 | 0.517 | 1.021 | 0.152 | 0.223 | 0.433 | 0.992 |
| <i>k_{MK6}</i> | 1.066 | 1.331 | 1.440 | 0.447 | 0.669 | 1.115 | 2.057 | 1.021 | 0.256 | 0.482 | 1.323 | 1.351 |
| <i>k_{GK}</i> | 0.993 | 1.649 | 8.543 | 68.828 | 0.240 | 0.388 | 1.559 | 15.689 | 0.127 | 0.160 | 0.315 | 2.564 |
| <i>k_{D1}</i> | 2.271 | 4.454 | 25.529 | 255.479 | 0.775 | 1.546 | 8.499 | 90.950 | 0.276 | 0.571 | 2.966 | 32.491 |
| <i>k_{D2}</i> | 2.262 | 4.423 | 24.962 | 239.680 | 0.772 | 1.539 | 8.426 | 89.428 | 0.276 | 0.571 | 2.963 | 32.392 |
| <i>k_{D3}</i> | 2.258 | 4.410 | 24.992 | 243.203 | 0.772 | 1.539 | 8.423 | 89.578 | 0.276 | 0.571 | 2.962 | 32.383 |
| <i>k_{D4}</i> | 2.033 | 3.879 | 20.483 | 196.406 | 0.682 | 1.334 | 7.096 | 68.843 | 0.269 | 0.548 | 2.773 | 29.479 |
| <i>k_{Y4}</i> | 0.473 | 0.486 | 0.426 | 0.304 | 0.166 | 0.188 | 0.213 | 0.113 | 0.155 | 0.203 | 0.245 | 0.177 |
| <i>k_{Y6}</i> | 0.900 | 0.874 | 0.841 | 0.830 | 0.033 | 0.028 | 0.022 | 0.020 | 0.303 | 0.297 | 0.281 | 0.271 |
| <i>k_{Y9}</i> | 0.959 | 1.412 | 3.019 | 4.193 | 0.583 | 0.958 | 2.336 | 4.484 | 0.225 | 0.413 | 1.302 | 3.507 |
| <i>k_{AS1}</i> | 1.383 | 2.543 | 14.131 | 130.708 | 0.528 | 0.948 | 4.748 | 49.801 | 0.235 | 0.431 | 1.730 | 17.759 |
| <i>k_{AS2}</i> | 1.086 | 0.997 | 0.835 | 0.955 | 0.181 | 0.141 | 0.079 | 0.046 | 0.474 | 0.461 | 0.391 | 0.207 |
| <i>k_{AS3}</i> | 0.782 | 0.989 | 1.157 | 0.661 | 0.468 | 0.669 | 1.143 | 1.044 | 0.209 | 0.342 | 0.700 | 0.848 |
| <i>k_{AY1}</i> | 2.271 | 4.453 | 25.521 | 255.345 | 0.775 | 1.546 | 8.498 | 90.939 | 0.276 | 0.571 | 2.966 | 32.490 |
| <i>k_{AY2}</i> | 2.196 | 4.254 | 24.053 | 238.648 | 0.761 | 1.509 | 8.201 | 87.268 | 0.275 | 0.567 | 2.920 | 31.666 |
| <i>k_{AY3}</i> | 1.838 | 3.336 | 16.931 | 154.109 | 0.670 | 1.260 | 6.154 | 60.380 | 0.265 | 0.528 | 2.428 | 23.287 |
| <i>k_{AY4}</i> | 2.127 | 4.080 | 22.812 | 224.917 | 0.747 | 1.474 | 7.937 | 84.108 | 0.274 | 0.563 | 2.876 | 30.919 |
| <i>k_{SK1}</i> | 0.383 | 0.424 | 0.468 | 0.370 | 0.159 | 0.196 | 0.288 | 0.223 | 0.135 | 0.185 | 0.285 | 0.385 |
| <i>k_{SK2}</i> | 0.628 | 0.785 | 1.115 | 1.047 | 0.321 | 0.442 | 0.834 | 1.101 | 0.194 | 0.309 | 0.660 | 1.144 |
| <i>k_{SK3}</i> | 0.532 | 0.655 | 0.904 | 0.852 | 0.276 | 0.368 | 0.639 | 0.734 | 0.180 | 0.277 | 0.537 | 0.894 |
| <i>k_{SK4}</i> | 0.585 | 0.728 | 1.051 | 1.045 | 0.316 | 0.430 | 0.770 | 0.989 | 0.191 | 0.301 | 0.617 | 1.067 |
| <i>k_{SK5}</i> | 0.630 | 0.793 | 1.195 | 1.279 | 0.352 | 0.487 | 0.899 | 1.267 | 0.198 | 0.319 | 0.685 | 1.236 |
| <i>k_{SK6}</i> | 0.523 | 0.530 | 0.658 | 0.979 | 0.065 | 0.054 | 0.025 | 0.019 | 0.098 | 0.094 | 0.083 | 0.146 |

Table 4: Estimated MSE when $n = 100$ and $p = 4$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|------------------------|--------|-------|-------|--------|--------|-------|-------|--------|--------|-------|-------|-------|
| | ρ | | | | ρ | | | | ρ | | | |
| | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| <i>PMLE</i> | 0.345 | 0.734 | 3.847 | 43.740 | 0.128 | 0.262 | 1.435 | 15.521 | 0.047 | 0.093 | 0.517 | 5.789 |
| <i>k_{HK1}</i> | 0.255 | 0.406 | 1.200 | 10.800 | 0.077 | 0.113 | 0.255 | 2.008 | 0.071 | 0.069 | 0.116 | 0.329 |
| <i>k_{HK2}</i> | 0.307 | 0.602 | 2.564 | 25.676 | 0.117 | 0.223 | 0.991 | 9.182 | 0.046 | 0.089 | 0.432 | 3.742 |
| <i>k_{HKB}</i> | 0.217 | 0.351 | 1.059 | 9.393 | 0.087 | 0.139 | 0.431 | 3.408 | 0.043 | 0.077 | 0.261 | 1.513 |
| <i>k_{LW1}</i> | 0.345 | 0.733 | 3.831 | 41.579 | 0.128 | 0.262 | 1.435 | 15.459 | 0.047 | 0.093 | 0.517 | 5.787 |
| <i>k_{LW2}</i> | 0.341 | 0.718 | 3.602 | 39.067 | 0.127 | 0.260 | 1.391 | 14.283 | 0.047 | 0.093 | 0.510 | 5.501 |
| <i>k_{HSL}</i> | 0.345 | 0.734 | 3.847 | 43.737 | 0.128 | 0.262 | 1.435 | 15.521 | 0.047 | 0.093 | 0.517 | 5.789 |
| <i>k_{AM}</i> | 0.483 | 0.510 | 0.506 | 0.494 | 0.081 | 0.074 | 0.085 | 0.075 | 0.193 | 0.202 | 0.158 | 0.349 |
| <i>k_{GM}</i> | 0.308 | 0.577 | 1.008 | 0.173 | 0.127 | 0.255 | 1.020 | 1.004 | 0.047 | 0.092 | 0.452 | 0.918 |
| <i>k_{KS}</i> | 0.252 | 0.449 | 1.482 | 12.809 | 0.096 | 0.184 | 0.628 | 4.889 | 0.045 | 0.083 | 0.333 | 2.146 |
| <i>k_{A1}</i> | 0.343 | 0.626 | 2.803 | 25.428 | 0.106 | 0.203 | 0.986 | 6.516 | 0.062 | 0.095 | 0.390 | 4.078 |
| <i>k_{A2}</i> | 0.353 | 0.567 | 2.212 | 15.840 | 0.094 | 0.166 | 0.706 | 2.948 | 0.080 | 0.105 | 0.310 | 3.120 |
| <i>k_{A3}</i> | 0.342 | 0.721 | 3.610 | 33.225 | 0.124 | 0.250 | 1.305 | 12.430 | 0.047 | 0.092 | 0.507 | 5.330 |
| <i>k_{MK4}</i> | 0.287 | 0.508 | 0.808 | 0.190 | 0.124 | 0.242 | 0.807 | 0.681 | 0.047 | 0.091 | 0.403 | 0.753 |
| <i>k_{MK5}</i> | 0.209 | 0.305 | 0.539 | 0.698 | 0.083 | 0.132 | 0.305 | 0.469 | 0.042 | 0.070 | 0.219 | 0.596 |
| <i>k_{MK6}</i> | 0.276 | 0.470 | 0.718 | 0.208 | 0.124 | 0.241 | 0.774 | 0.578 | 0.047 | 0.090 | 0.371 | 0.668 |
| <i>k_{GK}</i> | 0.255 | 0.406 | 1.196 | 10.444 | 0.077 | 0.113 | 0.255 | 1.999 | 0.071 | 0.069 | 0.116 | 0.329 |
| <i>k_{D1}</i> | 0.345 | 0.733 | 3.840 | 43.583 | 0.128 | 0.262 | 1.434 | 15.501 | 0.047 | 0.093 | 0.517 | 5.787 |
| <i>k_{D2}</i> | 0.345 | 0.733 | 3.827 | 42.508 | 0.128 | 0.262 | 1.431 | 15.363 | 0.047 | 0.093 | 0.516 | 5.783 |
| <i>k_{D3}</i> | 0.345 | 0.732 | 3.820 | 42.831 | 0.128 | 0.262 | 1.431 | 15.381 | 0.047 | 0.093 | 0.516 | 5.781 |
| <i>k_{D4}</i> | 0.344 | 0.705 | 3.551 | 39.257 | 0.125 | 0.253 | 1.367 | 14.115 | 0.049 | 0.093 | 0.503 | 5.601 |
| <i>k_{Y4}</i> | 0.253 | 0.325 | 0.364 | 0.204 | 0.091 | 0.142 | 0.268 | 0.179 | 0.045 | 0.074 | 0.190 | 0.263 |
| <i>k_{Y6}</i> | 0.510 | 0.496 | 0.476 | 0.456 | 0.025 | 0.023 | 0.019 | 0.018 | 0.120 | 0.117 | 0.108 | 0.101 |
| <i>k_{Y9}</i> | 0.224 | 0.399 | 0.859 | 1.125 | 0.114 | 0.207 | 0.613 | 1.166 | 0.044 | 0.083 | 0.330 | 0.950 |
| <i>k_{AS1}</i> | 0.271 | 0.506 | 2.071 | 22.417 | 0.113 | 0.205 | 0.811 | 8.067 | 0.046 | 0.088 | 0.378 | 3.095 |
| <i>k_{AS2}</i> | 0.719 | 0.733 | 0.535 | 0.431 | 0.129 | 0.106 | 0.069 | 0.035 | 0.330 | 0.343 | 0.228 | 0.124 |
| <i>k_{AS3}</i> | 0.238 | 0.358 | 0.495 | 0.217 | 0.115 | 0.207 | 0.548 | 0.558 | 0.045 | 0.082 | 0.273 | 0.416 |
| <i>k_{AY1}</i> | 0.345 | 0.733 | 3.839 | 43.565 | 0.128 | 0.262 | 1.434 | 15.498 | 0.047 | 0.093 | 0.517 | 5.786 |
| <i>k_{AY2}</i> | 0.341 | 0.719 | 3.674 | 40.680 | 0.127 | 0.260 | 1.398 | 14.842 | 0.047 | 0.093 | 0.513 | 5.670 |
| <i>k_{AY3}</i> | 0.319 | 0.639 | 2.786 | 27.174 | 0.122 | 0.240 | 1.129 | 10.547 | 0.047 | 0.092 | 0.473 | 4.478 |
| <i>k_{AY4}</i> | 0.337 | 0.706 | 3.522 | 38.244 | 0.127 | 0.257 | 1.364 | 14.264 | 0.047 | 0.093 | 0.510 | 5.559 |
| <i>k_{SK1}</i> | 0.196 | 0.260 | 0.360 | 0.212 | 0.082 | 0.127 | 0.243 | 0.160 | 0.041 | 0.065 | 0.175 | 0.353 |
| <i>k_{SK2}</i> | 0.239 | 0.378 | 0.615 | 0.445 | 0.102 | 0.173 | 0.430 | 0.527 | 0.044 | 0.080 | 0.292 | 0.640 |
| <i>k_{SK3}</i> | 0.223 | 0.335 | 0.537 | 0.400 | 0.098 | 0.163 | 0.375 | 0.393 | 0.043 | 0.076 | 0.249 | 0.559 |
| <i>k_{SK4}</i> | 0.231 | 0.356 | 0.583 | 0.458 | 0.102 | 0.172 | 0.411 | 0.479 | 0.044 | 0.079 | 0.270 | 0.610 |
| <i>k_{SK5}</i> | 0.236 | 0.372 | 0.626 | 0.531 | 0.105 | 0.180 | 0.445 | 0.575 | 0.045 | 0.081 | 0.286 | 0.661 |
| <i>k_{SK6}</i> | 0.248 | 0.251 | 0.310 | 0.624 | 0.036 | 0.035 | 0.019 | 0.015 | 0.035 | 0.035 | 0.028 | 0.036 |

Table 5: Estimated MSE when $n = 100$ and $p = 8$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|------------------------|--------|-------|-------|--------|--------|-------|-------|--------|--------|-------|-------|-------|
| | ρ | | | | ρ | | | | ρ | | | |
| | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| <i>PMLE</i> | 0.469 | 0.933 | 4.964 | 54.532 | 0.171 | 0.342 | 1.858 | 19.833 | 0.063 | 0.127 | 0.679 | 7.111 |
| <i>k_{HK1}</i> | 0.251 | 0.365 | 1.234 | 11.702 | 0.080 | 0.107 | 0.229 | 1.957 | 0.052 | 0.074 | 0.108 | 0.301 |
| <i>k_{HK2}</i> | 0.419 | 0.779 | 3.550 | 35.446 | 0.156 | 0.295 | 1.358 | 12.931 | 0.062 | 0.122 | 0.576 | 4.884 |
| <i>k_{HKB}</i> | 0.243 | 0.363 | 1.151 | 9.776 | 0.097 | 0.148 | 0.460 | 3.674 | 0.054 | 0.093 | 0.280 | 1.487 |
| <i>k_{LW1}</i> | 0.469 | 0.933 | 4.961 | 54.202 | 0.171 | 0.342 | 1.858 | 19.821 | 0.063 | 0.127 | 0.679 | 7.111 |
| <i>k_{LW2}</i> | 0.460 | 0.901 | 4.503 | 46.559 | 0.169 | 0.337 | 1.763 | 17.375 | 0.063 | 0.127 | 0.664 | 6.509 |
| <i>k_{HSL}</i> | 0.469 | 0.933 | 4.964 | 54.532 | 0.171 | 0.342 | 1.858 | 19.833 | 0.063 | 0.127 | 0.679 | 7.111 |
| <i>k_{AM}</i> | 0.596 | 0.549 | 0.490 | 0.572 | 0.076 | 0.057 | 0.045 | 0.063 | 0.212 | 0.234 | 0.219 | 0.171 |
| <i>k_{GM}</i> | 0.445 | 0.827 | 2.473 | 1.318 | 0.171 | 0.339 | 1.630 | 3.597 | 0.063 | 0.127 | 0.649 | 2.819 |
| <i>k_{KS}</i> | 0.325 | 0.533 | 2.057 | 19.401 | 0.134 | 0.232 | 0.793 | 6.954 | 0.060 | 0.112 | 0.401 | 2.861 |
| <i>k_{A1}</i> | 0.400 | 0.645 | 2.190 | 17.721 | 0.082 | 0.148 | 0.635 | 4.492 | 0.073 | 0.115 | 0.330 | 2.521 |
| <i>k_{A2}</i> | 0.399 | 0.512 | 1.147 | 7.921 | 0.058 | 0.084 | 0.278 | 1.440 | 0.101 | 0.129 | 0.217 | 1.131 |
| <i>k_{A3}</i> | 0.457 | 0.886 | 4.277 | 36.985 | 0.149 | 0.287 | 1.409 | 13.859 | 0.063 | 0.124 | 0.595 | 5.845 |
| <i>k_{MK4}</i> | 0.413 | 0.720 | 1.764 | 1.125 | 0.168 | 0.327 | 1.303 | 2.238 | 0.063 | 0.125 | 0.586 | 1.922 |
| <i>k_{MK5}</i> | 0.243 | 0.338 | 0.637 | 1.287 | 0.084 | 0.127 | 0.313 | 0.782 | 0.053 | 0.087 | 0.222 | 0.672 |
| <i>k_{MK6}</i> | 0.402 | 0.684 | 1.564 | 1.023 | 0.168 | 0.327 | 1.291 | 1.926 | 0.063 | 0.124 | 0.582 | 1.660 |
| <i>k_{GK}</i> | 0.251 | 0.365 | 1.233 | 11.653 | 0.080 | 0.107 | 0.229 | 1.955 | 0.052 | 0.074 | 0.108 | 0.301 |
| <i>k_{D1}</i> | 0.469 | 0.933 | 4.962 | 54.500 | 0.171 | 0.342 | 1.858 | 19.829 | 0.063 | 0.127 | 0.679 | 7.111 |
| <i>k_{D2}</i> | 0.469 | 0.932 | 4.954 | 54.226 | 0.171 | 0.342 | 1.855 | 19.785 | 0.063 | 0.127 | 0.679 | 7.109 |
| <i>k_{D3}</i> | 0.469 | 0.932 | 4.952 | 54.239 | 0.171 | 0.342 | 1.856 | 19.783 | 0.063 | 0.127 | 0.679 | 7.108 |
| <i>k_{D4}</i> | 0.460 | 0.893 | 4.620 | 48.659 | 0.164 | 0.328 | 1.750 | 18.226 | 0.064 | 0.127 | 0.660 | 6.836 |
| <i>k_{Y4}</i> | 0.274 | 0.323 | 0.346 | 0.222 | 0.090 | 0.130 | 0.210 | 0.184 | 0.056 | 0.091 | 0.180 | 0.224 |
| <i>k_{Y6}</i> | 0.737 | 0.715 | 0.674 | 0.659 | 0.021 | 0.018 | 0.014 | 0.012 | 0.242 | 0.235 | 0.218 | 0.209 |
| <i>k_{Y9}</i> | 0.338 | 0.594 | 1.792 | 4.279 | 0.159 | 0.296 | 1.100 | 3.363 | 0.059 | 0.115 | 0.506 | 2.246 |
| <i>k_{AS1}</i> | 0.349 | 0.617 | 2.851 | 30.387 | 0.151 | 0.272 | 1.127 | 11.037 | 0.061 | 0.118 | 0.498 | 3.991 |
| <i>k_{AS2}</i> | 0.940 | 0.888 | 0.783 | 0.483 | 0.162 | 0.120 | 0.070 | 0.031 | 0.408 | 0.435 | 0.411 | 0.228 |
| <i>k_{AS3}</i> | 0.352 | 0.551 | 0.990 | 0.873 | 0.152 | 0.268 | 0.768 | 1.328 | 0.060 | 0.112 | 0.389 | 0.934 |
| <i>k_{AY1}</i> | 0.469 | 0.933 | 4.962 | 54.496 | 0.171 | 0.342 | 1.857 | 19.828 | 0.063 | 0.127 | 0.679 | 7.111 |
| <i>k_{AY2}</i> | 0.465 | 0.922 | 4.845 | 52.727 | 0.170 | 0.340 | 1.833 | 19.433 | 0.063 | 0.127 | 0.677 | 7.037 |
| <i>k_{AY3}</i> | 0.439 | 0.834 | 3.873 | 37.575 | 0.164 | 0.318 | 1.543 | 14.677 | 0.063 | 0.125 | 0.630 | 5.744 |
| <i>k_{AY4}</i> | 0.462 | 0.911 | 4.737 | 51.132 | 0.170 | 0.338 | 1.810 | 19.065 | 0.063 | 0.127 | 0.674 | 6.966 |
| <i>k_{SK1}</i> | 0.218 | 0.276 | 0.391 | 0.416 | 0.083 | 0.122 | 0.241 | 0.315 | 0.050 | 0.080 | 0.178 | 0.357 |
| <i>k_{SK2}</i> | 0.306 | 0.459 | 0.925 | 1.206 | 0.125 | 0.203 | 0.522 | 1.127 | 0.058 | 0.105 | 0.327 | 0.991 |
| <i>k_{SK3}</i> | 0.280 | 0.403 | 0.742 | 0.982 | 0.116 | 0.185 | 0.446 | 0.818 | 0.057 | 0.100 | 0.289 | 0.764 |
| <i>k_{SK4}</i> | 0.295 | 0.436 | 0.845 | 1.171 | 0.125 | 0.203 | 0.512 | 1.007 | 0.058 | 0.104 | 0.323 | 0.894 |
| <i>k_{SK5}</i> | 0.306 | 0.459 | 0.931 | 1.367 | 0.131 | 0.218 | 0.571 | 1.196 | 0.059 | 0.107 | 0.351 | 1.003 |
| <i>k_{SK6}</i> | 0.290 | 0.287 | 0.313 | 0.592 | 0.039 | 0.039 | 0.023 | 0.008 | 0.051 | 0.052 | 0.044 | 0.050 |

Table 6: Estimated MSE when $n = 200$ and $p = 4$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|------------------------|--------|-------|-------|--------|--------|-------|-------|-------|--------|-------|-------|-------|
| | ρ | | | | ρ | | | | ρ | | | |
| | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| <i>PMLE</i> | 0.141 | 0.283 | 1.517 | 16.440 | 0.050 | 0.100 | 0.539 | 6.087 | 0.018 | 0.037 | 0.199 | 2.229 |
| <i>k_{HK1}</i> | 0.131 | 0.208 | 0.544 | 3.581 | 0.041 | 0.064 | 0.119 | 0.645 | 0.031 | 0.037 | 0.094 | 0.102 |
| <i>k_{HK2}</i> | 0.134 | 0.256 | 1.130 | 9.886 | 0.048 | 0.093 | 0.416 | 3.722 | 0.018 | 0.036 | 0.183 | 1.576 |
| <i>k_{HKB}</i> | 0.112 | 0.185 | 0.543 | 3.664 | 0.042 | 0.072 | 0.212 | 1.410 | 0.018 | 0.034 | 0.138 | 0.704 |
| <i>k_{LW1}</i> | 0.141 | 0.283 | 1.516 | 16.337 | 0.050 | 0.100 | 0.539 | 6.084 | 0.018 | 0.037 | 0.199 | 2.228 |
| <i>k_{LW2}</i> | 0.141 | 0.281 | 1.485 | 15.526 | 0.049 | 0.100 | 0.535 | 5.883 | 0.018 | 0.037 | 0.199 | 2.192 |
| <i>k_{HSL}</i> | 0.141 | 0.283 | 1.517 | 16.440 | 0.050 | 0.100 | 0.539 | 6.087 | 0.018 | 0.037 | 0.199 | 2.229 |
| <i>k_{AM}</i> | 0.382 | 0.455 | 0.379 | 0.531 | 0.060 | 0.058 | 0.059 | 0.070 | 0.162 | 0.187 | 0.133 | 0.225 |
| <i>k_{GM}</i> | 0.137 | 0.266 | 0.889 | 0.434 | 0.049 | 0.100 | 0.495 | 1.436 | 0.018 | 0.037 | 0.193 | 1.002 |
| <i>k_{KS}</i> | 0.124 | 0.224 | 0.767 | 5.030 | 0.044 | 0.086 | 0.306 | 2.044 | 0.018 | 0.036 | 0.151 | 1.020 |
| <i>k_{A1}</i> | 0.155 | 0.282 | 1.259 | 12.795 | 0.048 | 0.087 | 0.436 | 4.277 | 0.029 | 0.045 | 0.174 | 1.805 |
| <i>k_{A2}</i> | 0.175 | 0.288 | 1.067 | 10.368 | 0.047 | 0.079 | 0.359 | 2.828 | 0.042 | 0.058 | 0.154 | 1.489 |
| <i>k_{A3}</i> | 0.141 | 0.281 | 1.494 | 15.215 | 0.049 | 0.099 | 0.520 | 5.644 | 0.018 | 0.037 | 0.198 | 2.184 |
| <i>k_{MK4}</i> | 0.133 | 0.250 | 0.732 | 0.408 | 0.049 | 0.098 | 0.442 | 0.988 | 0.018 | 0.037 | 0.185 | 0.805 |
| <i>k_{MK5}</i> | 0.111 | 0.172 | 0.400 | 0.693 | 0.040 | 0.069 | 0.191 | 0.448 | 0.017 | 0.032 | 0.122 | 0.456 |
| <i>k_{MK6}</i> | 0.131 | 0.241 | 0.651 | 0.395 | 0.049 | 0.098 | 0.440 | 0.853 | 0.018 | 0.037 | 0.179 | 0.707 |
| <i>k_{GK}</i> | 0.131 | 0.208 | 0.544 | 3.567 | 0.041 | 0.064 | 0.119 | 0.645 | 0.031 | 0.037 | 0.094 | 0.102 |
| <i>k_{D1}</i> | 0.141 | 0.282 | 1.516 | 16.420 | 0.050 | 0.100 | 0.539 | 6.084 | 0.018 | 0.037 | 0.199 | 2.228 |
| <i>k_{D2}</i> | 0.141 | 0.282 | 1.515 | 16.353 | 0.050 | 0.100 | 0.539 | 6.074 | 0.018 | 0.037 | 0.199 | 2.228 |
| <i>k_{D3}</i> | 0.141 | 0.282 | 1.515 | 16.359 | 0.050 | 0.100 | 0.539 | 6.075 | 0.018 | 0.037 | 0.199 | 2.228 |
| <i>k_{D4}</i> | 0.144 | 0.284 | 1.457 | 15.603 | 0.049 | 0.099 | 0.525 | 5.886 | 0.019 | 0.037 | 0.198 | 2.189 |
| <i>k_{Y4}</i> | 0.131 | 0.198 | 0.310 | 0.218 | 0.043 | 0.074 | 0.191 | 0.255 | 0.019 | 0.034 | 0.117 | 0.276 |
| <i>k_{Y6}</i> | 0.364 | 0.355 | 0.334 | 0.318 | 0.016 | 0.015 | 0.012 | 0.011 | 0.080 | 0.078 | 0.071 | 0.067 |
| <i>k_{Y9}</i> | 0.112 | 0.208 | 0.609 | 1.079 | 0.047 | 0.091 | 0.348 | 0.993 | 0.018 | 0.035 | 0.162 | 0.716 |
| <i>k_{AS1}</i> | 0.127 | 0.235 | 0.904 | 8.471 | 0.047 | 0.091 | 0.357 | 3.146 | 0.018 | 0.036 | 0.173 | 1.251 |
| <i>k_{AS2}</i> | 0.619 | 0.702 | 0.500 | 0.205 | 0.104 | 0.095 | 0.063 | 0.028 | 0.294 | 0.331 | 0.222 | 0.143 |
| <i>k_{AS3}</i> | 0.122 | 0.204 | 0.441 | 0.307 | 0.048 | 0.092 | 0.336 | 0.683 | 0.018 | 0.035 | 0.152 | 0.446 |
| <i>k_{AY1}</i> | 0.141 | 0.282 | 1.516 | 16.418 | 0.050 | 0.100 | 0.539 | 6.084 | 0.018 | 0.037 | 0.199 | 2.228 |
| <i>k_{AY2}</i> | 0.141 | 0.281 | 1.487 | 15.756 | 0.049 | 0.100 | 0.534 | 5.936 | 0.018 | 0.037 | 0.199 | 2.208 |
| <i>k_{AY3}</i> | 0.137 | 0.267 | 1.254 | 11.229 | 0.049 | 0.097 | 0.472 | 4.508 | 0.018 | 0.037 | 0.193 | 1.896 |
| <i>k_{AY4}</i> | 0.140 | 0.279 | 1.459 | 15.159 | 0.049 | 0.100 | 0.528 | 5.798 | 0.018 | 0.037 | 0.198 | 2.188 |
| <i>k_{SK1}</i> | 0.107 | 0.155 | 0.296 | 0.296 | 0.040 | 0.068 | 0.175 | 0.223 | 0.017 | 0.031 | 0.105 | 0.318 |
| <i>k_{SK2}</i> | 0.121 | 0.207 | 0.509 | 0.561 | 0.045 | 0.083 | 0.263 | 0.593 | 0.018 | 0.035 | 0.152 | 0.565 |
| <i>k_{SK3}</i> | 0.116 | 0.189 | 0.436 | 0.513 | 0.044 | 0.080 | 0.244 | 0.463 | 0.018 | 0.034 | 0.137 | 0.484 |
| <i>k_{SK4}</i> | 0.118 | 0.198 | 0.472 | 0.568 | 0.045 | 0.083 | 0.262 | 0.533 | 0.018 | 0.035 | 0.145 | 0.526 |
| <i>k_{SK5}</i> | 0.120 | 0.204 | 0.502 | 0.631 | 0.046 | 0.085 | 0.278 | 0.606 | 0.018 | 0.035 | 0.151 | 0.562 |
| <i>k_{SK6}</i> | 0.143 | 0.142 | 0.153 | 0.325 | 0.021 | 0.023 | 0.015 | 0.007 | 0.019 | 0.020 | 0.017 | 0.015 |

Table 7: Estimated MSE when $n = 200$ and $p = 8$

| β_0 | -1 | | | | 0 | | | | 1 | | | |
|------------------------|--------|-------|-------|--------|--------|-------|-------|-------|--------|-------|-------|-------|
| | ρ | | | | ρ | | | | ρ | | | |
| | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 | 0.900 | 0.950 | 0.990 | 0.999 |
| Estimator | MSE | | | | MSE | | | | MSE | | | |
| <i>PMLE</i> | 0.134 | 0.264 | 1.411 | 15.475 | 0.049 | 0.099 | 0.520 | 5.705 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{HK1}</i> | 0.099 | 0.149 | 0.345 | 2.321 | 0.032 | 0.042 | 0.062 | 0.378 | 0.024 | 0.047 | 0.091 | 0.070 |
| <i>k_{HK2}</i> | 0.129 | 0.245 | 1.128 | 10.513 | 0.048 | 0.094 | 0.429 | 3.916 | 0.018 | 0.036 | 0.181 | 1.587 |
| <i>k_{HKB}</i> | 0.102 | 0.162 | 0.466 | 3.151 | 0.039 | 0.066 | 0.192 | 1.182 | 0.018 | 0.033 | 0.124 | 0.598 |
| <i>k_{LW1}</i> | 0.134 | 0.264 | 1.411 | 15.469 | 0.049 | 0.099 | 0.520 | 5.704 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{LW2}</i> | 0.134 | 0.262 | 1.374 | 14.292 | 0.049 | 0.099 | 0.514 | 5.446 | 0.018 | 0.036 | 0.192 | 2.022 |
| <i>k_{HSL}</i> | 0.134 | 0.264 | 1.411 | 15.475 | 0.049 | 0.099 | 0.520 | 5.705 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{AM}</i> | 0.449 | 0.442 | 0.437 | 0.371 | 0.059 | 0.046 | 0.029 | 0.045 | 0.158 | 0.183 | 0.200 | 0.125 |
| <i>k_{GM}</i> | 0.133 | 0.258 | 1.226 | 2.731 | 0.049 | 0.099 | 0.511 | 3.235 | 0.018 | 0.036 | 0.192 | 1.667 |
| <i>k_{KS}</i> | 0.119 | 0.203 | 0.715 | 5.778 | 0.044 | 0.075 | 0.257 | 2.313 | 0.018 | 0.035 | 0.141 | 1.083 |
| <i>k_{A1}</i> | 0.152 | 0.244 | 0.869 | 8.038 | 0.036 | 0.064 | 0.284 | 2.661 | 0.029 | 0.044 | 0.139 | 1.072 |
| <i>k_{A2}</i> | 0.192 | 0.251 | 0.576 | 4.339 | 0.033 | 0.046 | 0.160 | 1.282 | 0.049 | 0.063 | 0.120 | 0.582 |
| <i>k_{A3}</i> | 0.134 | 0.261 | 1.329 | 13.883 | 0.047 | 0.094 | 0.468 | 4.855 | 0.018 | 0.036 | 0.184 | 1.906 |
| <i>k_{MK4}</i> | 0.131 | 0.248 | 1.018 | 1.960 | 0.049 | 0.098 | 0.481 | 2.160 | 0.018 | 0.036 | 0.187 | 1.304 |
| <i>k_{MK5}</i> | 0.104 | 0.163 | 0.366 | 1.019 | 0.035 | 0.058 | 0.166 | 0.551 | 0.017 | 0.031 | 0.103 | 0.411 |
| <i>k_{MK6}</i> | 0.130 | 0.245 | 0.983 | 1.658 | 0.049 | 0.098 | 0.482 | 2.057 | 0.018 | 0.036 | 0.188 | 1.237 |
| <i>k_{GK}</i> | 0.099 | 0.149 | 0.345 | 2.320 | 0.032 | 0.042 | 0.062 | 0.378 | 0.024 | 0.047 | 0.091 | 0.070 |
| <i>k_{D1}</i> | 0.134 | 0.264 | 1.410 | 15.473 | 0.049 | 0.099 | 0.520 | 5.704 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{D2}</i> | 0.134 | 0.264 | 1.410 | 15.463 | 0.049 | 0.099 | 0.520 | 5.702 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{D3}</i> | 0.134 | 0.264 | 1.410 | 15.461 | 0.049 | 0.099 | 0.520 | 5.702 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{D4}</i> | 0.137 | 0.261 | 1.363 | 14.794 | 0.049 | 0.097 | 0.509 | 5.544 | 0.018 | 0.036 | 0.191 | 2.035 |
| <i>k_{Y4}</i> | 0.122 | 0.173 | 0.269 | 0.235 | 0.039 | 0.064 | 0.152 | 0.232 | 0.018 | 0.033 | 0.102 | 0.226 |
| <i>k_{Y6}</i> | 0.580 | 0.559 | 0.517 | 0.500 | 0.014 | 0.012 | 0.009 | 0.007 | 0.176 | 0.172 | 0.155 | 0.144 |
| <i>k_{Y9}</i> | 0.116 | 0.217 | 0.881 | 3.158 | 0.048 | 0.095 | 0.423 | 2.092 | 0.018 | 0.035 | 0.174 | 1.168 |
| <i>k_{AS1}</i> | 0.120 | 0.216 | 0.911 | 8.749 | 0.048 | 0.092 | 0.385 | 3.260 | 0.018 | 0.035 | 0.172 | 1.281 |
| <i>k_{AS2}</i> | 0.792 | 0.785 | 0.764 | 0.405 | 0.131 | 0.101 | 0.056 | 0.026 | 0.332 | 0.371 | 0.397 | 0.232 |
| <i>k_{AS3}</i> | 0.125 | 0.224 | 0.630 | 1.100 | 0.048 | 0.092 | 0.365 | 1.180 | 0.018 | 0.035 | 0.159 | 0.706 |
| <i>k_{AY1}</i> | 0.134 | 0.264 | 1.410 | 15.473 | 0.049 | 0.099 | 0.520 | 5.704 | 0.018 | 0.036 | 0.193 | 2.071 |
| <i>k_{AY2}</i> | 0.134 | 0.263 | 1.399 | 15.218 | 0.049 | 0.099 | 0.517 | 5.649 | 0.018 | 0.036 | 0.193 | 2.063 |
| <i>k_{AY3}</i> | 0.132 | 0.255 | 1.241 | 11.830 | 0.049 | 0.097 | 0.477 | 4.616 | 0.018 | 0.036 | 0.189 | 1.844 |
| <i>k_{AY4}</i> | 0.134 | 0.262 | 1.388 | 14.976 | 0.049 | 0.099 | 0.515 | 5.595 | 0.018 | 0.036 | 0.193 | 2.055 |
| <i>k_{SK1}</i> | 0.100 | 0.145 | 0.274 | 0.454 | 0.036 | 0.059 | 0.155 | 0.342 | 0.017 | 0.030 | 0.093 | 0.281 |
| <i>k_{SK2}</i> | 0.117 | 0.199 | 0.537 | 1.290 | 0.044 | 0.080 | 0.265 | 0.877 | 0.018 | 0.034 | 0.138 | 0.628 |
| <i>k_{SK3}</i> | 0.112 | 0.186 | 0.461 | 1.013 | 0.043 | 0.076 | 0.242 | 0.711 | 0.018 | 0.033 | 0.129 | 0.521 |
| <i>k_{SK4}</i> | 0.115 | 0.195 | 0.515 | 1.180 | 0.044 | 0.080 | 0.266 | 0.825 | 0.018 | 0.034 | 0.139 | 0.595 |
| <i>k_{SK5}</i> | 0.117 | 0.200 | 0.559 | 1.328 | 0.045 | 0.084 | 0.288 | 0.933 | 0.018 | 0.034 | 0.147 | 0.656 |
| <i>k_{SK6}</i> | 0.159 | 0.159 | 0.149 | 0.269 | 0.020 | 0.023 | 0.019 | 0.004 | 0.025 | 0.027 | 0.026 | 0.018 |

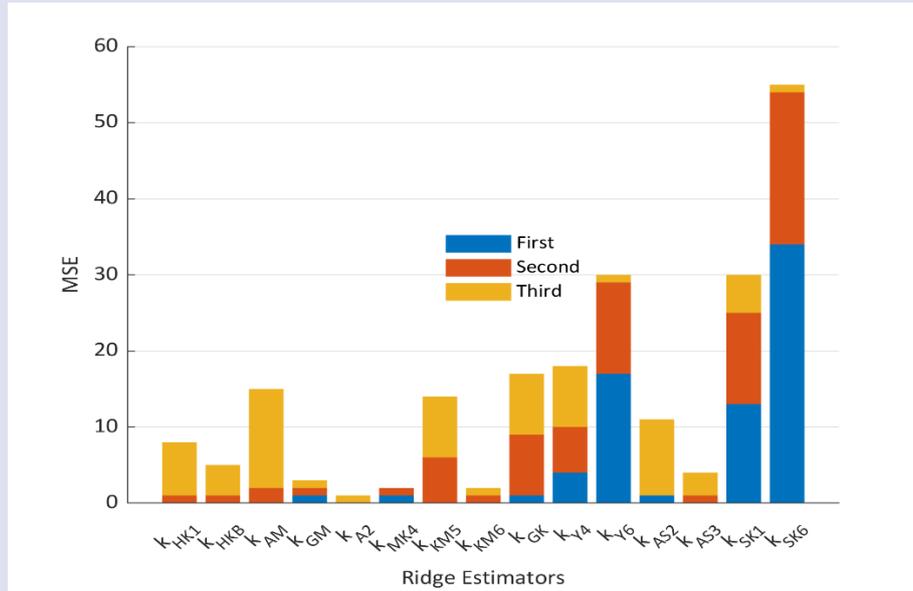


Figure 1: Estimators Achieving Top Three Positions in MSE Performance

Table 8: Poisson Regression Estimates for the aircraft damage dataset

| Estimator | β_0 | β_1 | β_2 | β_3 | MSE |
|------------------|-----------|-----------|-----------|-----------|-------|
| PMLE | -0.406 | 0.569 | 0.165 | -0.014 | 1.029 |
| k _{HK1} | -0.140 | 0.339 | 0.172 | -0.015 | 0.254 |
| k _{HK2} | -0.193 | 0.414 | 0.169 | -0.015 | 0.304 |
| k _{HKB} | -0.078 | 0.225 | 0.178 | -0.016 | 0.272 |
| k _{LW1} | -0.401 | 0.567 | 0.165 | -0.014 | 1.003 |
| k _{LW2} | -0.400 | 0.567 | 0.165 | -0.014 | 0.999 |
| k _{HSL} | -0.406 | 0.569 | 0.165 | -0.014 | 1.029 |
| k _{AM} | 0.001 | 0.003 | 0.035 | 0.004 | 0.503 |
| k _{GM} | -0.402 | 0.567 | 0.165 | -0.014 | 1.006 |
| k _{KS} | -0.168 | 0.380 | 0.171 | -0.015 | 0.273 |
| k _{A1} | 0.000 | 0.030 | 0.163 | -0.013 | 0.456 |
| k _{A2} | 0.003 | 0.013 | 0.112 | -0.006 | 0.479 |
| k _{A3} | -0.282 | 0.501 | 0.166 | -0.014 | 0.508 |
| k _{MK4} | -0.375 | 0.556 | 0.165 | -0.014 | 0.874 |
| k _{MK5} | -0.079 | 0.227 | 0.178 | -0.016 | 0.271 |
| k _{MK6} | -0.353 | 0.545 | 0.165 | -0.014 | 0.772 |
| k _{GK} | -0.140 | 0.339 | 0.172 | -0.015 | 0.254 |
| k _{D1} | -0.406 | 0.569 | 0.165 | -0.014 | 1.029 |
| k _{D2} | -0.406 | 0.569 | 0.165 | -0.014 | 1.029 |
| k _{D3} | -0.406 | 0.569 | 0.165 | -0.014 | 1.028 |
| k _{D4} | -0.399 | 0.566 | 0.165 | -0.014 | 0.994 |
| k _{Y4} | -0.151 | 0.355 | 0.172 | -0.015 | 0.259 |
| k _{Y6} | -0.103 | 0.273 | 0.176 | -0.016 | 0.254 |
| k _{Y9} | -0.261 | 0.483 | 0.166 | -0.015 | 0.446 |
| k _{AS1} | -0.168 | 0.380 | 0.171 | -0.015 | 0.273 |
| k _{AS2} | 0.000 | 0.001 | 0.011 | 0.006 | 0.512 |
| k _{AS3} | -0.333 | 0.534 | 0.165 | -0.014 | 0.687 |
| k _{AY1} | -0.406 | 0.569 | 0.165 | -0.014 | 1.029 |
| k _{AY2} | -0.402 | 0.567 | 0.165 | -0.014 | 1.009 |
| k _{AY3} | -0.311 | 0.521 | 0.165 | -0.014 | 0.605 |
| k _{AY4} | -0.399 | 0.566 | 0.165 | -0.014 | 0.990 |
| k _{SK1} | -0.058 | 0.180 | 0.180 | -0.016 | 0.300 |
| k _{SK2} | -0.187 | 0.406 | 0.169 | -0.015 | 0.296 |
| k _{SK3} | -0.125 | 0.314 | 0.174 | -0.016 | 0.250 |
| k _{SK4} | -0.160 | 0.369 | 0.171 | -0.015 | 0.266 |
| k _{SK5} | -0.184 | 0.401 | 0.170 | -0.015 | 0.291 |
| k _{SK6} | -0.003 | 0.039 | 0.171 | -0.014 | 0.445 |

Conclusion

This study evaluated the performance of several existing ridge estimators and introduced new ridge estimators which effectively address multicollinearity in Poisson regression models. Through an extensive simulation study, it was found that the proposed estimator, k_{SK6} , consistently outperformed all other estimators in terms of minimizing the MSE. Furthermore, estimators k_{SK1} and k_{Y6} ranked consistently as the second-best performers based on MSE values. When applied to a real dataset, the proposed estimator k_{SK1} demonstrated the lowest MSE, indicating its superior performance in practical applications. In contrast, the Poisson maximum likelihood estimator emerged as the least effective, with the highest MSE, underscoring its limitations in dealing with multicollinearity.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Agresti, *An Introduction To Categorical Data Analysis*, Third Edition. Wiley, 2019. [Online]. Available: <http://www.wiley.com/go/wsp>
- [2] Belsey D. A., Kuh E., Welsch R. E., *Regression Diagnostics*. Wiley, 2005.
- [3] Menard S., *Applied Logistic Regression Analysis*, 3rd ed. Sage, (2001).
- [4] Hoerl E., Kennard R. W., Ridge Regression: Biased Estimation for Nonorthogonal Problems, *Technometrics*, 12(1) (1970) 55–67.
- [5] Tibshirani R., Regression Shrinkage and Selection via the Lasso, *J. R. Statist. Soc. B*, 58(1) (1996) 267–288.
- [6] Kibria M. G., Performance of some New Ridge regression estimators, *Communications in Statistics Part B: Simulation and Computation*, 32(2) (2003) 419–435.
- [7] Shukur G., Choosing ridge parameter for regression problems, *Commun Stat Theory Methods*, 34(5) 1177–1182
- [8] Alkhamisi M., Khalaf G., Shukur G., Some modifications for choosing ridge parameters, *Commun. Stat. Theory Methods*, 35(11) (2006) 2005–2020.
- [9] Muniz G., Kibria B. M. G., On some ridge regression estimators: An empirical comparisons, *Commun. Stat. Simul. Comput.*, 38(3) (2009) 621–630
- [10] Kibria M. G., Månsson K., Shukur G., Performance of Some Logistic Ridge Regression Estimators, *Comput. Econ.*, 40(4) (2012) 401–414.
- [11] Asar Y., Genç A., A New Two-Parameter Estimator for the Poisson Regression Model, *Iran J. Sci. Technol. Trans. A Sci.*, 42(2) (2018) 793–803.
- [12] Lukman F., Arashi M., Prokaj V., Robust biased estimators for Poisson regression model: Simulation and applications, *Concurr. Comput.*, 35(7) (2023) 7594.
- [13] Lukman F., Adewuyi E., Månsson K., Kibria B. M. G., A new estimator for the multicollinear Poisson regression model: simulation and application, *Sci. Rep.*, 11(1) (2021) s41598-021-82582-w.
- [14] Rashad N. K., Algamal Z. Y., A New Ridge Estimator for the Poisson Regression Model, *Iran J. Sci. Technol. Trans. A Sci.*, 43(6) (2019) 2921–2928.
- [15] Myers R. H., Montgomery D. C., Vining G. G., Robinson T. J., *Generalized Linear Models with Application in Engineering and the Sciences*, Second edition. Wiley, 2012.
- [16] K. Månsson and G. Shukur, A Poisson ridge regression estimator, *Econ Model*, 28(4) (2011) 1475–1481.
- [17] Asar Y., Genç A., A note on some new modifications of ridge estimators, *Kuwait J. Sci*, 44(3) (2017) 75–82.
- [18] Khalaf G., A proposed ridge parameter to improve the least squares estimator, *Journal of Modern Applied Statistical Methods*, 11(2) (2012) 443–449.
- [19] Dorugade V., New ridge parameters for ridge regression, (2014) University of Bahrain.
- [20] Lawless J.F., Wang P., A Simulation Study Of Ridge And Other Regression Estimators, *Commun. Stat. Theory Methods*, 5(4) (1976) 307–323.
- [21] Hocking R. R., Speed F. M., Lynn M. J., A Class of Biased Estimators in Linear Regression, *Technometrics*, 18(4) (1976) 425–437.
- [22] V. Dorugade, On Comparison of Some Ridge Parameters in Ridge Regression, *Sri Lankan J. Appl. Stat.*, 15(1) (2014) 31–46.
- [23] Asar Y., Karaibrahimoğlu A., Genç A., Modified ridge regression parameters: A comparative Monte Carlo study, *Hacettepe Journal of Mathematics and Statistics*, 43(5) (2014) 827–841.

Monte Carlo Simulation Forecasting the Prices of Selected Stocks in the Automotive Sector

Mehmet Siddik Çadırcı^{1,a,*}

¹ Department of Statistics and Computer Sciences, Faculty of Science, Cumhuriyet University, 58140, Sivas, Türkiye

*Corresponding author

Research Article

History

Received: 01/05/2024

Accepted: 16/12/2024



This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

ABSTRACT

Investors are exposed to risk and uncertainty because of changes in financial markets' prices. Investors perceive the risks associated with changes in the market prices as higher due to inaccuracy in predicting future returns because of fluctuations in prices. For this reason, they adopt different risk management methods that reduce or eliminate these risks. This research relies on Monte Carlo Simulation technique in predicting forthcoming yield rates from three companies operating under Turkish automotive segment namely, Dogus Automotive (DOAS), Tofas (TOASO) and Ford Otosan (FROTO). The simulation, which runs from January 1, 2023, to December 31, 2023, gives investors research-based insights that help them make strategic investment choices in times of high volatility in the market. According to the results, by modeling prospective future scenarios, MCS can be employed as a viable means of predicting stock prices in financial markets which subsequently helps people make rational investments thereby securing profitable ventures. Furthermore, this study offers practical suggestions in the form of MCS-generated volatility ranges. Investors can determine when it is advisable to buy or sell stocks in order to reduce potential losses and increase profits by setting realistic price objectives and allocating the portfolio differently in accordance with these calls.

Keywords: Risk Management, Financial Markets, Automotive Industry, Stock Forecasting, Monte Carlo Simulation.

msiddikcadirci@cumhuriyet.edu.tr <https://orcid.org/0000-0001-7654-7609>

Introduction

Monte Carlo Simulation (MCS) is an eminently practical computational technique that embraces probabilistic modelling as part of quantifying the risks or uncertainties regarding wide-field analysis, including engineering, physics, and finance. It generates several random scenarios to forecast outcomes and quantify risk, hence becoming a key technique in risk assessment and decision-making. The automotive industry is one of the most volatile, meaning that its stock prices fluctuate dynamically. Therefore, it is also a very good area where MCS can be applied to perform investment planning and risk management.

MCS has been deployed in analyzes regarding stock price trends, sales forecasting, and optimization of investment strategies in the automotive field. This paper uses the Monte Carlo method to simulate the stock price movements regarding the selection of three key automotive companies listed in Borsa Istanbul: Dogus Automotive, Tofas, and Ford Otosan. For this, the work will give a model-based approach using historical stock price data from January 1, 2023, to December 31, 2023, so that the investor could hedge the volatility of the Turkish automotive sector. First, the following sections briefly outline the applications of MCS in risk management and automotive analysis. Following this will be the described methodology in detail, including the algorithm that has been undertaken for the Monte Carlo Simulations. Then it will present findings from the data simulations. This structure is necessary to ensure that investors and

researchers alike recognize the practical and theoretical contributions of this study.

Literature Review

The Monte Carlo Simulation is a strong computer technique that has seen its application in engineering, physics, and finance, among other areas. It constitutes one of the major tools for risk assessment and decision-making by the generation of diverse events, together with their probabilities. This section reviews pertinent literature concerning the application of MCS in financial risk management and the automotive sector, highlighting its versatility and relevance.

Monte Carlo Simulation in Risk Management

Monte Carlo Simulation has also shown versatility in the different industries in which risk management is necessary. MCS allows analysts to gauge possible outcomes of alternative scenarios, thus reducing uncertainty of outcome that one is trying to reach. In service quality management, [1] use MCS to delve deep into how risk influences service quality gaps. The findings place greater emphasis on the role of MCS in service management with respect to effecting plans and managing uncertainties, among other roles. A nested MCS approach with regression in respect to financial risk estimation is proposed in [2]. Such a model was applied in catering to financial risks and further expanded the scope of contribution by MCS in the management of critical

financial circumstances [3]. It has also presented such research that explains how MCS can enhance financial forecasting in modelling various risk factors, thereby enhancing overall accuracy and reliability in the prediction.

Application of Monte Carlo Simulation in the Automotive Sector

The automotive industry is fast-moving, sensitive to changes in the economy and market, and is increasingly integrating MCS to perform forecasting and optimization of operations. MCS has been estimated in the analysis of stock, filtering later accounting records, sales forecasting, and investment planning in the automotive industry. For instance, [4] combined MCS with DEA in support of a decision-making process about the selection of a green car. Their work involved how MCS could be extended to include multiple variables so that comprehensive risk and scenario analysis would be carried out, catering to particular needs of the automotive industry. The application of MCS in the automotive sector for sales forecasting was also presented by [5]. Their research showed that MCS has the capability of modeling periodic sales patterns, which would eventually help investors and automotive companies have a timely insight into making necessary decisions. These studies together give an example of the flexibility of MCS in solving multi-dimensional complex problems in the automotive industry. In computational finance, MCS has also been utilized for various applications, such as stock price analysis and risk evaluation within the automotive sector [6]. MCS's role in this domain emphasizes its capacity to provide financial insights by modelling stock market behaviour under different economic conditions. This capability is particularly relevant to the automotive sector, which must navigate fluctuating market dynamics and investor expectations. As such, most literature supports the application of MCS to stock market forecasting using its apparent capability for modeling complex financial phenomena. For instance, [7], while using an MCS approach, simulated the movement of stock prices; thus, this technique was applied to the uncertain nature of daily fluctuations in stock prices. Running several simulations based on historical data, the study by [7] provided valuable insights into possible future price ranges, which are crucial for both short-term and long-term investors. This study underscored the flexibility of MCS in generating realistic price predictions, especially for volatile markets where conventional models fall short.

Broader Applications of Monte Carlo Simulation

The effectiveness of Monte Carlo Simulation is not confined to either financial or automotive sectors. As an instance, research works have adopted MCS in the biomedical engineering field, such as the modelling of photon migration in tissues see in [8]. This, therefore, points out how flexible and precise MCS can become when applied within various contexts that demand probabilistic modelling-from design safety in vehicles to manufacturing optimization. Applications of MCS are so

varied that it confirms MCS's status as a basic instrument to solve problems that are complex and present elements of uncertainty and variability.

Novelty and Contributions

Although Monte Carlo Simulation is already well-established in many areas, such as finance and the automotive sector, for stock price forecasting in the automotive sector within Turkey, it has not been well explored. While past research, [9], has applied MCS in broader financial risk assessments and portfolio management contexts. The current paper tries to fill up this research gap by focusing on the application of MCS on stock prices of the companies Dogus Automotive: DOAS, Tofas: TOASO, and Ford Otosan: FROTO. It goes on to provide a customized model, taking into consideration unique market and economic conditions which impact the Turkish automotive industry, an extremely important emerging market that holds serious influence these days. Although new in regional focus, this study underlines the practical implications of MCS for investors in the Turkish automotive sector. Using a broad dataset and controlling the localized economic variables, it provides a framework that could be replicated for other emerging markets across similar dynamics. The methodology followed in this paper bridges the gap between theoretical modeling and real practice needs of investors; hence, it is a valuable tool deeply needed for risk assessment and investment planning. Though the Monte Carlo Simulation has proven helpful in financial forecasting and automotive analysis around the world, the application of this study to Turkish automotive stocks ushers in an entirely new dimension and enhances the understanding of localized market dynamics with much insight and practical applications that go long in contributing towards the greater literature with regard to stock market forecasting in emerging markets.

Assessing Simulation Accuracy and Reliability

In any case, ensuring the reliability of the Monte Carlo Simulation model for this study involves assessing the validity of simulated stock price distributions against historical data metrics like mean, variance, and volatility. While traditional MCS provides only a probabilistic range of future outcomes, the model's predictive reliability can be further advanced by estimating input parameters through advanced statistical techniques. Then, there is ABC that contains Approximate Bayesian Computation with Sequential Monte Carlo Sampling and Adaptive Importance Sampling as strong methods to update the input parameters without explicit usage of likelihood functions see in [10]. These likelihood-free Bayesian methods enable parameter estimation when direct computation is not feasible, thus enhancing model accuracy. Integrating these would allow better validation of simulation results through iterative tuning of input parameters, thereby improving the model's capability to accurately project underlying market dynamics. Although this approach was not used in the current study, its adoption offers a viable avenue for future research, particularly when it comes to the situation of stock price

simulation in emerging markets with limited historical data. Using this method will result in improved model accuracy that is consistent with current financial market stochastic modelling techniques rather than being ambiguous.

Methodology

This study tries to predict the future values of some selected stocks in the Turkish automotive industry, namely the equities of Dogus Automotive (DOAS), Tofas (TOASO), and Ford Otosan (FROTO) listed in BIST, by using the approach of Monte Carlo Simulation. The aim is to provide a probabilistic outlook for the future prices of stocks, hence enabling investors to have the insight they may need to make better decisions in these volatile markets. This analysis uses data from investing.com for end-of-day closing prices of the subject stocks from January 1, 2021, through December 31, 2023. There are 749 records of each stock in the dataset, representing three complete years of up and down changes. These historical prices form the basis for the estimation of future stock movements by simulation.

Python was required to be used for generating the Monte Carlo Simulation, as it can model the randomness and uncertainty at which the prices change. Through the simulation, the code will use the provided historical return values to calculate the value of daily and annual volatility for every security. For volatility here, or the standard deviation of daily returns, quantified, is very critical in characterizing the size and frequency of variations in price over time. Calculation of volatility includes computation of the standard deviation in daily changes and scaling up the same into an annual level, considering 365 or 252 days. MCS was run with 10.000 random values, running the stock prices ten times for generating a range of possible future scenarios. The algorithm follows a basic stochastic process:

Monte Carlo Simulation Algorithm

The process of Monte Carlo simulation used in this study is structured as follows:

Input Parameters

P_0 : Initial stock price (determined from historical data).

μ : Anticipated daily return (presumed to be 0 for simplicity).

σ : Daily volatility (determined by utilising historical data):

$$\sigma = \frac{\text{Annual Volatility}}{\sqrt{\text{Number of Trading Days}}}$$

T : The simulation's duration (244 days in this study).

N : The quantity of simulations (10 on each stock).

Δt : Time step (1 day).

Output

P_t^i : Stock prices simulated over T days for N simulations.

Algorithm

Initialize Parameters:

Extract P_0 based on the stock's initial price.

Compute σ (daily volatility) from historical data.

Run Simulations

For $i = 1$ to N :

$$\text{Set } P_t^i[0] = P_0.$$

For $t = 1$ to T :

Generate a random variable Z_t from $N(0,1)$.

Compute $P_t^i[t]$ using the formula:

$$P_t^i[t] = P_t^i[t-1] \times \exp\left((\mu - 0.5\sigma^2)\Delta t + \sigma\sqrt{\Delta t}Z_t\right)$$

Aggregate Simulation Results

Determine the mean, median, and standard deviation of descriptive statistics for each simulation.

Output Results:

Return the P_t^i range of potential stock prices together with related statistics.

These parameters form the basis of our Python simulation, which aims to mimic stock price evolution over a prescribed period. Each time this simulation is run, a new time series of stock prices is obtained, which can model the random behaviour of the market as well as historical volatility patterns. When iterated several times, the simulation captures the range of possible outcomes; thus, providing a probabilistic distribution of future stock prices. This method not only estimates the central tendency, i.e. the mean, but also outlines the extremes and variability of stock prices.

After the execution of MCS, the outputs were processed to derive a mean, median, standard deviation, and range of forecasted prices for all stocks. Descriptive statistical values, which are the mean price-DOAS at 101.96 TL, TOASO at 116.55 TL, and FROTO at 396.21 TL—are matched against their respective observed historical values to assure the correctness and relevance of the simulation. The critical evaluation of the volatility values for each stock is 0.0333 for DOAS, 0.0296 for TOASO, and 0.0298 for FROTO. The values give insight into each company's exposure to the swinging of prices and risk, which is significant for investors with a view to calculating market stability.

Furthermore, a sensitivity analysis has been done in this respect by changing the number of iterations and volatility levels to see its reflection in forecast accuracy. This will help in making the model robust under different scenarios and lend more credibility to the results. The following results identify some aspects of how external market forces—economic conditions or political events—might influence stock behaviour and point to the importance of a simulation-based approach in investment planning. The Python code used in this analysis section is hosted on the GitHub page of the project for reproducibility and transparency. Based on this, other researchers and practitioners can review the methodology and replicate it, hence allowing the model to be adapted for further studies or different market

conditions. This gives a systematic approach, showing in practice how the Monte Carlo Simulation applies to financial forecasting, tailor-made for specific automotive cases in Turkey. At the same time, this is an instrument for investors to devise strategic moves that reduce their risk and further optimize their investment portfolio. The Python codes and simulation results used in this study are made openly available on GitHub for transparency and reproducibility. Scripts for Monte Carlo simulations, stock price forecasting models, and data pretreatment are all included in the repository. Access to the codes is available on the Monte Carlo Simulation in Stock Price Forecasting.

A technique that generates random numbers is used in Monte Carlo Simulation to examine the impact of random variables on a system. By employing the inverse of the distribution function, the input variables are produced in accordance with a particular probability distribution. The inverse probability distribution function can be used by computers to convert uniformly quickly and easily generated random integers in the range [0-1] into any distribution. Assume, for instance, that the function f represents the relationship between the dependent variable Y and two random variables (X_1, X_2). The relationship in this instance can be stated as follows:

$$Y = f(X_1, X_2)$$

If the X_1 , distributions and X_2 , are known (such as the Normal Distribution), the distribution parameters (μ_x : Mean and σ_x : Standard Deviation) are also provided, random numbers for X values can be generated multiple times, and the associated Y values can be computed by using the function f . As a result, statistics regarding Y can be derived. A bigger sample size improves the accuracy of the results. To achieve optimal results when applying Monte Carlo Simulation to deterministic systems, a sufficient number of random numbers must be generated. These figures could be in the thousands or tens of thousands in some applications, underscoring the necessity for computing power. Monte Carlo simulation has shown to be a dependable method in stochastic simulation processes when properly planned and applied with a big enough sample size [11]. Monte Carlo Simulation can be used to produce desired distributions by producing an adequate number of random numbers with uniform distributions. Following a few guidelines and making the most of Monte Carlo Simulation at each stage of the process will guarantee high-quality simulation.

The Monte Carlo Method's Mathematical Analysis

Numbers uniformly distributed between 0 and 1 are a basic tool for numerically modelling an experiment or event in the Monte Carlo method. Usually represented by the letter q , these numbers are produced by computer programmes. These arbitrary numbers are gathered from measurement or experiment set values. The probability of each number, however, varies, with some numbers having larger or lower probabilities than others. As opposed to sets with equal probabilities, this results in a set with varied probabilities. A set of random numbers is said to be

uniformly distributed when all the probabilities are equal, meaning that any value has the same probability of happening. This technique is very helpful for projecting or modelling the results of a given experiment or occurrence. A significant number of random numbers, frequently produced by computers, are needed for the Monte Carlo approach. These numbers have the statistical characteristics of random numbers even though they are generated sequentially in accordance with a predetermined method, hence they are not completely random. Numerous numerical analysis and simulations employ this technique. The 'Mixed Congruential Method,' denoted by the following formula, can be used to create random numbers:

$$P_i = \text{int} \left(\frac{a \cdot x_i}{b \cdot r_i} \right) \tag{1}$$

$$X_{i+1} = a \cdot x_i - b \cdot r_i \tag{2}$$

$$q_i = \frac{x_{i+1}}{b} \tag{3}$$

The following relation can be used to display the algorithm for this method:

$$x_i = (a \cdot x_{i-1} + c) \pmod{m} \tag{4}$$

where the initial value of the positive integer sequence x_i is x_0 . m is a greater positive integer than a and b , which are also positive integers. After multiplying x_{i-1} by a , the sequence x_i is determined. The modulus regarding m is then calculated. The starting value, x_0 , in the "Mixed Congruential Method," is a positive integer. The created number sequence's elements are divided by m to create a new series with values between 0 and 1. The properties of the random number sequence are determined by the values a , c , and m . The sequence has a finite end and will eventually recur; the period of the repetition will depend on the values of m , a , and c . It can be extended the duration by choosing suitable values for m , a , and c .

Let us consider an event for which we want to model the probability distribution over an interval (a, b) , where the frequency function $f(x)$ gives each value x a specific likelihood of occurring. The following calculates the likelihood that a value will fall between x and $x + dx$:

$$P(x) \cdot dx = \frac{f(x) \cdot dx}{\int_a^b f(x) \cdot dx} \tag{5}$$

The Probability Density Function, or $P(x)$, is defined as follows for the Total Probability Density Function:

$$Q(x) = \int_a^x P(x') \cdot dx' \tag{6}$$

Over the interval (a, b) , the value of $Q(x)$ assumes random values in the 0– 1 range. We can relate $P(x)$ to T since the frequency function, which shows how often each value x occurs, has a uniform distribution:

$$T = Q(x) \quad (7)$$

The following is the derivation of the Fundamental Monte Carlo Principle using Equations (5), (6) and (7):

$$T = \frac{\int_a^x f(x') \cdot dx'}{\int_a^b f(x) \cdot dx} \quad (8)$$

The Fundamental Monte Carlo Principle is the name given to this equation. Depending on T , the following inversion transformation is obtained by inverting in Equation (8), [12]:

$$x = P^{-1}(T) \quad (9)$$

In this part of the study, a Monte Carlo simulation will be used to project the future value of specific automotive industry stocks for those who are interested in investing in the sector. This section will give an overview of the Turkish automotive industry and its volatility before doing a data analysis. The end-of-day closing prices of equities traded between January 1, 2021, and December 31, 2023, were the source of the data used in this study, which was obtained from investing.com. Monte Carlo simulation was used with Python programming language to generate 1,000 random values and simulate the equities ten times to forecast their prices. Educating prospective investors is the goal.

Findings and Discussion

The Automotive Sector in Turkey

Turkey's automobile industry is an essential part of the nation's industrial and economic framework. Turkey's industrialization progress has been greatly aided by the production of automobiles, a sector whose diversity and capacity are growing, and which makes the nation more competitive internationally. Turkey is one of the biggest automakers in Europe. The industry produces a wide range of goods, including as trucks, buses, commercial vehicles, and agricultural machines, in addition to passenger cars. Turkey's strength and adaptability in the automotive sector are strengthened by this diversity.

A significant number of the cars made are exported. Turkey's trade balance and economic growth are significantly influenced by its automobile exports. Turkish automakers provide premium cars at affordable prices to be competitive in global markets. Turkey's automobile industry has prospered mostly because of a strong supply chain and logistics system. Manufacturers maintain a streamlined production process by effectively utilising parts acquired from both domestic and international vendors. Furthermore, significant resources are allocated towards research and development (R&D) to consistently enhance production procedures and generate inventive resolutions. In conclusion, Turkey's automobile industry is essential to the nation's economic growth. With a robust supply chain, wide export network, high production capacity, and an emphasis on research and development, Turkey is likely to maintain its prominent position in the global automotive industry and flourish going forward.

Volatility

The term "volatility" refers to the variation in asset prices in financial markets during a specific time period. It

gauges how frequently and how strongly prices shift in unexpected ways. Put differently, volatility represents the speed and range of price changes for an asset. Usually, statistical computations like standard deviation or percentage change rates are used to quantify volatility. These measurements are employed to examine how the price of an asset changes over a given period. Low volatility denotes more steady and predictable market movements, whereas high volatility indicates the possibility of abrupt and significant price swings.

Volatility is a key concept for investors since low volatility offers more stable and low-risk investment options, while high volatility frequently raises investment risk. Investors can manage risk, create investment strategies, and optimise their portfolios to meet return goals by analysing volatility. As a result, investors keep a careful eye on volatility since it is seen as a key indication in financial markets. By calculating the standard deviation of daily fluctuations using historical data, annual stock volatility is computed. By dividing the yearly volatility by the square root of the number of trading days in a year, one may get the daily volatility.

Table 1. Descriptive Statistical Values of Three-Year Share Prices (TL) of DOAS, TOASO and FROTO

| Statistic | DOAS(TL) | TOASO(TL) | FROTO(TL) |
|--------------------|----------|-----------|-----------|
| Mean | 101,96 | 116,55 | 396,21 |
| Standard Error | 2,99 | 3,12 | 9,19 |
| Median | 71,27 | 79,06 | 294,58 |
| Mode | 23,00 | 26,99 | 154,46 |
| Standard Deviation | 82,07 | 85,33 | 251,45 |
| Variance | 6735,95 | 7280,773 | 63226,09 |
| Kurtosis | 0,93 | 0,76 | 0,85 |
| Skewness | -0,37 | -0,81 | -0,65 |
| Range | 297,48 | 279,06 | 854,31 |
| Minimum Value | 19,16 | 24,64 | 110,69 |
| Maximum Value | 316,64 | 303,70 | 965,00 |
| Volatility | 0,033 | 0,03 | 0,03 |
| Number of Points | 749 | 749 | 749 |

The study's descriptive statistics for the equities under examination are shown in

Table 1. There are 749 data points for the stocks that are included in this simulation. Dogus Automotive has an average price of 101.96TL, Tofas has an average price of 116.55 TL, and Ford Otosan has an average price of 396.21TL when examining the three-year averages of the stock prices. The range of prices for DOAS was 19.16 TL at the lowest and 316.64 TL at the most. The lowest TOASO price was 24.64 TL, while the highest price was 303.70 TL. The lowest and maximum prices for FROTO were 110.69 and 965 TL, respectively. These stock markets' three-year volatility values are as follows: The volatility of Dogus Automotive (DOAS), Tofas (TOASO), and Ford Otosan (FROTO) is 0.0333, 0.0296, and 0.0298, respectively. The difference in volatility between FROTO and DOAS is 0.0035, and the difference between TOASO and DOAS is 0.0037. The volatility of FROTO and TOASO differs by 0.0002. Therefore, in comparison to the other companies,

Dogus Automotive has displayed more notable price swings over time in terms of volatility.



Figure 1. Line chart of DOAS share prices between 01.01.2021-31.12.2023.

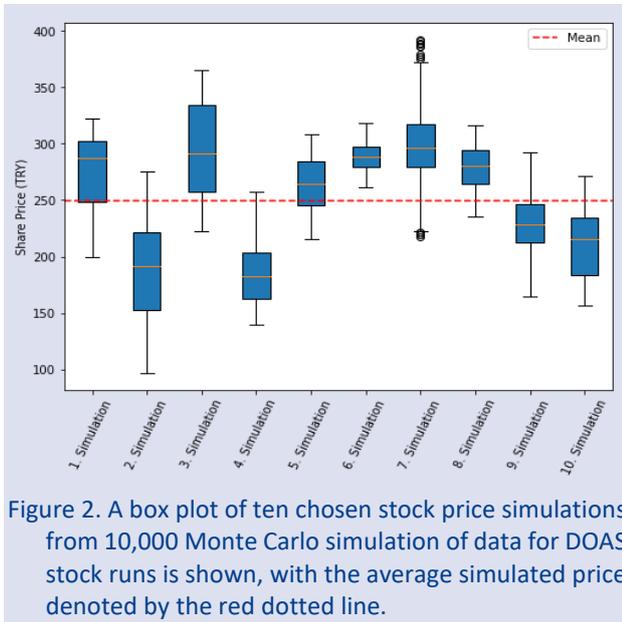


Figure 2. A box plot of ten chosen stock price simulations from 10,000 Monte Carlo simulation of data for DOAS stock runs is shown, with the average simulated price denoted by the red dotted line.

A line graph showing Dogus Automotive's share values over a three-year period is shown in Figure 1. Over a period of more than a year, from January 4, 2021, to March 4, 2022, the prices remained below 50 TL. The price

increased to over 185 TL after exceeding 50 TL, then decreased to approximately 135 TL. Prices demonstrated a rising trend from May 4, 2023, to September 4, 2023. During this time, they reached at 316.64 TL, following which they began to decrease

The results of ten simulations of the DOAS stock price over a three-year period are shown in the box plot in **Hata! Başvuru kaynağı bulunamadı..** From 200 to 300 TL forward, most of the simulations converge, with very little deviation until day 82 or so. This convergence suggests consistent early-stage patterns because it shows a shared beginning trajectory across the simulations. The simulations begin to deviate around day 82, suggesting more volatility in the behaviour of stock prices. The impact of economic variables, market dynamics, or other unforeseen occurrences influencing stock prices could be reflected in this disparity. As the simulated time ends, the results point to an increased number of possible outcomes; most simulations estimate stock prices in the range of 100 to 300 TL. The seventh, eighth, and tenth simulations exhibit more extreme behaviour, underscoring the need to consider the entire range of possible results and the risk of outliers. This box plot highlights the need of conducting several runs to accurately represent the range of possible market behaviours by illuminating the various degrees of uncertainty in long-term stock price simulations.

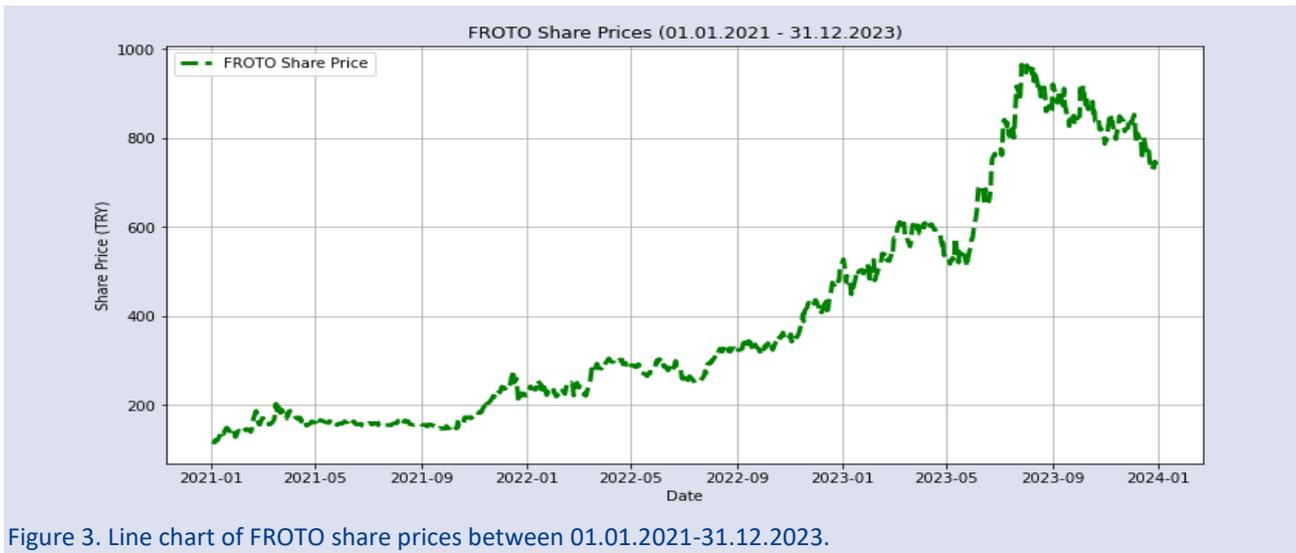


Figure 3. Line chart of FROTO share prices between 01.01.2021-31.12.2023.

A line graph illustrating the share prices of Ford Otosan during a three-year period appears in Figure 3. The share price remained below the 200 TL amount from January 4, 2021, until November 4, 2021. Prices changed between 200 and 400 TL on November 4, 2022, and November 4, 2021. On March 4, early in 2023, the share price launched an unsuccessful attempt to exceed 600 TL.

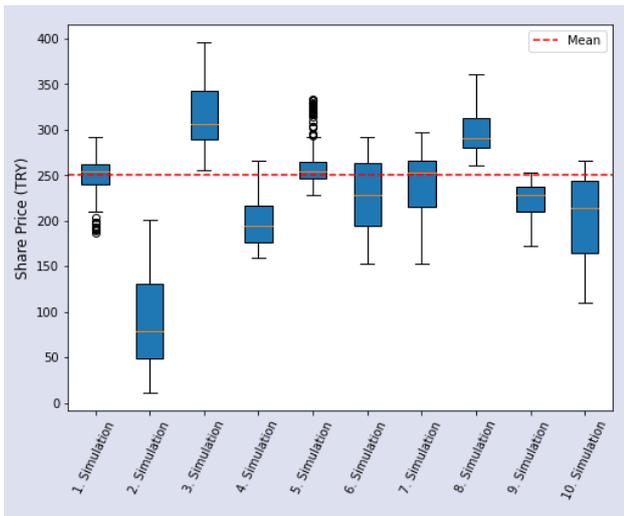


Figure 4. A box plot of ten chosen stock price simulations from 10,000 Monte Carlo simulation of data for FROTO stock runs is shown, with the average simulated price denoted by the red dotted line.

However, the share price increased above 600 TL between May 4, 2023, and July 4, 2023, eventually

peaking at 965 TL. The share price demonstrated a declining tendency after this level. **Hata! Başvuru kaynağı bulunamadı.** illustrate the outcomes of ten simulations for a three-year forecast of FROTO stock prices. As a steady baseline throughout simulations, share prices normally oscillate at the beginning between 600 and 850 TRY. This relative stability lasts until about day 127, at which point the simulations start to show more divergence, suggesting that as the timeframe goes on, stock price volatility would rise. Two unique patterns appear as the simulations approach closer to their endings. The eighth simulation in the first pattern presents a variety of share values that increase to 900 TRY. The remaining simulations illustrate the second pattern, demonstrating share price variations within a narrower range, usually between 600 and 800 TRY, indicating a more cautious growth trajectory. In the previous simulation, there is an outlier that reaches over 900 TRY, indicating an uncommon occurrence or extreme circumstance. With an expanded distribution and the highest median of any simulation, the eighth one suggests more variability or uncertainty. The diversity of these simulations highlights varying market circumstances and possible stock price movements, highlighting the intrinsic uncertainty of long-term forecasts. A range of results from this divergence can be utilised to evaluate risk and guide investment plans.



Figure 5. Line chart of TOASO share prices between 01.01.2021-31.12.2023.

A line graph representing the share prices of Tofas during a three-year period can be seen in Figure 5. At less than 50 TL, the share price remained until November 4, 2021. There was a 50–100 TL variation in the share price from November 4, 2021, and November 4, 2022. Between July 4, 2023, and September 4, 2023, Tofas reached its peak value of 303.70 TL. After that, it continued to decline and eventually reached approximately 200 TL.

After ten simulation tests utilising a three-year period of TOASO stock data, the box plot in Figure 6 displays a starting range for share values between 150 and 250 TL. This first convergence suggests a consistent starting point for the simulations, implying that they started with remarkably similar values.

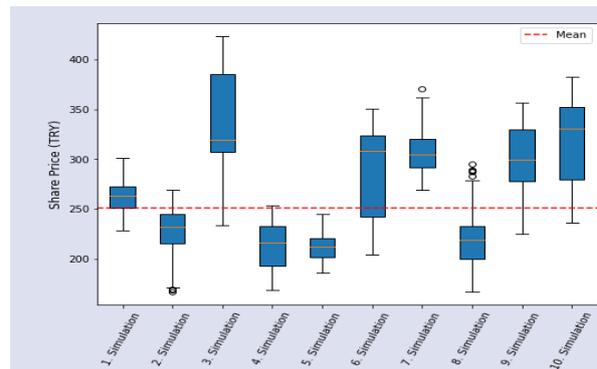


Figure 6. A box plot of ten chosen stock price simulations from 10,000 Monte Carlo simulation of data for TOASO stock runs is shown, with the average simulated price denoted by the red dotted line.

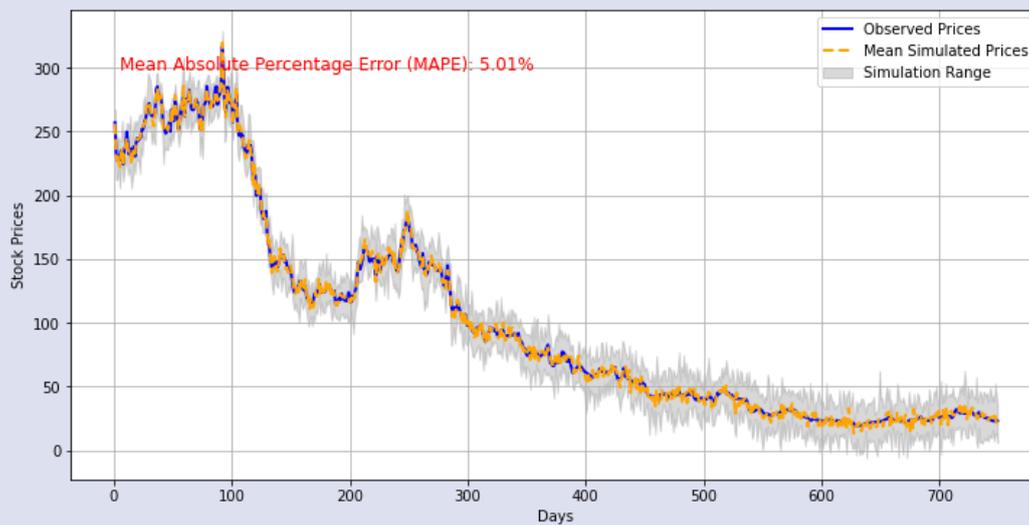


Figure 7. A comparison of the mean absolute percentage error (MAPE) of 5.01% between the observed and simulated stock prices.

The Figure 7 highlights the model's performance by comparing observed stock prices over a certain period with their simulated counterparts. The orange dashed line (mean simulated prices), and the grey shaded region (simulation range) represent the simulated data, and the blue line represents the observed prices, which are used as a reference to assess the correctness of the simulated data. The simulation approach effectively represents the

underlying patterns and variability of the observed stock prices, as evidenced by the comparatively low Mean Absolute Percentage Error (MAPE) of 5.01%. Although the fluctuation within the simulation range highlights the need for additional refinement to reduce uncertainty in forecasts, this alignment points to the model's potential for real-world applications in forecasting or risk assessment.

We calibrated the simulation parameter (standard deviation) using Approximate Bayesian Computation (ABC) to guarantee simulation correctness. This strategy minimized the distance between the simulated and observed stock price statistics (mean and standard deviation). The calibration procedure appears in Figure 8 a), where candidate standard deviation values were assessed according to simply how far they deviated from historical measures.

The ideal standard deviation, which minimises the discrepancy between simulated and observed statistics, is shown by the red dotted line. By bringing the model into line with actual stock price movements, this calibration improves simulation accuracy. b) uses the ideal parameter to compare the distributions of stock prices in the simulation with the real world. The blue and red dashed lines display the observed and simulated means, respectively, and the histogram depicts the simulated prices. The simulation's authenticity is confirmed by the tight alignment of both means and the comparable distribution spread.

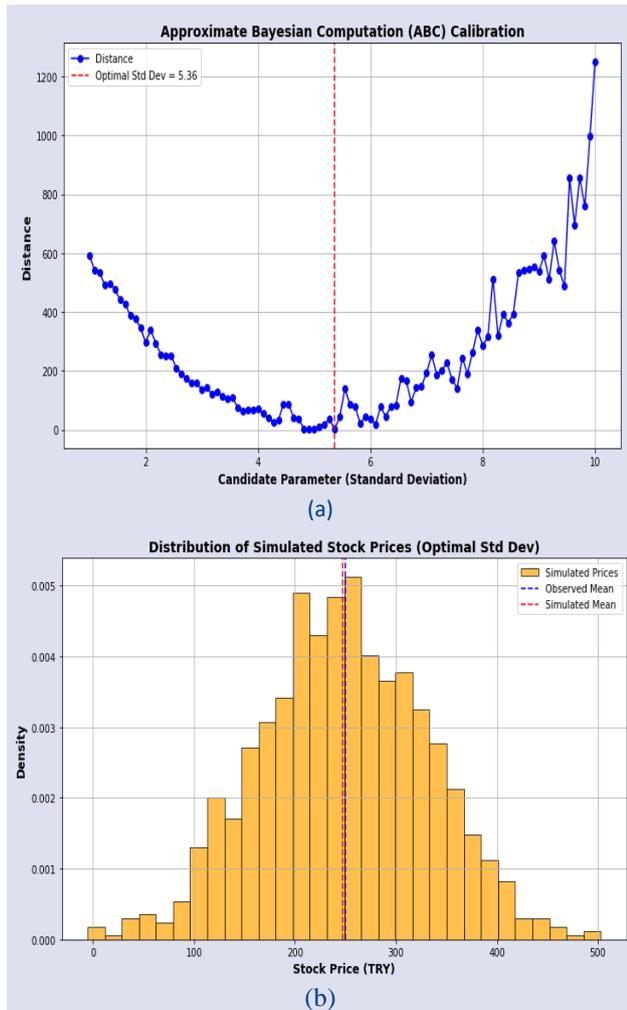


Figure 8. a) The difference in stock price statistics between simulations and observations as a function of the chosen standard deviation parameter. b) A comparison between the observed historical mean (blue dashed line) and the simulated stock price distribution (orange histogram).

The prices largely followed a similar trajectory over the course of the simulations, suggesting little variation or variance in the results. This initial phase stability raises the possibility that common underlying forces are driving the stock price, resulting in a consistent pattern across simulations. But as the simulations continued, especially in the last part of the year, the box plot clearly demonstrates a divergence, with share prices changing dramatically from simulation to simulation. This greater range suggests that as time passed on, outside influences or arbitrary movements had a bigger impact on the stock price, which increased its volatility and unpredictability. The inherent unpredictability of long-term stock price simulations is demonstrated by this box plot pattern, underscoring the significance of performing repeated simulations to capture a wider range of outcomes. The final divergence highlights the possibility of large fluctuations in stock price, which is important for risk assessment and financial planning.

Table 2. Monte Carlo Simulation Values of Shares (TL)

| | DOAS | TOASO | FROTO |
|--------------------|----------|----------|----------|
| Starting Price | 257 | 210 | 739,5 |
| Daily Volatility | 3,33% | 2,96% | 2,99 % |
| Annual Volatility | 52,78% | 47,02% | 47,4% |
| Mean | 254,9790 | 203,2083 | 742,6018 |
| Median | 221,3507 | 181,4836 | 681,8091 |
| Standard Deviation | 137,1685 | 96,3172 | 344,0281 |
| Quartiles | | | |
| 25 % | 157,8000 | 136,2762 | 504,5080 |
| 50 % | 221,3507 | 181,4836 | 681,8091 |
| 75 % | 318,7190 | 248,3489 | 923,4383 |

Investigating at the annual volatility rates presented in Table 2, we can see that Dogus Automotive has an annual volatility of 52.78%, Tofas has an annual volatility of 47.02%, and Ford Otosan has an annual volatility of 47.4%. This suggests that Tofas and Ford Otosan have comparable annual volatility, although Dogus Automotive has marginally greater volatility. When comparing the averages, the share prices of Dogus Automotive, Tofas, and Ford Otosan are 254.98 TL, 203.21 TL, and 742.60 TL, respectively.

Conclusion

Transformations and volatility in financial markets cause risk to become more pronounced for both small and large investors. In this process, managing and minimizing risk effectively is becoming increasingly important. For this reason, there is an increasing interest in risk-related studies in various disciplines. Predicting risk and identifying all possible scenarios in advance and taking precautions is becoming increasingly critical for investors. The complexity and uncertainties in the global financial structure encourage the continuous development of new approaches and strategies in risk management. Investors are increasingly relying on analytical tools and risk management techniques based on mathematical models

to reduce the risks in their portfolios and minimize potential losses. This contributes to the adoption of a more informed and well-founded approach to financial decision-making. The Monte Carlo Simulation method allows us to obtain different measurement results with random numbers. Therefore, it is necessary to look at the environment in which the companies are located when analysing the values that emerge in this simulation research. This study has been conducted to give an insight to people who desire to purchase shares in the automotive sector. The annual prices of Dogus Automotive, Tofas and Ford Otosan stocks for three years are estimated by Monte Carlo Simulation. The results of Monte Carlo Simulation were analysed in the study. According to the simulation study, the annual forecast values of the three stocks examined were reached with ten different simulations. When we look at the volatilities of the stocks, it is observed that the annual volatility of Dogus Automotive is 52.78%, the annual volatility of Tofas is 47.02% and the annual volatility of Ford Otosan is 47.40%. In this case, the annual volatilities of Tofas and Ford Otosan are close and Dogus Automotive is slightly higher than the others. When daily volatilities are analysed, it is observed that Dogus Automotive, Tofas and Ford Otosan have 3.33%, 2.96% and 2.99%, respectively. The daily volatilities of all three stocks are close to the annual volatilities of Tofas and Ford Otosan, while Dogus Automotive is slightly higher. It is observed that all three stocks experience similar fluctuations in the financial markets. In this simulation, when the initial and average

prices of the shares are analysed, it is observed that the initial price of Dogus Automotive in this simulation is TL 257 and the average price is TL 254.9790. In the simulation for Tofas, the initial price was TL 210, and the average price was TL 203.2083 and in the simulation for Ford Otosan, the initial price was 739.5 and the average price was TL 742.6018. In the comparison of the standard deviations of the three stocks in the simulation, it was observed that the standard deviation of Dogus Automotive data was 137.1685. The standard deviation of Tofas's data is 96.3172. Finally, the standard deviation of the data belonging to Ford Otosan was observed as 344.0281. When only the standard deviations of the three shares are considered, it is observed that there is a difference between them. As a result of the simulations, it was observed that all three stocks were similar, their volatilities were very close, and Tofas was slightly ahead of the others when looking at their standard deviations. However, it should be kept in mind that various factors such as market conditions, economic conditions, political events, trade volume, natural disasters, company activities, speculation will affect the prices of the shares. In this simulation, share prices are estimated only to give an idea to the investor.

Conflicts of interest

There are no conflicts of interest in this work.

References

- [1] Edjossan-Sossou A. M., Sustainable risk management strategy selection using a fuzzy multi-criteria decision approach, *Int. J. Disaster Risk Reduct.*, 45 (2019) 101474.
- [2] M. Broadie, Operations Research jUffl rms Risk Estimation via Regression, 63(5) (2024) 1077–1097.
- [3] David Iyanuoluwa A., Rhoda Adura A., Onyeka Franca A., Oluwaseyi Rita O., Binaebi Gloria ., Ndubuisi Leonard N., Review of Ai Techniques in Financial Forecasting: Applications in Stock Market Analysis, *Financ. Account. Res. J.*, 6(2) (2024) 125–145.
- [4] Prakash A., Mohanty R.P., DEA and Monte Carlo simulation approach towards green car selection, *Benchmarking*, 24(5) (2017) 1234–1252.
- [5] Sivabalan S., Minu R.I., *Statistical Sales Forecasting Using Machine Learning Forecasting Methods for Automotive Industry*, vol. 686 LNNS. Springer Nature Singapore, (2023).
- [6] Čečević B. N., Antić L., Jevtić A., Stock Price Prediction of the Largest Automotive Competitors Based on the Monte Carlo Method, *Econ. Themes*, 61(3) (2023) 419–441.
- [7] Borchert P., Zellmer-Bruhn D. M., Reproduced with permission of the copyright owner . Further reproduction prohibited without, *J. Allergy Clin. Immunol.*, 130(2) (2010) 556.
- [8] Zhao X., Liu H., Zhang B., Liu F., Luo J., Bai J., Fast photon-boundary intersection computation for Monte Carlo simulation of photon migration, *Opt. Eng.*, 52(1) (2013) 019001.
- [9] Twumasi C., Twumasi J., Machine learning algorithms for forecasting and backcasting blood demand data with missing values and outliers: A study of Tema General Hospital of Ghana, *Int. J. Forecast.*, 38(3) (2022) 1258–1277.
- [10] Twumasi C., Cable J., Pepelyshev A., Mathematical Modelling of Parasite Dynamics: A Stochastic Simulation-Based Approach and Parameter Estimation via Modified Sequential-Type Approximate Bayesian Computation, vol. 86, no. 5. Springer US, 2024.