DECEMBER '25

INTERNATIONAL REVIEWS, RESEARCH AND STUDIES IN THE FIELD OF NATURAL SCIENCES

EDITORS
PROF. DR. EDİP BAYRAM
PROF. DR. HASAN AKGÜL

SERÜVEN YAYINEVİ

Genel Yayın Yönetmeni / Editor in Chief • C. Cansın Selin Temana

Kapak & İç Tasarım / Cover & Interior Design • Serüven Yayınevi

Birinci Basım / First Edition • © Aralık 2025

ISBN • 978-625-8682-07-6

© copyright
Bu kitabın yayın hakkı Serüven Yayınevi'ne aittir.
Kaynak gösterilmeden alıntı yapılamaz, izin almadan hiçbir yolla çoğaltılamaz. The right to publish this book belongs to Serüven Publishing. Citation can not be shown without the source, reproduced in any way without permission.

Serüven Yayınevi / Serüven Publishing
Türkiye Adres / Turkey Address: Kızılay Mah. Fevzi Çakmak 1. Sokak
Ümit Apt No: 22/A Çankaya/ANKARA
Telefon / Phone: 05437675765
web: www.seruvenyayinevi.com
e-mail: seruvenyayinevi@gmail.com

Baskı & Cilt / Printing & Volume
Sertifika / Certificate No: 47083

INTERNATIONAL REVIEWS, RESEARCH AND STUDIES IN THE FIELD OF NATURAL SCIENCES

EDITORS

PROF. DR. EDİP BAYRAM

PROF. DR. HASAN AKGÜL

CONTENTS

Chapter 1

DEVELOPMENT OF GATE DIELECTRICS IN SEMICONDUCTOR DEVICES AND MODERN CMOS STRUCTURES

MERVE EREN YAKIŞIKLIER[1]

1 Assistant Professor, Yozgat Bozok University, Department of Medical Services and Techniques, Orcid: 0000-0002-6920-359X

The Strategic Importance of Dielectric Materials in Chip Manufacturing

In the development of semiconductor technology, dielectric materials, especially the thin insulating layer forming the gate dielectric in MOSFET structures, are one of the most critical components directly determining device performance. In the traditional MOS structure, this layer is formed with silicon dioxide (SiO_2) and has existed as the industry standard for over forty years. SiO_2's high band gap, strong chemical stability, and perfect interface with silicon have resulted in low trap density, high electron mobility, and stable electrical characteristics (Nicollian & Brews, 1982). Despite these advantages, the downscaling trend predicted by Moore's Law has led to reaching the physical limits of SiO_2, making a re-evaluation of dielectric materials necessary. Particularly since the late 1990s, as the thickness of the MOSFET gate oxide has decreased to a few nanometers, gate leakage currents have rapidly increased due to quantum tunneling effects, becoming a critical problem threatening both power consumption and device reliability (Kingon, Maria & Streiffer, 2000). The increasing importance of the EOT (Equivalent Oxide Thickness) concept due to scale-down has made it impossible to achieve the same capacitance values with smaller physical thicknesses using SiO_2. Therefore, the transition to new materials with high dielectric constant (high-k) has emerged as a technological necessity where not only performance requirements but also power efficiency and reliability criteria are decisive. This transformation has not only meant selecting a new insulating layer; it has also changed the entire MOSFET design paradigm in terms of gate electrode, interface layer, gate stack architecture, and manufacturing processes.

With high-k materials, polysilicon gates have also reached their performance limits, and metal gate technologies (HKMG – High-k/Metal Gate) have become the industry's standard solution (Robertson, 2006).

These developments have moved the semiconductor industry from the classic SiO_2-based MOS design to a completely new era of insulator-conductor engineering. Therefore, the evolution of dielectric materials in chip manufacturing is a multi-dimensional development area where not only materials science but also electrical stability, scalability, thermal process management, and nanostructure interface engineering intersect. This section will discuss in detail the advantages and limitations of SiO_2, the emergence of nitride-doped dielectrics, the necessity of transitioning to high-k dielectrics, the introduction of HfO_2-based insulators to the industry, and the predicted dielectric innovations for sub-1 nm processes in the future.

The Evolution of SiO_2 Gate Dielectrics

Silicon dioxide (SiO_2) has been the most critical dielectric material in the semiconductor industry from the 1960s, when the metal-oxide-semiconductor (MOS) structure emerged, until the early 2000s. The reason SiO_2 has been the standard for so long is its exceptional compatibility with silicon and its ability to form a high-quality interface.

The growth processes of silicon oxide, especially thermal oxidation, have yielded both controlled and reproducible results, thus achieving low trap density, high breakdown voltage, and long-term reliability in MOS gate dielectrics (Deal & Grove, 1965). The low defect rate at the Si–SiO_2 interface has supported high electron and hole mobility and played a critical role in the development of MOSFET architecture.

As MOSFET channel lengths have shrunk from micrometer levels to nanometers, the thickness of the gate oxide has also been reduced proportionally. By the late 1990s, SiO_2 layers in industrial production had thinned to levels of 1.5–2 nm. However, at this point, quantum mechanical tunneling effects became prominent, and gate leakage began to increase dramatically (Lombardi et al., 1988). At these levels, electron tunneling through the oxide layer increased static power consumption and

became a fundamental physical limit threatening device reliability. Another factor limiting the scalability of SiO_2 is the concept of equivalent oxide thickness (EOT). EOT is a metric that expresses the capacitance advantage provided by high-k dielectrics in terms of SiO_2. For example, to achieve a 1.2 nm EOT, SiO_2 must be physically approximately 1.2 nm thick; however, this thickness increases quantum tunneling leakage to unacceptable levels. Therefore, physical thicknesses around 1 nm are no longer feasible with SiO_2 (Frank et al., 2001). This clearly demonstrates that scaling down targets are not sustainable with SiO_2. SiO_2 growth processes have undergone significant improvements over time.

Methods such as dry oxidation, wet oxidation, and chlorine-doped oxidation (HCl, TCA – trichloroethylene) have improved interface quality, reduced total trap density, and strengthened breakdown properties. Despite this, the physical limits have not changed, and achieving an EOT below 1 nm has not been possible for SiO_2. Therefore, the fundamental problem that led to the "gate dielectric crisis" in industry in the early 2000s was that SiO_2 could no longer meet the needs of MOSFET scaling. However, SiO_2 has not been completely abandoned; it has continued to exist as an interface layer in high-k dielectric stacks. Because direct contact between high-k materials and silicon creates problems such as high interface trap density and low carrier mobility. Therefore, in a typical modern gate-stack architecture, a layer of SiO_2 or SiON a few angstroms thick is still a critical component (Robertson, 2006).

In conclusion, the evolution of SiO_2 formed the backbone of MOSFET history; however, reaching the limits of scalability made the transition to new dielectric materials inevitable. In the next period, nitride-doped dielectrics and later high-k oxides filled this gap, initiating a new paradigm in semiconductor manufacturing.

Transition from SiO_2 to SiON and Nitride Derivatives

As the physical limits of SiO_2 gate dielectrics were approached, intermediate solutions were developed to improve the properties of the dielectric layer in order to continue MOSFET scaling. In this context, silicon oxynitride (SiON) and nitride derivative gate dielectrics played a critical bridging role in industry from the late 1990s to the mid-2000s. The main reason for the emergence of SiON was that when SiO_2 thickness was reduced below 1.5–2 nm, leakage currents increased unacceptably and the dielectric constant ($k \approx 3.9$) could no longer provide the desired capacitance (Wong & Iwai, 2006). SiON layers mitigated some of these problems and also contributed to the adaptation of production lines to new chemical processes before the transition to high-k dielectrics. SiON dielectrics are generally obtained by nitrogen implantation, plasma nitriding (RPN – Radical Plasma Nitridation), or thermal nitriding methods. This structure, obtained by diffusion of nitrogen atoms into SiO_2, reduces electron tunneling by increasing both the band gap and barrier height, thus enabling the creation of thinner physical layers with low leakage current levels (Chau et al., 2005). In addition, since the dielectric constant of SiON is slightly higher than that of SiO_2 ($k \approx 4.5–6$), a limited improvement in EOT values is achieved. This property has made SiON the main gate dielectric in industry, especially in 90 nm and 65 nm CMOS processes. Another important advantage of SiON is that it forms a barrier against boron diffusion and dopant penetration. In polysilicon-gated MOSFETs, diffusion of dopant atoms through SiO_2 into the channel leads to significant threshold voltage (V_TH) shifts, while the SiON layer significantly reduces this diffusion (Hobbs et al., 2001). Therefore, SiON acts as both a dielectric and a diffusion barrier, increasing fabrication reliability. However, SiON dielectrics also have significant limitations in scaling. As nitrogen concentration increases, interface trap density increases, leading to both decreased electron mobility and deterioration of device stability. In addition, increasing nitride content negatively affects band alignments, leading to undesirable deviations in gate leakage (Robertson, 2006). Therefore, although effective for the short term, SiON dielectrics have not been able to approach the EOT values that high-k materials can

provide in the long term. The most important role of SiON in industry is to provide the infrastructure for the transition to high-k dielectrics. Direct contact between silicon and high-k materials like HfO_2 creates problems of high trap density and low mobility, while thin SiON layers act as buffer layers at this interface, maintaining stability. Even today, many high-k gate stack architectures utilize 0.3–0.6 nm thick SiON or SiO_2 interface layers. Consequently, SiON and nitride-derived dielectrics have mitigated the limitations of conventional SiO_2 at a critical stage of MOSFET scaling and played a significant role as an intermediate generation in the transition to high-k/metal gate technologies. This transition has enabled the simultaneous evolution of materials science, surface chemistry, and manufacturing processes in industry.

The Necessity of Transitioning to High-k Dielectrics

The development of MOSFET technology clearly revealed that gate dielectrics reached their physical limits by the 2000s. While thinning classical dielectric materials such as SiO_2 and SiON provided advantages for a long time in terms of strengthening gate control and maintaining stable threshold voltage (V_TH), quantum mechanical tunneling leakage increased dramatically as the physical thickness fell below 1.2 nm. This not only led to increased power consumption but also created an unacceptable problem in terms of device reliability (Wong & Iwai, 2006). Therefore, redefining the dielectric material became necessary to sustain scaling. One of the main motivations for transitioning to high-k materials is the concept of EOT (Equivalent Oxide Thickness). EOT expresses the high capacitance provided by high-k dielectrics in terms of SiO_2. Maintaining a high gate capacitance in MOSFETs is critical for strengthening electrostatic control across the channel. However, due to the relatively low dielectric constant ($k \approx 3.9$) of SiO_2, achieving the desired capacitance value around 1 nm is physically impossible. In contrast, with materials that have a high dielectric constant (e.g., HfO_2, $k \approx 20–25$), the EOT remains the same while the physical thickness can be increased; thus, tunneling current is reduced while gate control is preserved (Kingon et al., 2000). This property is one of the fundamental theoretical justifications for replacing SiO_2 with high-k materials. Another factor that necessitates high-k materials is power consumption and thermal management issues. In nanometer-scale devices, gate leakage can reach approximately 30% of the total power consumption of the entire chip; this directly impacts battery life in mobile processors and total energy costs in large-scale data centers. With the thinning of SiO_2 layers at 45 nm and below technology nodes, the gate leakage current has increased by the order of 6–9, turning into a serious efficiency problem that has challenged the industry (Bohr, 2007).

However, the transition to high-k dielectrics is not solely about physical thickness; it also means undergoing a profound structural change in the MOSFET architecture. Most high-k materials have higher interface trap density, lower band alignments, or thermal instability compared to SiO_2. Therefore, a true high-k solution has required not only the selection of a new dielectric but also the transition to metal gate technology (HKMG – High-k/Metal Gate), the redesign of the interface layer, and the optimization of manufacturing processes (especially ALD) (Robertson, 2006).

Therefore, the industry has worked on a wide list of candidates for high-k materials for many years; Oxides such as Al_2O_3, ZrO_2, TiO_2, and La_2O_3 were tested, but only a few materials approached industrial standards. Among these, hafnium dioxide (HfO_2) emerged as the most successful candidate due to both its thermal stability and suitable band alignment, and was finally officially introduced to the industry by Intel in 2007 using a 45 nm manufacturing process. Consequently, the transition to high-k dielectrics is not merely a material change, but a technological milestone that has enabled the semiconductor industry to overcome its scale limitations. This transition has ensured the continuity of modern CMOS architectures in terms of power efficiency, reliability, and performance.

The Development of High-k Materials and the Industrial Introduction of HfO_2

When the need to transition to high-k dielectrics became clear, the semiconductor industry accelerated the search for alternative materials that could provide higher dielectric constants while maintaining the unique interface quality offered by SiO_2. In this process, dozens of oxides, nitrides, and hybrid compounds were evaluated; candidate materials were compared based on criteria such as dielectric constant, band alignments, thermal stability, interface trap density, and application method (Robertson, 2006). However, only a few materials from this large pool of candidates have achieved true industrial maturity.

The main goal in selecting high-k materials is to keep the eot (electron leakage) low while increasing physical thickness, thus reducing leakage current. In this context, a high dielectric constant ($k > 15$) and suitable band offsets (≥ 1 eV) are critical design parameters. For example, TiO_2 ($k \approx 80$), despite offering a very high dielectric constant, is insufficient in terms of gate leakage due to its low band offsets. Oxides such as La_2O_3 and ZrO_2 have shown problems in terms of chemical stability and thermal resistance (D'Emic et al., 2006). Therefore, the industry has turned to hafnium-based oxides that are suitable for band alignment with silicon and offer stable physical properties. HfO_2 (hafnium dioxide) has emerged as the most promising material among these candidates. The main reasons for this are:

• Its dielectric constant $k \approx 20–25$, approximately 5 times higher than that of SiO_2.

• It is chemically stable with silicon at high temperatures.

• It reduces leakage currents by several steps.

• It provides excellent thickness control with ALD (Atomic Layer Deposition).

• Band offsets are at a sufficient level for both electrons and holes.

These properties have made HfO_2 ideal for the industry in terms of both performance and production compatibility. Indeed, initial studies have shown that HfO_2 can be used in place of SiO_2 in MOS gates and that mobility losses can be corrected through interface engineering (Wilk, Wallace & Anthony, 2001).

The critical breakthrough in integrating high-k materials occurred in 2007 with Intel's announcement of its 45 nm HKMG architecture. By using a metal gate (TiN-based) together with a high-k HfO_2 layer, Intel solved two fundamental problems:

1. Fermi-level pinning in the polysilicon gate,

2. Tunneling leakage, which is inevitable in thin SiO_2 layers.

This innovation was considered a major technological breakthrough in terms of reducing mobility losses and stabilizing the threshold voltage (Chau et al., 2007). Intel's HKMG integration became a turning point in the semiconductor industry and was soon adopted by manufacturers such as TSMC, IBM, and Samsung. The dielectric properties of HfO_2 stem from the diversity in its crystalline phase structures. While HfO_2 exists in a monoclinic phase at room temperature, it can be stabilized into tetragonal and cubic phases in thin-film form. Each of these phases offers different dielectric constant values, and high-k properties can be optimized with appropriate doping or heat treatment (Materlik et al., 2015). Furthermore, special phases of HfO_2 that can exhibit ferrolelectric properties have paved the way for next-generation gate stack designs in modern memory devices (FeFETs).

The design of the interface layer has also played a decisive role in the success of the HfO_2 architecture. Since it is not possible to completely remove SiO_2, a typical 0.3–0.6 nm SiO_2 or SiON layer is incorporated into the gate stack architecture to maintain interface quality. This structure both reduces

trap density and improves transport properties (Wong & Iwai, 2006). Consequently, the introduction of HfO_2-based high-k dielectrics to the industry is one of the most critical material innovations enabling the continuation of MOSFET technology following the end of SiO_2 scaling. The success of high-k dielectrics has not only reduced tunneling currents but has also standardized the HKMG design, shaping the future of chip architecture.

Transition to Metal Gate Technology (High-k / Metal Gate – HKMG)

The success of the transition to high-k dielectrics depends not only on the selection of a new insulating material but also on the complete redesign of the MOSFET's gate electrode architecture. Polysilicon gates, used in traditional MOSFETs, have been preferred for many years due to their ease of fabrication and compatibility with silicon processes. However, the use of polysilicon with high-k materials has begun to create physical and electrical problems that severely limit device performance (Hobbs et al., 2001). Therefore, the transition to metal gate technology in 45 nm and below processes has become imperative; this change has been one of the turning points in modern CMOS architecture. The biggest problem with polysilicon gates is the Fermi-level pinning phenomenon that occurs with high-k materials. When dense interface traps form at the interface between polysilicon and high-k oxides such as HfO_2, the gate's Fermi level cannot be set to the desired position, resulting in a loss of threshold voltage (V_TH) control (Shiraishi et al., 2003). This causes undesirable voltage drifts for NMOS and PMOS transistors and complicates device design. Another critical problem is the gate depletion that occurs when the polysilicon gate comes into contact with a high-k dielectric. Since polysilicon has a limited carrier density, a depletion layer forms under the gate, reducing the effective capacitance. This increases the EOT (Effective Transmission), partially negating the advantage provided by high-k technology (Mayer et al., 2003). Metal gates, on the other hand, do not form a depletion region because their free carrier density is much higher, thus maximizing gate capacitance. The transition process of metal gate technology is actually based on two main approaches:

(1) Gate-first architecture

In this approach, the metal gate is formed in the early stage along with a high-k dielectric. However, the gate-first method has been limited, especially in advanced nodes, because subsequent heat treatments lead to phase transformation, work function shifts, and diffusion effects in the metal gates.

(2) Gate-last (replacement gate) architecture

This is the method that is common in the industry today. In the first stage, a temporary polysilicon gate resistant to low temperatures is formed. At the end of the production process, the polysilicon is completely removed and replaced with a metal gate with a suitable work function. This approach ensures a more stable and controlled structure in the high-k + metal gate stack because it prevents heat processes from damaging the metal gate (Chau et al., 2007).

Common metal gate materials used in gate-stack architecture are:

• For NMOS: TiAlN, TiN, HfN

• For PMOS: TaN, Ru, Ir, WN, Pt

• General conductive filling: W (tungsten)

The choice of metal gate is based on work function engineering, which is one of the key factors determining the threshold voltage of the MOSFET. For example, the target work function for NMOS is approximately 4.1–4.3 eV, while the ideal range for PMOS is 4.9–5.1 eV (Oshima et al., 2005).

Therefore, the success of the HKMG architecture depends on the precise engineering optimization of the gate-stack structure.

With Intel's announcement of the HKMG system at 45 nm, the gate-depletion problem was eliminated, threshold voltage stability was increased, and gate leakage currents were reduced by multiple steps. This innovation also minimized interface disturbances that cause carrier mobility to decrease, significantly improving CMOS performance. In a short time, manufacturers such as Samsung, IBM, TSMC, and GlobalFoundries also incorporated HKMG into their standard process flows, and today all 14 nm, 7 nm, 5 nm, and 3 nm FinFET/GAAFET architectures are based on HKMG technology. HKMG is the result of a complete redefinition of not only the dielectric change but also the gate stack architecture. This structure has strengthened gate control in modern devices, reduced tunneling losses, and enabled the infrastructure of three-dimensional architectures such as FinFET/GAAFET.

Physical Problems Encountered in High-k Dielectrics

While the introduction of high-k dielectrics into MOSFET manufacturing has provided a significant advantage in terms of scaling down, it has also brought about a number of complex physical problems compared to SiO_2. These problems stem from the chemical properties of the material, its crystalline phase structure, interface characteristics, and manufacturing methods. Therefore, the integration of high-k insulators has become not only a material change but also a multi-layered optimization problem in gate stack engineering.

Interface Trapping and Transport Distortion

When high-k materials are in direct contact with Si, a flawless interface is not formed as in the Si–SiO_2 system. In oxides such as HfO_2 and ZrO_2, dense oxygen vacancies and metal-oxygen bond deficiencies increase the interface trap density. (Houssa et al., 2006). These traps:

• Cause carrier scattering, reducing mobility.

• It leads to threshold voltage (V_TH) instability.

• It impairs the long-term reliability of the device. Therefore, the necessity of a thin (0.3–0.6 nm) SiO_2/SiON interface layer in modern gate-stack structures has become one of the most critical design requirements of high-k technology.

Fixed Charges, Hysteresis, and V-TH Shifts

Another problem frequently observed in high-k dielectrics is the hysteresis behavior caused by fixed charges and polarized bonds within the structure. This phenomenon leads to:

• Transient threshold voltage shifts,

• Increased drain-induced barrier lowering (DIBL) effect,

• Deterioration of cyclic stability (Pei et al., 2008). Oxygen vacancies in HfO_2 films, especially since they lead to positive charge trapping, cause significant threshold shifts in PMOS devices.

Gate Leakage: Fowler–Nordheim and Direct Tunneling Mechanisms

Although high-k materials can be used in thick forms, gate leakage currents are still a limiting factor affecting device stability. Leakage current in HfO_2 sheets is generally explained by two mechanisms:

• Fowler–Nordheim tunneling (at high electric fields)

• Direct tunneling (in very thin interface layers)

Lower band alignments compared to SiO_2 can increase leakage current, especially in terms of electron tunneling (Khakifirooz & Antoniadis, 2006). Therefore, in modern gate-stacks, the choice of metal gate and interface thickness are precisely optimized.

Thermal Stability and Phase Transformations

The crystalline phase of HfO_2 films is one of the most important parameters determining their dielectric properties. At room temperature, the monoclinic phase predominates; however, during heat treatments in thin films, transformations to:

• tetragonal,

• cubitic,

• orthorhombic

phases can be observed. These phases have different dielectric constants. For example, the tetragonal phase has a higher k value. However, returning to the monoclinic phase at high temperatures can lead to undesirable EOT growth (Materlik et al., 2015). Therefore, heat treatment budget is one of the critical design constraints in high-k devices**.**

Oxygen Vacancies, Trap States, and Electrical Fatigue

Oxygen vacancies, common in high-k oxides like HfO_2, create disadvantages in terms of both conductivity and dielectric stability:

• They cause charge closure in the dielectric.

• They cause permanent degradation under electrical stress.

• They increase the Bias Temperature Instability (BTI) effect.

Negative Bias Temperature Instability (NBTI) in particular is one of the most important causes of performance losses in PMOS devices (Grasser et al., 2007).

High-k + Metal Gate Stack Integration Problems

The main integration challenges encountered in HKMG structures are:

• Reactions between the metal gate and high-k

• Maintaining work function stability

• Preventing metal diffusion

• Suitable doping for tetragonal phase stabilization (e.g., Al, Si, La doping)

If these problems are not solved, threshold voltage shifts, mobility degradation, and long-term instability occur.

Mobility Degradation

One of the most persistent problems of high-k dielectrics is the degradation of carrier mobility. The reasons for this degradation are:

• Polar phonon scattering

• Interfacial disturbances

• Coulomb scattering

• New scattering mechanisms starting with metal gates

Therefore, strain engineering, SiGe channels, and new interface designs with 2D materials have come to the forefront to maintain mobility in modern devices.

Next Generation Dielectrics and Industry Roadmap (IRDS – Beyond 1 nm)

The adoption of high-k dielectrics (especially HfO_2-based systems) in industry is part of a transformation that began with the 45 nm node. However, the downscaling process, progressing in line with Moore's Law, has pushed gate-stack design to a new limit at nodes such as 7 nm, 5 nm, and 3 nm. Today, IRDS (International Roadmap for Devices and Systems) reports indicate that sub-1 nm EOT, atomic-scale gate-stack engineering, 2D material-based MOSFETs, ferroelectric HfO_2 derivatives, and entirely new dielectric classes are on the agenda for the period after 2030. This transformation is shaped along three main axes:

High-k Materials After HfO_2

Although HfO_2-based dielectrics are currently the standard insulator for FinFET and GAAFET architectures, some limitations are becoming apparent in sub-1 nm processes:

• Interface quality not as good as SiO_2

• Charge trapping limiting mobility

• Phase transformations damaging stability

• EOT not being able to be reduced below 0.5 nm

Therefore, IRDS proposes new candidate dielectrics for production after 2030 (IRDS, 2023). These include:

• La-doped HfO_2 (for mobility and EOT optimization)

• ZrO_2–HfO_2 hybrids (due to high band offsets)

• Praseodymium oxide (Pr_2O_3)

• Gadolinium oxide (Gd_2O_3)

• Aluminum lanthanum oxide ($LaAlO_3$)

• Rare-earth-based oxides (Er_2O_3, Dy_2O_3, Y_2O_3)

The common goal of these oxides is to both create a compatible interface with Si and overcome the limited band alignment problem of HfO_2 (Robertson, 2015).

Ferroelectric HfO_2 and Nanoscale Memory + Logic Integration

In recent years, the ferroelectric behavior of the orthorhombic phase discovered in HfO_2-based dielectrics has taken gate-stack design to a completely new dimension. Ferroelectric HfO_2:

• Enables FeFET (Ferroelectric Field-Effect Transistor) memories,

• Allows "non-volatile logic" architecture in logic devices,

• Improves nanoscale energy efficiency. Therefore, IEDM sessions predict that FeFET-based L4/Cache memory units will enter commercial production within 5 years (IEDM 2022 Highlights).

This has a significant impact on gate-stack design:

In addition to the classical dielectric behavior of HfO_2, its switchable polarization property must now also be optimized. Thus, the dielectric becomes not only an insulator but also an active component that stores information.

2D Materials and Dielectrics Suitable for Van der Waals Interfaces

One of the most important candidates for sub-1 nm CMOS architecture is 2D semiconductor channels (MoS_2, WSe_2, WS_2, hBN). When these materials are in direct contact with high-k Si-based materials, problems such as:

• High interfacial trap density,

• Band alignment problems,

• Mobility reduction

are encountered (Fiori et al., 2014). Therefore, completely new types of dielectrics are needed for 2D devices:

(a) 2D–2D compatible dielectrics

• hBN (hexagonal boron nitride)

• AlN monolayer insulators

• 2D oxides (multilayer vdW dielectrics)

(b) ALD-to-2D nucleating layers

It is difficult to initiate ALD growth in 2D materials with inert surfaces such as MoS_2. Therefore:

• Ozone pre-treatment

• NO_2 functionalization

• Seed-layer engineering

Techniques such as these determine the quality of gate dielectrics (Sharma et al., 2019).

IRDS's 2035 prediction for 2D FETs:

EOT ≈ 0.4 nm and gate leakage $<10^{-3}$ A/cm².

These values are not possible with classical HfO_2; this necessitates new dielectric classes.

Atomic Layer Dielectrics and "Gate-Stack Layer-by-Layer" Architecture

ALD (Atomic Layer Deposition) enables atomic-level control of modern dielectric design. In sub-1 nm processes, gate-stack architecture is based on the following concepts:

• sub-nanometer interface layers

• atomic doping (Al, La, Si doping)

• laminated dielectric stacks (such as HfO_2/Al_2O_3)

• interfacial scavenging techniques

• oxygen vacancy engineering

In particular, "oxygen scavenging" processes can reduce the EOT value by 0.1–0.2 nm; this is a critical design method for sub-1 nm targets.

Conclusion and Evaluation

In this section, the strategic role of gate dielectric materials in the development of semiconductor technology is discussed within a historical framework, and the transformation process from classical SiO_2-based MOS architectures to high-k/metal gate (HKMG) structures is comprehensively evaluated. The literature and industry applications examined clearly demonstrate that dielectric material selection is not merely an insulator choice; it is a multi-dimensional engineering problem that directly determines device performance, power consumption, scalability, and reliability.

SiO_2, thanks to its low-defect interface with silicon, has formed the basis of MOSFET technology for decades; however, with the reduction of physical thickness to sub-nanometer scales, quantum tunneling effects have driven gate leakage currents to critical levels. This situation has shown that the scaling down goals pursued in line with Moore's Law cannot be continued with SiO_2. Although SiON and nitride-doped dielectrics have offered a temporary solution in this transition process, they have not been sufficient to overcome the EOT limits.

In this context, the introduction of high-k dielectrics, especially HfO_2-based systems, to industry represents a turning point for CMOS technology. HfO_2's high dielectric constant, thermal stability, and atomic-precision growth via ALD have provided critical advantages in reducing gate leakage and maintaining gate control. However, interface traps, mobility degradation, and threshold voltage instabilities arising from direct contact of high-k materials with Si have shown that this technology alone is insufficient. Therefore, the transition to metal gate technology has become imperative, and the HKMG architecture has formed the basis of modern FinFET and GAAFET structures. The study also emphasizes that the physical problems encountered in high-k dielectrics—interface trapping, static charges, BTI effects, phase transformations, and oxygen vacancies—are still active research topics today. These problems demonstrate that dielectric engineering now requires not only material selection but also an approach based on interface chemistry, atomic layer engineering, and gate-stack optimization. From a future perspective, sub-1 nm EOT targets in line with IRDS roadmaps necessitate new dielectric classes beyond classical HfO_2-based systems. Rare earth oxides, HfO_2–ZrO_2 hybrids, ferroelectric HfO_2 derivatives, and van der Waals dielectrics compatible with 2D materials are among the strongest candidates for the coming period. In particular, the fact that ferroelectric HfO_2 enables new device architectures that combine both logic and memory functions shows that dielectrics are moving from being a passive component to an active functional layer. Consequently, the evolution of gate dielectrics is one of the fundamental driving forces enabling the scalability of semiconductor technology. Success in future CMOS and beyond devices will depend not on high-k values, but on the simultaneous optimization of flawless interfaces, low trap density,

band alignment, and atomic-precision fabrication techniques. In this respect, dielectric engineering will continue to be one of the most critical research areas shaping the future of nanoelectronics.

REFERENCES

Bohr, M. (2007). The evolution of scaling from the homogeneous era to the heterogeneous era. *IEDM Technical Digest*, 1–6.

Chau, R., Datta, S., Doczy, M., Doyle, B., Jin, B., & Kavalieros, J. (2005). High-k/metal gate stack and its MOSFET characteristics. *IEEE Electron Device Letters, 25*(7), 408–410.

Chau, R., Doyle, B., Datta, S., Kavalieros, J., & Zhang, K. (2007). Integrated CMOS transistors using high-k gate dielectric and metal gate. *IEDM Technical Digest*, 1–4.

D'Emic, C., Cartier, E., & Buchanan, D. (2006). High-k dielectrics for future generation CMOS devices. *MRS Bulletin, 31*(6), 518–522.

Deal, B. E., & Grove, A. S. (1965). General relationship for the thermal oxidation of silicon. *Journal of Applied Physics, 36*(12), 3770–3778.

Fiori, G., et al. (2014). Electronics based on two-dimensional materials. *Nature Nanotechnology, 9*, 768–779.

Frank, D. J., Laux, S. E., & Wong, H. S. P. (2001). Generalized scale length model for long-channel MOSFETs. *IEEE Transactions on Electron Devices, 48*(11), 2120–2127.

Grasser, T., Kaczer, B., Roussel, P., et al. (2007). A unified model for BTI in high-k dielectrics. *IEDM Technical Digest*, 801–804.

Hobbs, C. C., Taylor, B., Grant, J. M., et al. (2001). Fermi-level pinning at the polysilicon/metal oxide interface. *IEEE Transactions on Electron Devices, 48*(5), 923–931.

Houssa, M., Pantisano, L., Ragnarsson, L. Å., et al. (2006). Electrical properties of high-k gate stacks. *Materials Science and Engineering: R, 51*(4–6), 37–85.

IEDM Technical Highlights. (2022–2023).

International Roadmap for Devices and Systems (IRDS). (2023). *More Moore Report*.

Khakifirooz, A., & Antoniadis, D. (2006). MOSFET performance scaling—Part II: Compact modeling of ground-plane and channel doping effects. *IEEE Transactions on Electron Devices, 53*(11), 2792–2798.

Kingon, A. I., Maria, J.-P., & Streiffer, S. K. (2000). Alternative dielectrics to silicon dioxide for memory and logic devices. *Nature, 406*(6799), 1032–1038.

Lombardi, C., Manzini, S., Saporito, A., & Vanzi, M. (1988). A physically based mobility model for numerical simulation of nonplanar devices. *IEEE Transactions on Computer-Aided Design, 7*(11), 1164–1171.

Materlik, R., Kersch, A., & Schroeder, U. (2015). Origin of the ferroelectric phase in HfO_2-based thin films. *Journal of Applied Physics, 117*, 134109.

Mayer, T., Goel, A., Ponoth, S., et al. (2003). Gate depletion and mobility considerations for polysilicon and metal gate MOSFETs. *IEEE Electron Device Letters, 24*(4), 242–244.

Nicollian, E. H., & Brews, J. R. (1982). *MOS (Metal Oxide Semiconductor) physics and technology*. Wiley.

Oshima, Y., Natori, K., & Kamimuta, Y. (2005). Work function engineering in metal gate MOSFETs. *Microelectronic Engineering, 80*, 256–259.

Pei, Y., Song, L., He, Y., et al. (2008). Charge trapping in HfO_2-based high-k gate stacks. *Applied Physics Letters, 92*(10), 102903.

Robertson, J. (2006). High dielectric constant oxides. *European Physical Journal – Applied Physics, 28*(3), 265–291.

Robertson, J. (2015). High dielectric constant oxides: Current status. *Journal of Applied Physics, 117*, 112806.

Sharma, A., et al. (2019). ALD dielectrics for 2D materials. *Applied Physics Reviews, 6*, 021311.

Shiraishi, K., Terao, T., Takagi, Y., et al. (2003). First-principles study of Fermi-level pinning at the high-k/Si interface. *Japanese Journal of Applied Physics, 42*, 4660–4665.

WIlk, G. D., Wallace, R. M., & Anthony, J. M. (2001). High-k gate dielectrics: Current status and materials properties considerations. *Journal of Applied Physics, 89*(10), 5243–5275.

Wong, H.-S. P., & Iwai, H. (2006). On the scaling issues and high-k replacement of ultra-thin gate dielectrics for nanoscale MOS transistors. *Microelectronics Engineering, 83*(10), 1867–1904.

Chapter 2

PARA*f* (3,2,1)-STRUCTURES ON MANIFOLDS[1]

Mustafa GÖK[2]

1 This book chapter is produced from the author's doctoral thesis.

2 Associate Professor, National Defence University, Turkish Military Academy, 06420 Ankara, Turkey, Orcid: 0000-0001-6346-0758, mustafa.gok@email.com

Introduction

A pseudo f-structure (Matsumoto, 1976) on a differentiable manifold M is given by a non-zero tensor field φ of type $(1,1)$ of constant rank k satisfying $\varphi^3 - \varphi = 0$. Bucki (Bucki, 1989) called it a para f-structure. One of the well known examples of para f-structures is almost product structures. Explicity, an almost product structure φ is a para f-structure of rank $k = \dim M$. Bucki (Bucki, 2003) defined the integrability of a para f-structure and showed that the integrability condition of such a structure is equivalent to the vanishing of its Nijenhuis tensor.

On the other hand, a tensor field Φ of type $(1,1)$ on a differentiable manifold M with the property $\Phi^2 = \Phi + I$ is referred to as a golden structure by Crâşmăreanu and Hreţcanu (Crâşmăreanu & Hreţcanu, 2008), where I is the identity map on M. There is a close relationship between golden structures and almost product structures. In spite of such a relationship, golden structures are not a class of para f-structures.

To fill the above-mentioned gap in the literature, this paper is devoted to a polynomial structure of degree 3 on differentiable manifolds, called a para $f(3,2,1)$-structure, which contains the class of golden structures. Initially, the relationship between para $f(3,2,1)$-structures and golden structures is established. Later, the basic features of a para $f(3,2,1)$-structure and the integrability of its associated distributions are examined. Afterwards, the concepts of partially integrability and integrability of para $f(3,2,1)$-structures are defined and investigated.

Main Results

In this section, we introduce the notion of a para $f(3,2,1)$-structure on differentiable manifolds and study some of its fundamental properties.

Definition 1. A non-zero tensor field F of type $(1,1)$ on an m-dimensional differentiable manifold M is named a para $f(3,2,1)$-structure if it verifies the equation

$$F^3 = F^2 + F. \tag{1}$$

In this situation, the pair (M, F) is called a para $f(3,2,1)$-manifold.

Now, we present a trivial example of para $f(3,2,1)$-structures.

Example 1. A golden structure on a differentiable manifold is a para $f(3,2,1)$-structure.

Proposition 1. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. The characteristic values of F are $1 - \phi$, 0 and ϕ, where ϕ is the golden ratio, i.e., $\phi = \frac{1+\sqrt{5}}{2}$.

Proof. If λ is a characteristic value of F, then we have $FU = \lambda U$ for all $U \in \Gamma(TM)$. Thus, from (1), we get

$$\lambda^3 - \lambda^2 - \lambda = 0, \tag{2}$$

which implies that the characteristic values of F are $\frac{1-\sqrt{5}}{2}$, 0 and $\frac{1+\sqrt{5}}{2}$. This completes the proof.

We consider a para $f(3,2,1)$-structure F on a differentiable manifold M. Let $\ker F$ and $\mathrm{Im} F$ be the kernel and image of F, respectively. That is,

$$\ker F = \bigcup_{x \in M} (\ker F)_x \tag{3}$$

and

$$\mathrm{Im} F = \bigcup_{x \in M} (\mathrm{Im} F)_x, \tag{4}$$

where

$$(\ker F)_x = \{U_x \in T_x M : F_x U_x = 0\} \tag{5}$$

and

$$(\mathrm{Im} F)_x = \{V_x \in T_x M : V_x = F_x U_x \text{ for some } U_x \in T_x M\} \tag{6}$$

at each point $x \in M$.

Proposition 2. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then the following statement holds:

$$ker F \cap Im F = \{0\}. \tag{7}$$

Proof. We assume that $U \in \Gamma(ker F \cap Im F)$. Then we get $FU = 0$. At the same time, U is written in the form

$$U = FV, \tag{8}$$

where V is a vector field on M. Hence, we find

$$F^3 V = F^2 V = 0. \tag{9}$$

From (1), (8) and (9), it follows that

$$U = 0, \tag{10}$$

so we derive (7).

Remark 1. According to Proposition 2, for a para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M, we have

$$TM = ker F \oplus Im F. \tag{11}$$

Proposition 3. A para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M is a golden structure if only one of the following statements is satisfied:

(a) $(\ker F)_x = \{0\}$ for each point $x \in M$,
(b) $rank F = m$,
(c) F is injective.

Proof. We assume that (a) or (b) holds. Then F is injective, so there is the inverse of F, denoted by F^{-1}. Applying F^{-1} from the left hand side to (1), we get

$$F^2 - F - I = 0, \tag{12}$$

that is, F is a golden structure. Thus, the proof is done.

Example 2. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Define a tensor field by $\Phi = F_{|\mathrm{Im}F}$. If $U \in \Gamma(\ker\Phi)$, then we have $U \in \Gamma(kerF \cap ImF)$. Thus, by Proposition 2, we obtain $U = 0$. As a result, Proposition 3 implies that Φ is a golden structure.

Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. We consider two operators defined by the rules

$$p = F^2 - F \tag{13}$$

and

$$q = -F^2 + F + I, \tag{14}$$

respectively, where I stands for the identity map on M.

Theorem 1. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then p and q are complementary projection operators with the properties

$$p + q = I, \tag{15}$$
$$p^2 = p, \tag{16}$$
$$q^2 = q \tag{17}$$

and

$$pq = qp = 0. \tag{18}$$

Proof. By means of (1), the proof is a straightforward consequence of (13) and (14).

We denote by P and Q the distributions with regard to the projection operators p and q, respectively. Then we have

$$TM = P \oplus Q. \tag{19}$$

Proposition 4. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then the following statements hold:

$$P = \mathrm{Im}F = \ker q \tag{20}$$

and

$$Q = \ker F = \ker p. \tag{21}$$

So we have

$$Fp = pF = F, \tag{22}$$
$$(F^2 - F)p = p, \tag{23}$$
$$qF = Fq = 0 \tag{24}$$

and

$$(-F^2 + F + I)q = q. \tag{25}$$

Proof. Taking into consideration (1), the proof is a direct consequence of (13) and (14).

Proposition 5. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then the following statements are valid:

(a) F acts on the distribution P as a golden structure,

(b) F acts on the distribution Q as a null operator.

Proof. Using (23) and (24), respectively, the proofs of (a) and (b) are obtained.

Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M and we consider its Nijenhuis tensor N_F defined by

$$N_F(U,V) = [FU,FV] - F[FU,V] - F[U,FV] + F^2[U,V] \tag{26}$$

for all $U,V \in \Gamma(TM)$. Thus, by means of (23), we get

$$N_F(U,V) = [FU,FV] - F[FU,V] - F[U,FV] + F[U,V] + p[U,V]. \tag{27}$$

Proposition 6. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then we have

$$qN_F(U,V) = qN_F(pU,pV) = q[FU,FV] \tag{28}$$

for all $U,V \in \Gamma(TM)$.

Proof. From (18), (24) and (27), we obtain

$$qN_F(U,V) = q[FU,FV] \tag{29}$$

for all $U,V \in \Gamma(TM)$. Putting pU and pV into U and V in (29), respectively, we infer from (22) that

$$qN_F(pU,pV) = q[FU,FV]. \tag{30}$$

Thus, (29) and (30) together imply that (28).

Theorem 2. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. The distribution P is integrable if and only if

$$qN_F(pU,pV) = 0 \tag{31}$$

for all $U,V \in \Gamma(TM)$.

Proof. It is well known that P is integrable if and only if

$$q[pU,pV] = 0 \tag{32}$$

for all $U,V \in \Gamma(TM)$. Also, because of (22), (32) is equivalent to the condition

$$q[FU,FV] = 0. \tag{33}$$

Thus, using (28), we get (31). Conversely, we assume that (31) holds. It follows from (28) that

$$q[FU,FV] = 0. \tag{34}$$

In (34), putting $(F-I)U$ and $(F-I)V$ into U and V, respectively, we obtain from (22) and (23) that

$$q[pU,pV] = 0, \tag{35}$$

in other words, P is integrable.

Proposition 7. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then the following statements are valid:

$$N_F(qU,qV) = pN_F(qU,qV) = F[qU,qV] + p[qU,qV] \tag{36}$$

and

$$(2I - F)N_F(qU, qV) = p[qU, qV] \tag{37}$$

for all $U, V \in \Gamma(TM)$.

Proof. Considering (27), from (16), (22) and (24), we have

$$N_F(qU, qV) = F[qU, qV] + p[qU, qV] \tag{38}$$

and

$$pN_F(qU, qV) = F[qU, qV] + p[qU, qV] \tag{39}$$

for all $U, V \in \Gamma(TM)$. Hence, by (38) and (39), we obtain (36). In addition, using (22) and (23), obviously, (37) follows from (36) .

Theorem 3. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. The distribution Q is integrable if and only if

$$N_F(qU, qV) = 0 \text{ or } pN_F(qU, qV) = 0 \tag{40}$$

for all $U, V \in \Gamma(TM)$.

Proof. Combining (22) with (36), we get

$$N_F(qU, qV) = Fp[qU, qV] + p[qU, qV] \tag{41}$$

for all $U, V \in \Gamma(TM)$. Moreover, we remember that the integrability of Q is equivalent to the condition

$$p[qU, qV] = 0 \tag{42}$$

for all $U, V \in \Gamma(TM)$. We suppose that Q is integrable. Then (41) implies that we have (40). Conversely, if one of the relations in (40) is valid, then from (37), we obtain

$$p[qU, qV] = 0, \tag{43}$$

so Q is integrable.

Theorem 4. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Both the distributions P and Q are integrable if and only if the Nijenhuis tensor N_F of F is given by the form

$$N_F(U, V) = pN_F(pU, pV) + N_F(pU, qV) + N_F(qU, pV) \tag{44}$$

for all $U, V \in \Gamma(TM)$.

Proof. We assume that P and Q are both integrable. In view of (15), from Theorems 2 and 3, a direct calculation gives that N_F has the form in (44). The converse follows from that the intersection of P and Q is zero.

We suppose that P is integrable. Let M^P be an arbitrary integral manifold of P. We consider an operator F^P defined by

$$F^P U^P = FU^P \tag{45}$$

for all $U^P \in \Gamma(TM^P)$. In this situation, F^P acts on the tangent bundle TM^P as an invariant operator. Besides, as is seen, F^P is a golden structure on M^P. Denoting by N^P the Nijenhuis tensor of F^P induced from F on M^P, we have

$$N^P(pU, pV) = N_F(pU, pV) \tag{46}$$

for all $U, V \in \Gamma(TM)$.

Now, we define the partial integrability of a para $f(3,2,1)$-structure.

Definition 2. A para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M is said to be partially integrable if the following statements are valid:

(a) The distribution P is integrable,
(b) The golden structure F^P induced from F on the integral manifold M^P is integrable.

Proposition 8. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then the following statements are equivalent:

(a) $N_F(pU, pV) = 0$ for all $U, V \in \Gamma(TM)$,
(b) $N_F(FU, FV) = 0$ for all $U, V \in \Gamma(TM)$.

Proof. Considering (27), the proof is evident from (22) and (23).

Theorem 5. A para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M is partially integrable if and only if

$$N_F(pU, pV) = 0, \text{or equivalently } N_F(FU, FV) = 0 \tag{47}$$

for all $U, V \in \Gamma(TM)$.

Proof. As is known, F^P is integrable if and only if

$$N^P(pU, pV) = 0 \tag{48}$$

for all $U, V \in \Gamma(TM)$. Hence, taking account of (46), the proof is derived from Theorem 2 and Proposition 8.

Proposition 9. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. The distribution Q is integrable and F is partially integrable if and only if

$$N_F(U, V) = N_F(pU, qV) + N_F(qU, pV) \tag{49}$$

for all $U, V \in \Gamma(TM)$.

Proof. The proof is deduced immediately from Theorems 3, 5 and Proposition 8.

The Lie derivative of a para $f(3,2,1)$-structure F on a differentiable manifold M in regard to an arbitrary vector field $V \in \Gamma(TM)$ is given by

$$(\mathcal{L}_V F)(U) = F[U, V] - [FU, V] \tag{50}$$

for all $U \in \Gamma(TM)$.

Proposition 10. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Then we have

$$N_F(pU, qV) = F\big(\mathcal{L}_{qV}F\big)(pU) \tag{51}$$

and

$$(F - I)N_F(pU, qV) = p\big(\mathcal{L}_{qV}F\big)(pU) \tag{52}$$

for all $U, V \in \Gamma(TM)$.

Proof. Applying F from the left hand side to (50), we get

$$F\big(\mathcal{L}_{qV}F\big)(pU) = F^2[pU, qV] - F[FU, qV] \tag{53}$$

for all $U, V \in \Gamma(TM)$. On the other hand, substituting pU and qV for U and V in (26), respectively, it is derived from (22) and (24) that

$$N_F(pU, qV) = F^2[pU, qV] - F[FU, qV]. \tag{54}$$

Hence, in view of (53) and (54), we obtain (51). Also, from (22) and (23), we have (52).

We assume that P and Q are both integrable. Then there is a local coordinate system $\{(\Omega, \varphi): (x^i, y^\alpha)\}$ on M that verifying the following properties:

(a) The integral manifolds of P are represented by $l = m - k$ local coordinates

$$y^{k+1} = \text{constant}, \dots, y^m = \text{constant}, \tag{55}$$

(b) The integral manifolds of Q are represented by other k local coordinates

$$x^1 = \text{constant}, \dots, x^k = \text{constant}. \tag{56}$$

Such a coordinate system $\{(\Omega, \varphi): (x^i, y^\alpha)\}$ is said to be an adapted coordinate system (Yano & Kon, 1983).

Proposition 11. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Assume that the distributions P and Q are both integrable. The local components of F are independent of the coordinates which are constant along the integral manifolds of P in an adapted coordinate system if and only if

$$p\big(\mathcal{L}_{qV}F\big)p = 0 \text{ or } \mathcal{L}_{qV}F = 0 \tag{57}$$

for all $U, V \in \Gamma(TM)$.

Proof. In an adapted coordinate system $\{(\Omega, \varphi): (x^i, y^\alpha)\}$, by virtue of the fact that

$$\frac{\partial}{\partial x^i} \in \Gamma(P) \tag{58}$$

and

$$\frac{\partial}{\partial y^\alpha} \in \Gamma(Q), \tag{59}$$

F has the matrix form

$$F = \begin{pmatrix} F_j^i & 0 \\ 0 & F_\beta^\alpha \end{pmatrix}. \tag{60}$$

That is, F is written in the form

$$F = F_j^i \frac{\partial}{\partial x^i} \otimes dx^j + F_\beta^\alpha \frac{\partial}{\partial y^\alpha} \otimes dy^\beta. \tag{61}$$

Furthermore, p and q have the components

$$p = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \text{ or } p = \frac{\partial}{\partial x^i} \otimes dx^i \tag{62}$$

and

$$q = \begin{pmatrix} 0 & 0 \\ 0 & I_l \end{pmatrix} \text{ or } q = \frac{\partial}{\partial y^\alpha} \otimes dy^\alpha, \tag{63}$$

respectively. Hence, we obtain

$$\mathcal{L}_{\frac{\partial}{\partial y^\gamma}} F = \frac{\partial F_j^i}{\partial y^\gamma} \frac{\partial}{\partial x^i} \otimes dx^j + \frac{\partial F_\beta^\alpha}{\partial y^\gamma} \frac{\partial}{\partial y^\alpha} \otimes dy^\beta, \tag{64}$$

which implies that

$$p\left(\mathcal{L}_{\frac{\partial}{\partial y^\gamma}} F\right)p = \frac{\partial F_j^i}{\partial y^\gamma} \frac{\partial}{\partial x^i} \otimes dx^j. \tag{65}$$

At the same time, from (22) and (24), it becomes evident that $F_\beta^\alpha = 0$. Thus, F has the form

$$F = \begin{pmatrix} F_j^i & 0 \\ 0 & 0 \end{pmatrix} \text{ or } F = F_j^i \frac{\partial}{\partial x^i} \otimes dx^j. \tag{66}$$

Moreover, (64) takes the form

$$\mathcal{L}_{\frac{\partial}{\partial y^\gamma}} F = \frac{\partial F_j^i}{\partial y^\gamma} \frac{\partial}{\partial x^i} \otimes dx^j = p\left(\mathcal{L}_{\frac{\partial}{\partial y^\gamma}} F\right)p. \tag{67}$$

For this reason, by means of (65) and (67), if $p(\mathcal{L}_{qV}F)p = 0$ or $\mathcal{L}_{qV}F = 0$, then we get $\frac{\partial F_j^i}{\partial y^\gamma} = 0$, which says that the local components of F are independent of the coordinates which are constant along the integral manifolds of P. The converse follows directly from (67).

Proposition 12. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Suppose that the distributions P and Q are both integrable. Then the local components of F are independent of the coordinates which are constant along the integral manifolds of P in an adapted coordinate system if and only if

$$N_F(pU, qV) = 0 \tag{68}$$

for all $U, V \in \Gamma(TM)$.

Proof. Taking account of (51) and (52), it follows that the following statements are equivalent to each other:

$$p(\mathcal{L}_{qV}F)p = 0 \tag{69}$$

and

$$N_F(pU, qV) = 0 \tag{70}$$

for all $U, V \in \Gamma(TM)$. Hence, the proof is obvious from Proposition 11.

Proposition 13. Let F be a para $f(3,2,1)$-structure on an m-dimensional differentiable manifold M. Assume that the distributions P and Q are both integrable. Then the local components of F are independent of the coordinates which are constant along the integral manifolds of P in an adapted coordinate system if and only if

$$N_F(U,V) = pN_F(pU,pV) \tag{71}$$

for all $U, V \in \Gamma(TM)$.

Proof. Because P and Q are both integrable, N_F is expressed by

$$N_F(U,V) = pN_F(pU,pV) + N_F(pU,qV) + N_F(qU,pV) \tag{72}$$

for all $U, V \in \Gamma(TM)$. We suppose that local components of F are independent of the coordinates which are constant along the integral manifolds of P. In this case, based on Proposition 12, (72) gives rise to (71). Conversely, we assume that (71) is satisfied. Substituting pU and qV for U and V in (71), respectively, we obtain

$$N_F(pU,qV) = 0. \tag{73}$$

Thus, from Proposition 12, the local components of F are independent of the coordinates which are constant along the integral manifolds of P.

Now, we introduce the definition of the integrability of a para $f(3,2,1)$-structure.

Definition 3. A para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M is integrable if the following statements hold:

(a) F is partially integrable,
(b) The distribution Q is integrable,
(c) The local components of F are independent of the coordinates which are constant along the integral manifolds of the distribution P in an adapted coordinate system.

Theorem 6. A para $f(3,2,1)$-structure F on an m-dimensional differentiable manifold M is integrable if and only if $N_F = 0$.

Proof. The proof follows from Theorems 3, 5 and Proposition 13.

Theorem 7. A para $f(3,2,1)$-structure F of rank k on an m-dimensional differentiable manifold M is integrable if and only if there is a coordinate system in which F has the constant components

$$F = \begin{bmatrix} \phi I_a & 0 & 0 \\ 0 & (1-\phi) I_b & 0 \\ 0 & 0 & 0 \end{bmatrix}, a + b = k, \tag{74}$$

where I_a and I_b are the identity matrices of types $a \times a$ and $b \times b$, respectively.

Proof. If F is integrable, then its components in an adapted coordinate system are given by

$$F = \begin{bmatrix} F^P & 0 \\ 0 & 0 \end{bmatrix}, \tag{75}$$

where F^P stands for a matrix of type $k \times k$ and its elements are independent of the coordinates which are constant along the integral manifolds of P. Because F^P is a golden structure on an arbitrary integral manifold of P, it is possible to make a change of the adapted coordinate system such that

$$F^P = \begin{bmatrix} \phi I_a & 0 \\ 0 & (1-\phi) I_b \end{bmatrix}. \tag{76}$$

Therefore, we get

$$F = \begin{bmatrix} \phi I_a & 0 & 0 \\ 0 & (1-\phi) I_b & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{77}$$

Conversely, we suppose that there is a coordinate system in which F has the constant components

$$F = \begin{bmatrix} \phi I_a & 0 & 0 \\ 0 & (1-\phi) I_b & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{78}$$

Because of (13) and (14), p and q have the following components:

$$p = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \tag{79}$$

and

$$q = \begin{bmatrix} 0 & 0 \\ 0 & I_l \end{bmatrix}, \tag{80}$$

respectively. Hence, P and Q are both integrable. At the same time, as is clear, the local components of F are independent of the coordinates which are constant along the integral manifolds of P. Also, F^P is given by

$$F^P = \begin{bmatrix} \phi I_a & 0 \\ 0 & (1-\phi) I_b \end{bmatrix}, \tag{81}$$

i.e., it is a linear combination of tensor products $\frac{\partial}{\partial x^i} \otimes dx^j$ and $\frac{\partial}{\partial y^\alpha} \otimes dy^\beta$ with constant coefficients. In this case, we have $N^P = 0$, so F^P is integrable. Moreover, since P is integrable, F is partially integrable by Definition 2. Thus, all conditions of Definition 3 are satisfied. Consequently, the proof is complete.

References

Bucki, A. (1989). Para-φ-structures with parallelizable kernel on manifolds. *Tensor (N.S.), 48*(1), 36-45.

Bucki, A. (2003). Para-f-Lie groups. *International Journal of Mathematics and Mathematical Sciences, 49*, 3149-3152.

Bucki, A., & Miernowski, A. (1985). Almost r-paracontact structures. *Annales Universitatis Mariae Curie-Skłodowska. Sectio A. Mathematica, 39*, 13-26.

Crâşmăreanu, M. C., & Hreţcanu, C. E. (2008). Golden differential geometry. *Chaos, Solitons & Fractals, 38*(5), 1229-1238.

Matsumoto, K. (1976). On a structure defined by a tensor field f of type $(1,1)$ satisfying $f^3 - f = 0$. *Bulletin of the Yamagata University. Natural Science, 9*(1), 33-46.

Yano, K., & Kon, M. (1983). *CR Submanifolds of Kaehlerian and Sasakian Manifolds.* Boston: Birkhäuser.

Chapter 3

BIAS-VARIANCE TRADEOFF IN MACHINE LEARNING MODELS: A COMPARATIVE ANALYSIS

Pelin AKIN[1]

1 Dr.Öğr.Üyesi Pelin AKIN, Çankırı Karatekin Üniversitesi, Fen Fakültesi, İstatistik Bölümü, pelinakin@karatekin.edu.tr, ORCID: 0000-0003-3798-4827.

INTRODUCTION

Today, machine learning has become an indispensable component of data analytics processes. Machine learning is widely used to develop prediction and decision support systems in many fields such as healthcare, finance, education, and engineering (Jordan & Mitchell, 2015). When evaluating the performance of these models, prediction accuracy is an important criterion and is expressed by error terms in regression models. When selecting the best prediction model, consideration should be given not only to how the model generalizes to the training data but also to how it generalizes to new data (Hastie et al., 2009). At this point, two fundamental concepts that affect model performance come to the fore: bias and variance.

Bias is an error component arising from a model's inability to adequately learn the true relationship (Geman, Bienenstock & Doursat, 1992). In other words, it is the failure to capture important patterns in the data due to the simplicity of the model's structure. High bias causes the model to overlook fundamental relationships between features in the training data and the target variable, ultimately leading to underfitting (Bishop & Nasrabadi, 2006). Simple models such as linear regression may be inadequate for learning complex and nonlinear relationships (James et al., 2013). In such cases, errors remain high in both the training and test data. Variance is the error component resulting from a model being overly sensitive to the data it was trained on. A model with high variance overfits the training data by learning even random noise in it. In this case, the model performs perfectly with very low error on the training data but fails on new, unseen test data because the details it learned are not generalizable. Complex models (deep neural networks, decision trees, etc.) in particular have high variance. They can fit the training data almost perfectly, but because they also memorize the data noise, errors increase on the test data.

The total overall error of a model is a combination of bias and variance. Underfitting describes a situation of high bias and low variance; the model fails on both training and test data because its learning capacity is insufficient. Overfitting, on the other hand, is a situation of low bias and high variance; the model memorizes the training data, resulting in low training error but high test error. Ideally, both bias and variance are kept at reasonable levels, and a model of moderate complexity is selected, thus balancing training and test performance.

Recent theoretical studies have shown that highly parametric models behave beyond the classical bias-variance framework. For example, Rocks and Mehta (2022) performed a bias-variance decomposition on two simple models (linear regression and a two-layer neural network). In this study, a phase transition was observed where, as model complexity increased, the training error decreased to zero and the test error increased due to variance; however, beyond this threshold, both bias and variance decreased, and the test error began to decline. Thus, contrary to expectations, increasing capacity in overly parametric neural networks can improve the overall error rate by reducing both bias and variance.

Liu and Mazumder (2025) present an important result in terms of ensemble methods. In their case study on random forests, they show that forests, which are generally considered to only reduce variance, can actually reduce bias as well. They defined the underlying mechanism as “hidden patterns” and emphasized that forests, which capture patterns missed by the bagging method through random feature selection, can simultaneously reduce both variance and bias.

Similar results have been obtained in applied studies. In an experimental study conducted on a health insurance cost dataset, the performance of various regression models such as linear regression, SVR, Ridge, Random Forest, and Neural Network was compared. A finely tuned Random Forest model was found to have the lowest training error; however, it was determined that this model

had a higher test error than the other models, meaning it overfitted (Jui,Rivas,2024) . In contrast, both the training and test errors of the linear regression and SVR models remained high, and no significant difference was observed between these errors; this situation indicates that these models are high bias and low variance, and therefore prone to underfitting.

Yao and colleagues (2023) compared XGBoost, Random Forest, and SVM models using US health insurance premium data. The results showed that Random Forest had the lowest average absolute error (MAE) (≈2535); XGBoost had a slightly higher error (≈2789), while SVM had a much higher error (≈6446) The authors emphasized that the Random Forest model provided the most accurate results. The SVM model, with its high error, underfitted by failing to capture the complex relationships in the data (high bias and low variance).

This section discusses the concepts of bias, variance, underfitting, and overfitting; bias-variance balance is examined comparatively on linear regression, support vector regression (SVR), and Random Forest models. Experimental analyses were performed in R using the Insurance Charges (Medical Cost Personal) dataset. The performance of each model was measured using error metrics (RMSE, MAE) and interpreted in terms of bias-variance.

INSURANCE CHARGES DATASET

This study uses the Insurance Charges (Medical Cost Personal) dataset, a publicly available dataset containing individuals' demographic and health characteristics along with their annual health insurance costs (Mirichoi, 2018). The dataset consists of 1,338 observations and includes variables such as age, gender, body mass index (BMI), number of children, smoking status, and region. In the dataset, the independent variables (age, sex, BMI, children, smoker, region) and the dependent variable charges are used.

There are no missing values in the dataset; moreover, only one duplicate row was detected and cleaned when necessary. Therefore, the data is ready to be fed into the models. Categorical variables (sex, smoker, region) were converted to factor type in the R environment for proper processing during the modeling phase (Figure 1) . Below is the code for loading and preprocessing the dataset in R:

```
install.packages("readr")

url    <-    "https://raw.githubusercontent.com/stedy/Machine-Learning-with-R-
datasets/master/insurance.csv"

insurance <- read_csv(url)

insurance$sex<-as.factor(insurance$sex)

insurance$smoker<-as.factor(insurance$smoker)

insurance$region<-as.factor(insurance$region)

summary(insurance)
```

Figure 1. Loading the insurance cost dataset into the R

In Figure 2, the dataset is split into two parts: 80% for training and 20% for testing.

```
set.seed(42)
train_index <- sample(1:nrow(insurance), 0.8 * nrow(insurance))
train_data <- insurance[train_index, ]
test_data  <- insurance[-train_index, ]
```

Figure 2. Creation of training and test subsets (80% training, 20% test).

Let's briefly explain the structure of the models in order, then discuss their application with R and the results.

LINEAR REGRESSION

Linear regression is a fundamental method that models the relationship between a dependent variable and one or more independent variables using a linear equation (Hastie et al., 2009). The multiple linear regression model,

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

Here, x_i represents the independent variables and $\beta_i$represents the coefficients that the model attempts to learn. The coefficients are typically trained using the Ordinary Least Squares method; this method finds the β values that minimize the mean square error (MSE) between the actual values and the predicted values.
The linear regression model was applied to the Insurance dataset. In R, multiple linear regression can be set up using the lm() function.

```
lm_model <- lm(charges ~ age + sex + bmi + children + smoker + region, data = train_data)
summary(lm_model)
```

Figure 3. Setting up a multiple linear regression model in R (using the lm() function).

Table 1 shows that, according to the linear regression model output, the model explains approximately 74.9% of the total variance in the model charges (insurance fee) variable ($R^2 = 0.7489$). This indicates that the model has a very strong explanatory power. The overall significance of the model is confirmed by the F-statistic ($F = 395.6$, $p < 0.001$), indicating that at least one independent variable has a statistically significant effect on the target variable.

Among the independent variables in the model, age, body mass index (BMI), number of children, and smoking show a statistically significant positive effect on the charges variable ($p < 0.05$). Smoking, in particular, stands out as the strongest determinant variable, with an additional cost of approximately $23,771. The age variable causes an average increase of $261 in charges for each year of age. Similarly, it is observed that charges increase as BMI increases, and this relationship is significant. An increase in the number of children is also associated with higher charges.

Although the gender variable (being male) is included in the model, it does not show a significant effect ($p > 0.5$). Similarly, only some of the region variables (e.g., southeast and southwest regions) were found to be statistically significant, but their effects are relatively low and negative. The model's average prediction error (residual standard error) is approximately $6,100, indicating the presence of some individual prediction deviations.

Table 1. Coefficient estimates and statistical test results for the multiple linear regression model predicting insurance premiums.

Variable	**Estimate**	**Std. Error**	**t value**	**p-value**
Intercept	-12386.93	1122.34	-11.037	< 0.001
Age	261.04	13.47	19.378	< 0.001
Sex (Male)	-230.54	375.70	-0.614	0.540
BMI	353.48	32.41	10.906	< 0.001
Children	516.05	154.12	3.348	0.001
Smoker (Yes)	23771.14	465.43	51.074	< 0.001
Region: Northwest	-544.63	537.05	-1.014	0.311
Region: Southeast	-1274.10	541.39	-2.353	0.019
Region: Southwest	-1325.16	537.80	-2.464	0.014

After training the model, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used to measure its performance. The following code calculates the RMSE and MAE values of the linear model on the test data:

```
pred_lm <- predict(lm_model, newdata = test_data)
rmse_lm <- sqrt(mean((pred_lm - test_data$charges)^2))
mae_lm  <- mean(abs(pred_lm - test_data$charges))
rmse_lm; mae_lm
```

Figure 4. Calculation of the performance metrics (RMSE, MAE) of the linear regression model

As a result of this calculation, the error of the linear regression model on the test data was found to be RMSE = $5,926.983 and MAE = $4,148.585. The errors on the training data were RMSE = $6,073.812 and MAE = $4,195.38. It can be seen that the training and test errors are quite close to each other. However, the relatively high errors in both data sets indicate that the model does not represent the data well enough. This situation points to an underfitting problem, which stems from the linear regression model's inability to capture the non-linear relationships in the insurance cost data due to its simple structure. On the other hand, the similar magnitude of training and testing errors indicates that the model does not overfit and that the variance problem is low. In this context, when the error components of the model are examined, it can be said that the bias error is dominant, while the variance error is relatively low.

SUPPORT VECTOR REGRESSION

Support Vector Regression (SVR) is an adaptation of support vector machines for regression problems. Unlike linear regression, SVR aims to keep all observations within a specified error band defined by a certain epsilon (ε) error margin. The model ignores errors within the ε band and only penalizes points outside this band. Thus, a more flexible regression line/hyperplane is found within a certain tolerance. In short, SVR is a regression technique that defines how much error is acceptable in the model; it places the most suitable linear line or, in the non-linear case, hyperplane according to the specified error tolerance (Smola & Schölkopf, 2004; Vapnik, 1995).

SVR models typically use kernel functions to capture non-linear relationships. From a bias-variance perspective, the behavior of SVR depends on the complexity of the selected kernel and its hyperparameters. If a linear kernel is used, SVR essentially becomes a linear regression-like model (which may have high bias). With kernels like RBF, the model becomes more flexible, can capture non-linear patterns, and thus bias error may decrease compared to a linear model.

Application with R: In the R environment, the svm() function from the e1071 package can be used for the SVR model. Below, an SVR model has been trained with a radial basis function (RBF) kernel :

```
library(e1071)
svr_model <- svm(charges ~ ., data = train_data, kernel = "radial",
                 epsilon = 0.1, cost = 1)
```

Figure 5. Installation of the RBF kernel SVR model in R (using the e1071 package).

An SVR has been set up with default values of epsilon = 0.1 and cost = 1. Furthermore, since scale=TRUE is the default, numerical features have been scaled during training (so that variables of different scales are balanced during training). The model is expected to learn nonlinear interactions using the RBF kernel during the training process. After training is complete, let's make predictions on the test data and calculate the error metrics:

```
pred_svr <- predict(svr_model, newdata = test_data)
rmse_svr <- sqrt(mean((pred_svr - test_data$charges)^2))
mae_svr  <- mean(abs(pred_svr - test_data$charges))
rmse_svr; mae_svr
```

Figure 6. RMSE and MAE values of the SVR model on the test data.

The performance of our SVR model on the test data was found to be better than that of the linear model. The approximate values obtained are: RMSE=5035.476, MAE=2785.582 dollars. In other words, SVR has significantly reduced the error compared to the linear model by lowering the average error amount to around 5 thousand dollars. The decrease in the MAE value, in particular, indicates that SVR makes more accurate predictions. This success demonstrates that SVR learns curvilinear relationships that linear regression cannot capture, thanks to the RBF kernel. For example, it has been observed that the SVR model partially captures the interaction between the smoker variable and BMI, thereby better predicting individuals who smoke at a very high cost compared to the linear model.
Looking at the training error of the SVR model, RMSE=4702.182, MAE=2451.135, we see that the errors in the training data are close to those in the test data. This indicates that the model did not completely memorize the training data and gained a certain generalization ability. Despite its flexibility, SVR controls model complexity through the ε-insensitive loss function and regularization, which helps prevent excessive variance.

RANDOM FOREST

Random Forest is an ensemble learning method created by combining multiple decision trees. It is a popular algorithm that can be used for both classification and regression problems (Breiman, 1996; Breiman, 2001). Decision trees can be powerful learners on their own, but a single deep tree is often prone to overfitting. Random Forest mitigates this problem by using the bagging (bootstrap aggregating) method: For each tree, a random sample is drawn from the original data using bootstrapping, and a random subset of features is used at each node split (Ho, 1998). This way, the trees make errors as independently as possible, reducing correlation between them.
For regression problems, each tree in the random forest produces an estimate, and the final estimate is obtained by averaging the outputs of all trees. For example, in a forest of 100 trees, the estimate of an individual's healthcare costs would be the average of the 100 trees' estimates. This ensemble approach reduces error compared to a single tree model: Specifically, the variance of the trees (fluctuation in predictions) decreases along with the mean, while bias remains low compared to a single simple model because the trees have high flexibility. For this reason, Random Forest generally offers strong generalization performance by capturing a combination of low bias and reduced variance.

```
library(randomForest)
rf_model <- randomForest(charges ~ ., data = train_data,
                         ntree = 500, mtry = 3, importance = TRUE)
```

Figure 7. Setting up the Random Forest model in R and calculating variable importance.

In R, we can use the randomForest package for the random forest model. This code creates a forest model with 500 trees (ntree=500). The mtry=3 argument specifies the number of variables each tree will randomly select at each node. This ensures sufficient diversity among the trees. We also requested that variable importance scores be recorded with importance=TRUE; with the post-model importance(rf_model) command, we can see which variables are more effective in predictions (smoker, age, and bmi will likely be the most important features, which is a reasonable result).
After training the model, let's make predictions on the test data and examine the performance:

```
pred_rf <- predict(rf_model, newdata = test_data)
rmse_rf <- sqrt(mean((pred_rf - test_data$charges)^2))
mae_rf  <- mean(abs(pred_rf - test_data$charges))
rmse_rf; mae_rf
```

Figure 8. RMSE and MAE values of the Random Forest model on the test data.

The results obtained show that the random forest model performed best in this problem: RMSE = 4697.881, MAE = 2598.298 dollars were calculated. These values are significantly lower compared to the previous linear and SVR models.
The random forest model's error on the training data, RMSE = 2508.193, MAE = 1350.389, is also lower than that on the test data. While this may appear to be overfitting when viewed in isolation, the low test error is due to the variance-reducing structure of the random forest. In other words, even though the model learns the training data almost perfectly, it does so through hundreds of different trees and takes an average, so it does not make as many errors as a single, overly branched tree model when encountering new examples in the test data. This allows us to obtain a complex model with high generalization ability.
In summary, the random forest model has been extremely successful in predicting insurance costs. The model has correctly learned relationships such as predicting very high spending for a smoker, increased costs for individuals above a certain age, and increased costs for high BMI values. In terms of bias–variance, this has led to a significant reduction in the total error.

CONCLUSION AND DISCUSSION

In this section, we examined the concept of bias-variance balance in machine learning models by comparing linear regression, SVR, and random forest models on the Insurance Charges dataset. The dataset was then introduced, and expectations regarding the effects of features on the target variable were explained.
When comparing the experimental results obtained with the three different models, the linear regression model, being the simplest model, failed to fully capture the true relationship due to high bias, resulting in poor performance. Random forest, as the most complex model, had low bias but a tendency toward high variance in individual trees. However, it balanced the variance through its ensemble mechanism and achieved the best generalization performance. The SVR model performed between these two, managing to keep bias and variance errors at a moderate level with appropriate complexity and regularization.
A particularly noteworthy point is how signs of overfitting appear in the models: Random forest's very low training error and higher test error (a clear difference between train and test) is a classic potential sign of overfitting. If there were only a single decision tree, this difference would likely translate to a much larger error in the test; however, in the random forest, this difference is still acceptable because variance is reduced.

On the other hand, similar and high training and test errors in linear regression indicate that the model fails both in the training data and in the test, meaning its capacity is insufficient (a sign of underfitting). In SVR, however, a moderate difference between training and test errors may be observed; the model fits the data slightly better in training, but part of this fit is generalizable to the test – this is a sign of a successful bias-variance balance.
Our experimental findings are consistent with the generally accepted view in the literature: Simple models (e.g., linear regression) have high bias and low variance, while highly complex models (such as deep trees) have low bias and high variance[9][10]. The best generalization is achieved by selecting an appropriate model or method between these two extremes. In this particular study, random forest was the model that best achieved this balance.
In summary, the bias–variance balance is a critical guide when determining a model's complexity and capacity. A simple model like linear regression will fail (underfit) if the relationship it needs to learn is more complex; in this case, the model's complexity must be increased. However, if we increase the

model's complexity too much (e.g., decision trees with unlimited depth), the model will learn the noise in the data and errors will increase in new data (overfitting). It is possible to control the variance of complex models using ensembles such as random forests or regularization techniques (such as ridge and lasso). In this study, it was observed that SVR does this with the epsilon-insensitive approach and random forest with the bagging approach.

In conclusion, selecting the best model in a machine learning problem is not just about minimizing training error; the real goal is to ensure that the model performs well on unknown new data. This is only possible with a model/method that balances bias and variance appropriately. The performance improvements we observed when moving from linear regression to SVR and then to random forest provide important lessons on when increasing model complexity and capacity can be beneficial and when it can be detrimental.

KAYNAKLAR

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.

Breiman L. Bagging predictors. Machine Learning. 1996;24(2):123–140. doi:10.1007/BF00058655

Breiman L. Random forests. Machine Learning. 2001;45(1):5–32. doi:10.1023/A:1010933404324

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural computation, 4(1), 1-58.

Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer; 2009.

James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: with Applications in R. Springer, New York; 2013

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Jui, T. D., & Rivas, P. (2024, December). Analyzing Insurance Cost Estimation: A Supervised Regression Approach. In International Conference on Computational Science and Computational Intelligence (pp. 243-253). Cham: Springer Nature Switzerland.

Liu, B., & Mazumder, R. (2025). Randomization can reduce both bias and variance: A case study in random forests. Journal of Machine Learning Research, 26(150), 1-49.

Mirichoi M. Medical Cost Personal Datasets. Kaggle; 2018. https://www.kaggle.com/datasets/mirichoi0218/insurance

Rocks, J. W., & Mehta, P. (2022). Bias-variance decomposition of overparameterized regression with random linear features. Physical Review E, 106(2), 025304.

Rocks, J. W., & Mehta, P. (2022). Memorizing without overfitting: Bias, variance, and interpolation in overparameterized models. Physical review research, 4(1), 013201.

Smola AJ, Schölkopf B. A tutorial on support vector regression. Statistics and Computing. 2004;14(3):199–222. doi:10.1023/B:STCO.0000035301.49549.88.

Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer; 1995.

Yao, D., Li, J., & Shen, Y. (2023). Ensemble learning for insurance premium prediction: a comparative analysis of XGBoost, Random Forest, and SVM.

Chapter 4

LIGHT CURVE ANALYSIS OF THE ECLIP-SING BINARY STAR CSS_J155848.8+215034 BASED ON SUPERWASP, ASAS-SN, CRTS AND ZTF DATA

Burcu ÖZKARDEŞ[1]

1 Assoc. Prof. Burcu ÖZKARDEŞ, Department of Space Sciences and Technologies, Faculty of Science, Çanakkale Onsekiz Mart University, Terzioğlu Kampüsü, TR-17020, Çanakkale, Türkiye. ORCID ID: 0000-0002-6764-9299

1. INTRODUCTION

W Ursae Majoris (W UMa) type eclipsing binaries constitute a well-defined class of close binary systems with extremely short orbital periods, typically between 0.2 and 1.0 days. These systems are classified as overcontact binaries, meaning that both stellar components fill their Roche lobes and are embedded in a shared common envelope. The fundamental classification of W UMa systems was established by Binnendijk (1970), who emphasized their photometric characteristics and divided them into A- and W-type subclasses. Due to their compact configuration and strong interaction, W UMa binaries provide valuable insights into the physics of mass transfer and angular momentum loss in close stellar systems (Rucinski, 1998).

Observationally, W UMa-type binaries display continuous and nearly sinusoidal light curves with primary and secondary eclipses of comparable depth. This behavior implies that both components have nearly equal surface temperatures, even when their masses differ significantly (Lucy 1968). The stars are typically late-type main-sequence objects with spectral types ranging from F to K and show signs of strong magnetic activity, including starspots and chromospheric emission. Such activity can lead to light-curve asymmetries, most notably the O'Connell effect, which is frequently observed in these systems (Rucinski, 1998).

The defining structural feature of W UMa systems is the presence of a common convective envelope, formed as both stars overfill their Roche lobes. Within this envelope, efficient energy transfer occurs from the more massive component to the less massive one, resulting in nearly identical effective temperatures at the stellar surfaces (Lucy 1968). The system geometry is strongly influenced by tidal forces, leading to synchronous rotation and significant stellar deformation. Convective energy transport within the shared envelope is thought to play a crucial role in maintaining thermal equilibrium across the system (Eggleton, 2006).

The evolutionary origin of W UMa-type binaries is generally attributed to initially detached close binaries that evolve into contact through angular momentum loss mechanisms. Magnetic braking driven by stellar winds gradually reduces orbital separation, eventually causing Roche lobe overflow and the onset of mass transfer (Stepien 2006). During this process, the mass ratio may be reversed, such that the originally more massive star becomes the less massive component. Observed orbital period variations in many W UMa systems support the presence of ongoing mass and angular momentum exchange (Qian, 2001).

The long-term evolution and ultimate fate of W UMa-type binaries remain subjects of active investigation. Some systems appear capable of maintaining stable contact configurations for several gigayears, while others may evolve toward merger events, producing single rapidly rotating stars such as FK Comae-type objects (Stepien, 2006). Additionally, the frequent detection of tertiary companions suggests that three-body interactions may significantly influence the orbital evolution and stability of W UMa systems (Yakut and Eggleton, 2005). Understanding these processes is essential for developing comprehensive models of close binary evolution and stellar mergers.

We focus on the investigation of the neglected binary star CSS_J155848.8+215034 photometrically in this work. We aimed to obtain physical parameters of the system by analyzing four different archival data in the literature, for the first time.

There are limited studies for the eclipsing pair CSS_J155848.8+215034 in the literature. Drake et al. (2014) presented nearly 47,000 periodic variables obtained from the analysis of 5.4 million variable star candidates in the Catalina Surveys Data Release-1 (CSDR1) covered the system. In that study the variability classification for the system is given as a W UMa type contact binary. Heinze et al. (2018) presented the first data catalog of variable stars qualified by the ATLAS (Asteroid Terrestial-impact Last Alert System) operations. In that catalog, the variable classification of the system has been given to be CBH (Close binary, half period). Qian et al. (2020) cataloged new stellar atmospheric parameters of 9149 EW type contact binaries using low and medium resolution spectroscopic surveys of LAMOST including the binary star CSS_J155848.8+215034. Basic photometric features of the system, obtained from the SIMBAD database, are given in Table 1.

Table 1

Basic Features of CSS_J155848.8+215034Based on SIMBAD Database. Symbols are as follows: The equatorial coordinates (α, δ); visual magnitude in V band (Vmag);GAIA parallax (π; Gaia Collaboration, 2022)

Star Name	α (h:m:s)	δ (°:′:″)	*V* (mag.)	π (mas)
CRTS J155848.8+215034	15:58:48	+21:50:34	13.37	1.2781

2. O-C DIAGRAM ANALYSIS

First, the minimum light times required to generate the *O-C* (Observed minus Calculated) diagram were searched from the literature. Only two minimum times were found, one in the VSX[1] (The International Variable Star Index) database and the other in the ASAS-SN[2] (All-Sky Automated Survey for Supernovae Sky Patrol) database. In this work, we obtained 19 primary minima times (Min I) and 15 secondary minima times (Min II) from SuperWASP *V* light curve. Table 2 provides these eclipsing times in HJD time.

Table 2

Eclipsing Times of CSS_J155848.8+215034 from SuperWASP light curve

HJD (+2400000)	Min. Type	Errror	HJD (+2400000)	Min. Type	Errror
53855.5221	I	0.0002	53832.6939	II	0.0003
53920.4344	I	0.0003	54204.5611	II	0.0003
54170.7389	I	0.0002	54206.6256	II	0.0002
54202.6729	I	0.0004	54208.6878	II	0.0004
54212.6331	I	0.0002	54214.5188	II	0.0005
54213.6636	I	0.0005	54215.5534	II	0.0005
54230.4912	I	0.0002	54235.4661	II	0.0003
54231.5168	I	0.0004	54236.4932	II	0.0004
54232.5478	I	0.0004	54256.4085	II	0.0004
54233.5777	I	0.0002	54257.4404	II	0.0002
54235.6425	I	0.0002	54268.4266	II	0.0002
54262.4219	I	0.0002	54269.4620	II	0.0002

[1] https://www.aavso.org/vsx/index.php?view=search.top
[2] https://asas-sn.osu.edu/?ra=239.70348&dec=21.84296

54263.4533	I	0.0003	54270.4910	II	0.0004
54264.4807	I	0.0004	54290.4066	II	0.0004
54265.5116	I	0.0002	54291.4394	II	0.0002
54275.4661	I	0.0002	-	-	-
54285.4302	I	0.0002	-	-	-
54286.4613	I	0.0002	-	-	-
54295.3940	I	0.0002	-	-	-

The *O-C* diagram has been created using all minimum light times given in Table 2 for studying the period variation of the eclipsing binary. The *O-C* diagram of the system is shown in Figure 1 with the residuals.

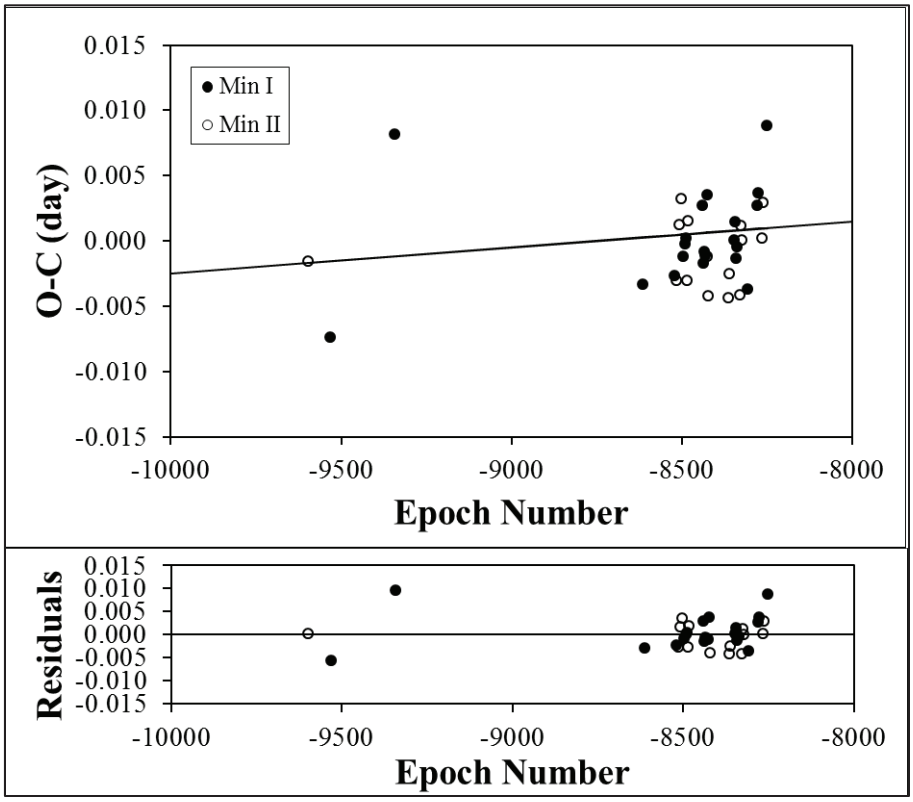


Figure 1. Top: *O-C* diagram for CSS_J155848.8+215034; bottom: residuals.

It can be clearly seen nonvariation in the system's orbital period in Figure 1. We just corrected the light elements using the linear regression fit to the *O-C* values. The resulting linear ephemerides (namley T_0 and P) are given below. These values were used in the phasing of the light curves.

$$\text{Min. I (HJD)} = 2457128.1908\pm0.0114+0^{d}.3433704\pm0.0000018\times E. \quad (1)$$

3. LIGHT CURVE MODELING

Databases containing the system's photometric data were searched, and data from four databases (SuperWASP, ASAS-SN, CRTS and ZTF) were used in the light curve analysis. The phase

folded light curves was modelled with the Wilson-Devinney (WD) method (Wilson & Devinney, 1971) supplemented with the Monte Carlo (MC) algorithm (Kreiner et al., 2003).

The primary component's temperature was adopted as 5901±200 from Qian et al. (2020), and was fixed in the analysis. The coefficients g_{12} and A_{12} were fixed at 0.32 and 0.5, respectively (Lucy, 1967; Rucinski, 1969). By adopting the square-root limb-darkening law, Claret (2017)'s calibration tables were used for limb-darkening coefficients' values. Synchronoused rotation was assumed for the components. The circular orbit was adopted for system's orbit. The third light (l_3) was assumed to be zero because of nonvariation in the orbital period.

The adjustible parameters in the search were as follows: q (mass ratio), i (orbital inclination), φ (phase-shift), T_2 (the secondary component's temperature), $\Omega_1=\Omega_2$ (non-dimensional surface potential parameters of components), L_1 (the primary's luminosity). Due to the absence of spectroscopic mass ratio, the value range for q was given between 0.1 and 5.0 in the MC search. The resulting parameters are listed in Table 3. The agreement between best theoretical fits derived from the MC solution and observational data is given comparatively in Figure 2a,b,c,d. The Roche geometry (Bradstreet and Steelman, 2002) of the system is shown in Figure 3.

Table 3

The Light Curve Solution Parameters of CSS_J155848.8+215034

Parameters	**Value**	**Error**	**Parameters**	**Value**	**Error**
i (°)	63.9	0.2	q (M_2/M_1)	0.67774	0.0131
Phase shift (φ)	0.0057	0.0008	$\Omega_1 = \Omega_2$	3.1239	0.0241
T_1 (K)	5901	Fixed	r_1 (mean)	0.43	0.01
T_2 (K)	5774	201	r_2 (mean)	0.36	0.01
Fill out (%)	20	-			
L_1/L_{total} (SWASP-*V*)	0.56	0.01	$\Sigma(O-C)^2$ (SWASP *V*)	2.295	
L_1/L_{total} (ASAS-SN *V*)	0.56	0.01	$\Sigma(O-C)^2$ (ASAS-SN *V*)	0.135	
L_1/L_{total} (CRTS *V*)	0.57	0.01	$\Sigma(O-C)^2$ (CRTS *V*)	0.231	
L_1/L_{total} (ZTF *g*)	0.56	0.01	$\Sigma(O-C)^2$ (ZTF *g*)	0.184	
L_1/L_{total} (ZTF *r*)	0.56	0.01	$\Sigma(O-C)^2$ (ZTF *g*)	0.193	
L_1/L_{total} (ZTF *i*)	0.55	0.01	$\Sigma(O-C)^2$ (ZTF *g*)	0.051	

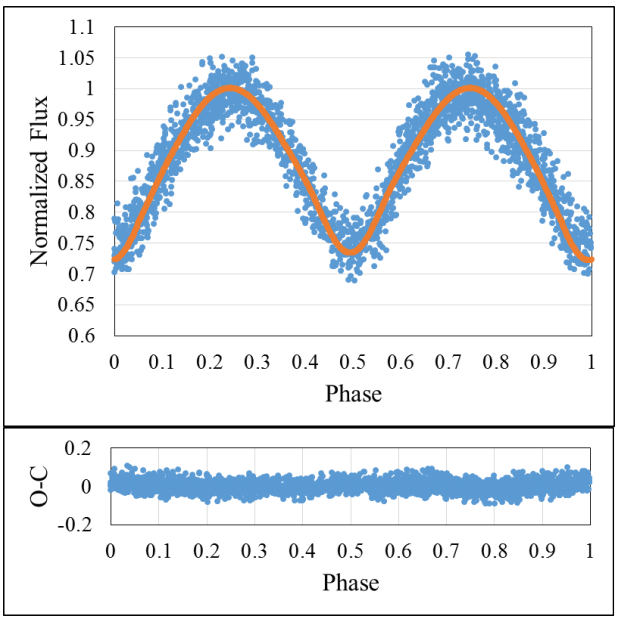


Figure 2a. Comparision between the theoretical light curve and SuperWASP *V* band light curve. Residuals (O-C) are given in bottom panel.

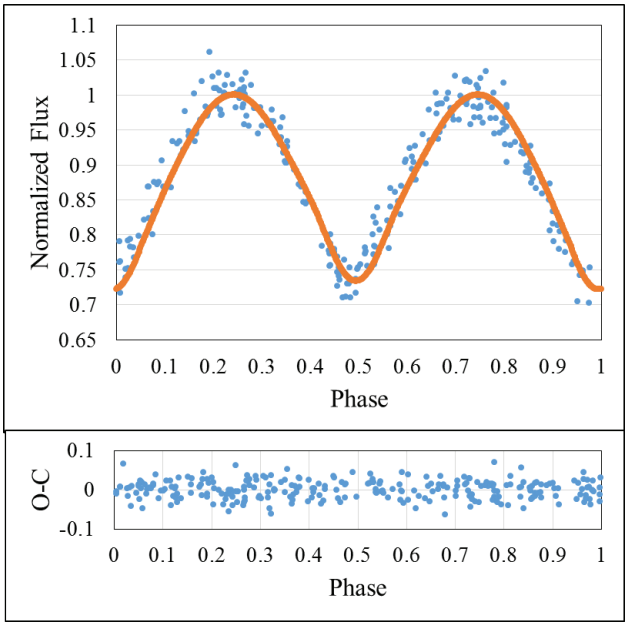


Figure 2b. Same as Figure 2a, but for ASAS-SN *V* band.

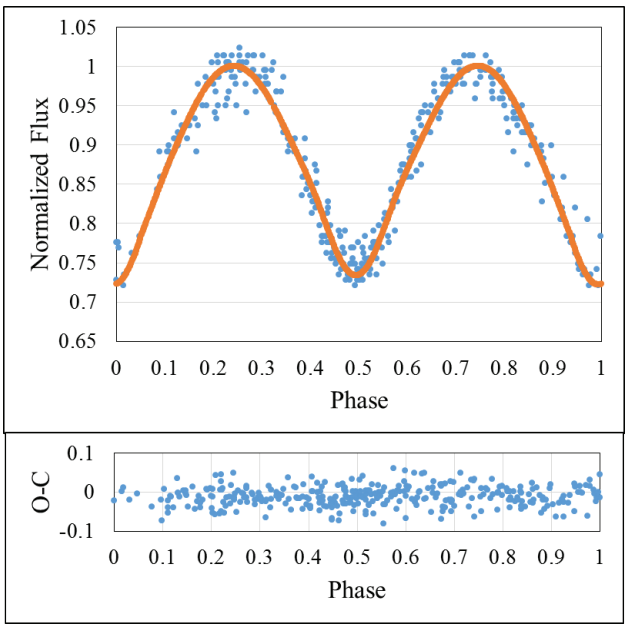


Figure 2c. Same as Figure 2a, but for CRTS *V* band.

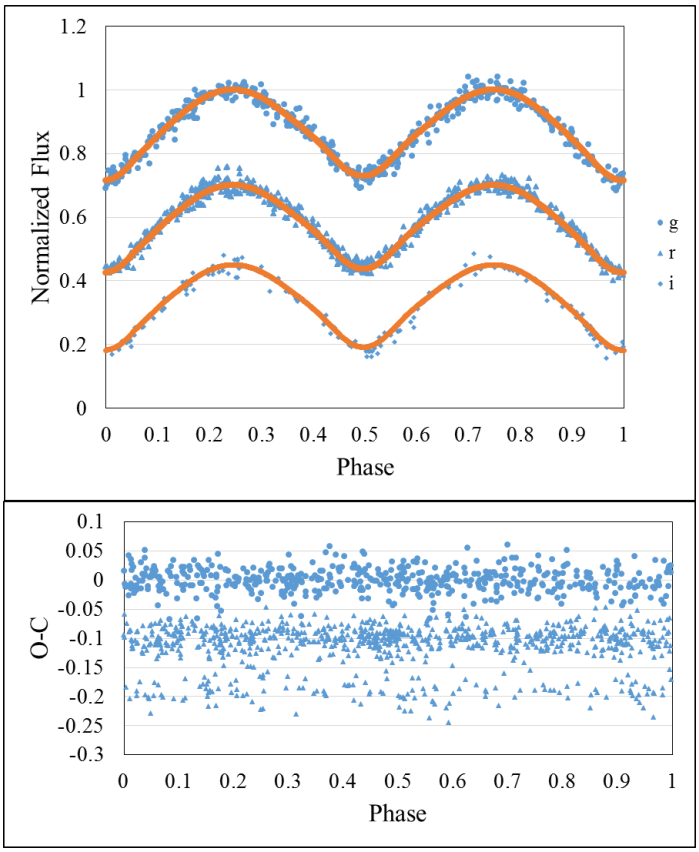


Figure 2d. Same as Figure 2a, but for ZTF *gri* band.

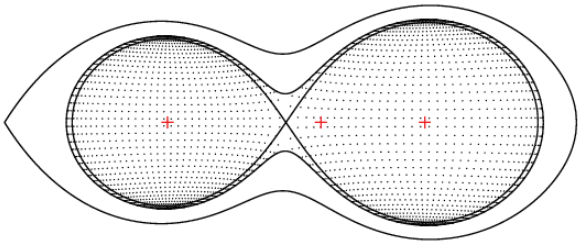


Figure 3. The 2D surface geometry of the system.

4. RESULTS AND CONCLUSION

CSS_J155848.8+215034's global parameters were estimated due to non-spectroscopic mass ratio of the system, The primary component's mass was assumed to be 1.05±0.03 $M_\odot$, using the calibration table of the empirical MS stars (Drilling and Landolt, 2000). The mass of the secondary component was determined as 0.71±0.06 $M_\odot$. The semi-major axis of the binary orbit was calculated as 2.49±0.01 $R_\odot$ using the Keplerian orbital equation. The radii of the components were calculated to be 1.07±0.01 $R_\odot$ and 0.90±0.01 $R_\odot$, respectively, using the equation (2).

$$r_i = \frac{R_i}{a} \tag{2}$$

The surface gravity values and bolometric magnitudes of the component stars were calculated using Pecaut and Mamajek (2013)'s the values for the sun. The physical parameters are listed in Table 4. The distance of the system was computed to be 735±44 pc considering the interstellar reddening. Here, galactic extinction A_V was taken to be 0.192 from Schlafly and Finkbeiner (2011), and color excess E(*B-V*) are calculated as 0.061. It seems that the distance determined in this study are in agreement with the parallax of the system given by Gaia DR3 (Mowlavi et al., 2023).

Table 4

Absolute Parameters of The Eclipsing Binary CSS_J155848.8+215034

Parameters	Value	Error	Parameters	Value	Error
a ($R_\odot$)	2.49	0.01	L_2 ($L_\odot$)	0.81	0.11
M_1 ($M_\odot$)	1.05	0.03	log g_1 (cgs)	4.40	0.01
M_2 ($M_\odot$)	0.71	0.06	log g_2 (cgs)	4.38	0.03
R_1 ($R_\odot$)	1.07	0.01	d (pc)	735	44
R_2 ($R_\odot$)	0.90	0.01	d_Gaia (pc)	782	8
L_1 ($L_\odot$)	1.25	0.17			

The positions of both components ofCSS_J155848.8+215034's are plotted in the log (T_{eff}) – log ($L/L_\odot$) plane (see Figure 4). The theoretical lines of ZAMS, TAMS and the evolutionary tracks given by Girardi et al. (2000) for Z=0.019 were used. Selected A-subtype components shown in the diagram were taken from Latkovic et al. (2021) and listed in Table 5. The primary component is located within the main-sequence band while the secondary one is located close to the ZAMS. The positions of the component stars are found to be consistent when compared with the archival A type contact binary stars. We tried to estimate the age of the system using the evolutionary isochrones from Girardi et al. (2000), and the best fit was found at the age of 3.981 Gyr.

Table 5

Temperatures and luminosities of the selected A subtype contact binaries.

Star Name	**T_1 (K)**	**T_2 (K)**	**L_1 ($L_\odot$)**	**L_2 ($L_\odot$)**
V712 Mon	5600	5302	0.596	0.480
V692 Pup	5600	5398	0.876	0.756
V695 Pup	5625	5394	0.714	0.502
HH Boo	5680	5386	0.749	0.570
V899 Her	5700	5677	2.32	1.40
CRTS J004259.3+410629	5724	5706	0.69	0.24
II Per	5740	4464	6.86	0.25
ASAS J024155-2507.8	5746	5537	1.37	0.50
AO Aqr	5750	5708	2.62	0.70
V658 Lyr	5752	5628	1.52	0.33
USNO-A2.0 1200-18678842	5767	5562	1.436	1.106
IK Boo	5781	5422	0.641	0.729
ASAS J035020-8017.4	5790	5675	3.14	0.70
EQ Tau	5800	5730	1.327	0.61
Cl* NGC 104 WS V95	5800	5707	1.12	0.229
GZ And	5810	5500	1.033	0.451
OGLE BW1 V38	5830	5770	2.151	0.771
OGLE J004619.65-725056.2	5850	5718	1.19	0.99
BQ Ari	5850	5659	0.91	0.748
1SWASP J210318.76+021002.2	5850	5778	1.59	1.422
BG Vul	5868	5520	1.632	1.151
UY UMa	5900	5882	1.95	0.52
TU Boo	5900	5870	1.107	0.575
Cl* NGC 6838 YM V5	5900	5829	1.31	0.19
NSVS 908513	5923	5615	1.031	0.646
OU Ser	5940	5759	1.488	0.292
FZ Ori	5940	4983	1.5	1.34
BO Ari	5940	5908	1.325	0.29
V410 Aur	5950	5898	2.1	0.37
TV Mus	5980	5808	3.33	0.71
ASAS J102556+2049.3	6000	5930	1.41	0.23

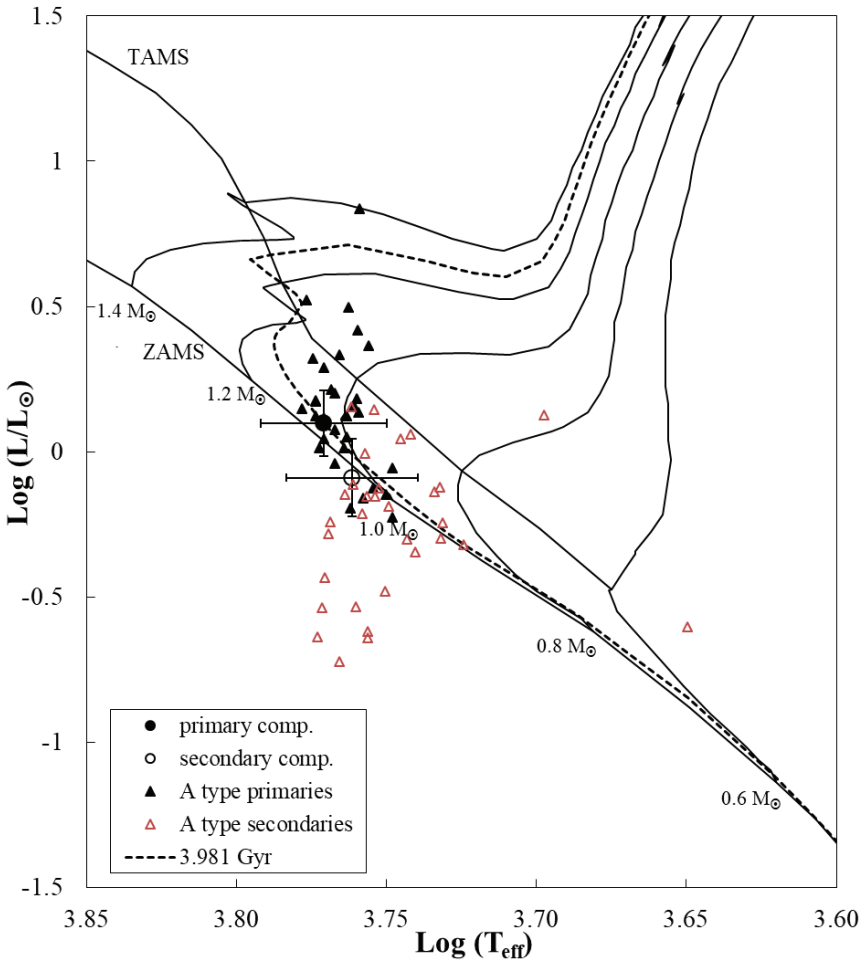


Figure 4. Position of the components of the eclipsing binary CSS_J155848.8+215034 in the log (T_{eff}) – log ($L/L_{\odot}$) diagram.

Analysis of the archival light curves resulted in the system being classified as an A type contact binary. Determining the precise absolute parameters of this binary, investigating its evolutionary stage in detail and identifying the presence of a possible third component requires multi-color light curves and spectroscopic observations.

REFERENCES

1. Binnendijk, L. (1970). The orbital elements of W Ursae Majoris systems, Vistas in Astronomy, vol. 12, Issue 1, pp.217-256.

2. Bradstreet, D.H., Steelman, D.P., (2002). AAS 201, 75.02.

3. Claret, A., (2017). Limb and gravity-darkening coefficients for the TESS satellite at several metallicities, surface gravities, and microturbulent velocities. Astronomy and Astrophysics, 600, A30.

4. Drake, A. J., Graham, M. J., & Djorgovski, S. G. et al. (2014). The Catalina Surveys Periodic Variable Star Catalog. *The Astronomical Journal Supplement Series* , 213(1), 9, 29 pp.

5. Drilling, J.S., Landolt, A.U. (2000). In: Cox, Arthur N. (Ed.), Allen's Astrophysical Quantities, fourth ed. AIP Press; Springer, New York, ISBN: 0-387-98746-0, p. 381.

6. Eggleton, P. (2006). Evolutionary Processes in Binary and Multiple Stars, Evolutionary Processes in Binary and Multiple Stars, by Peter Eggleton, pp. . ISBN 0521855578. Cambridge, UK: Cambridge University Press.

7. Gaia Collaboration, (2022). VizieR Online Data Catalog: Gaia DR3 Part 1. Main source. VizieR On-line Data Catalog: I/355, DOI: 10.26093/cds/vizier.1355.

8. Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. (2000). Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 Msun, and from Z=0.0004 to 0.03. *Astronomy and Astrophysics Supplement*, 141, 371-383.

9. Heinz, A. N., Tonry, J. L., & Denneau, L. et al. (2018). A First Catalog of Variable Stars Measured by the Asteroid Terrestrial-impact Last Alert System (ATLAS). *The Astronomical Journal*, 156(5), 241, 49 pp.

10. Kreiner, J. M., Rucinski S. M., Zola S., Niarchos P., Ogloza W., Stachowski G., Baran A., Gazeas K., Drozdz M., Zakrzewski, B., Pokrzywka B., Kjurkchieva D. and Marchev D. (2003). Physical Parameters Of Components In Close Binary Systems. I. *Astronomy and Astrophysics*, 412: 465-471.

11. Latkovic, O., Ceki, A., & Lazarevic, S. (2021). Statistics of 700 Individually Studied W UMa Stars. *The Astrophysical Journal Supplement Series*, 254(1), id.10, 18 pp.

12. Lucy, L. B. (1967). Gravity-Darkening for Stars with Convective Envelopes. *Zeitschrift für Astrophysik*, 65, p.89.

13. Mowlavi, N.; Holl, B.; Lecoeur-Taïbi, I., et al. (2023). Gaia Data Release 3. The first Gaia catalogue of eclipsing-binary candidates. Astronomy & Astrophysics, 674, A16.

14. Qian, S-B., Zhu, Li, Liu, L. (2020). Contact binaries at different evolutionary stages, Research in Astronomy and Astrophysics, Volume 20, Issue 10, id.163, 20 pp

15. Pecaut, M. J. and Mamajek, E. E. (2013). Intrinsic Colors, Temperatures, and Bolometric Corrections of Pre-main-sequence Stars. *The Astrophysical Journal Supplement*, 208(1), id. 9, 22 pp.

16. Rucinski, S. M. (1969). The Proximity Effects in Close Binary Systems. II. The Bolometric Reflection Effect for Stars with Deep Convective Envelopes. *Acta Astronomica*, 19, p. 245.

17. Schlafly, E.F., Finkbeiner, D.P. (2011). Measuring reddening with sloan digital sky Survey stellar spectra and recalibrating SFD. The Astrophysical Journal, 737 (2),id 103.

18. Wilson, R. E. and Devinney, E. J. (1971). Realization of Accurate Close-Binary Light Curves: Application to MR Cygni. *Astrophysical Journal*, 166, 605-619.

Chapter 5

STATISTICAL INFERENCE FOR KUMARASWAMY GUMBEL DISTRIBUTION: AN APPLICATION TO WAVE HEIGHT DATA

Gamze GÜVEN[1], Birdal ŞENOĞLU[2]

1 Assoc. Prof., Eskisehir Osmangazi University, ORCID:0000-0002-8821-3179, gamzeguven@ogu.edu.tr
2 Prof., Ankara University, ORCID: 0000-0003-3707-2393, senoglu@science.ankara.edu.tr

1 Introduction

Gumbel distribution, also known as the extreme value Type I distribution, proposed by Gumbel [1941] is commonly used in various fields especially in engineering, insurance, meteorology, finance and climatology. For example, Rodríguez et. al. (2009) proposed a goodness of fit test for Gumbel distribution and used it for modeling rainfall data. Solomon and Prince (2013) used Gumbel distribution for modeling the annual maximum discharge of the Osse River for a period of 20 years. Kang et al. (2015) conducted a study on Jeju Island, South Korea to examine the suitability of various extreme wind values for the Gumbel distribution and demonstrated that daily maximum wind speeds were the best extreme wind values for the Gumbel distribution regardless of topographical conditions. Aydin and Senoglu (2015) compared the performances of the seven different parameter estimation methods for the Gumbel distribution and illustrated their findings via a real-life example taken from hydrology literature. Purohit and Lalit (2022) modeled the risk-neutral density function for European call option pricing using Gumbel distribution.

If X follows a Gumbel distribution with location parameter μ and scale parameter σ, then the cumulative distribution function (cdf) and probability density function (pdf) of X is given by

$$F_{Gum}(x) = e^{-e^{-(x-\mu)/\sigma}}, -\infty < x < \infty,\ -\infty < \mu < \infty, \sigma > 0 \tag{1}$$

and

$$f_{Gum}(x) = \frac{1}{\sigma} e^{-\left((x-\mu)/\sigma + e^{-(x-\mu)/\sigma}\right)}, \tag{2}$$

respectively and it is denoted by $X \sim \text{Gumbel}(\mu, \sigma)$. The mean, variance, mode and median of Gumbel distribution are given by $\mu + \gamma\sigma$, $\left(\frac{\pi^2}{6}\right)\sigma^2$, μ and $\mu - \sigma \ln(\ln 2)$, respectively. Here, $\gamma \approx 0.57721566$ is the Euler constant. The skewness and kurtosis of Gumbel distribution are about 1.14 and 5.4, respectively. It should be realized that Gumbel is not a flexible distribution since it has constant skewness and kurtosis values. Therefore, various generalizations of it have been proposed in literature for modeling data with broad range of skewness and kurtosis. For example, Nadarajah and Kotz (2004) introduced beta Gumbel (BG) distribution by applying the logit transformation to a beta random variable. They examined mathematical properties of the BG distribution in detail and used maximum likelihood (ML) methodology for parameter estimation. Nadarajah (2006) proposed exponentiated Gumbel (EG) distribution by applying the same power-transformation idea used for the exponentiated exponential distribution. Cooray (2010) presented generalized Gumbel (GG) distribution by taking a logarithmic transformation of an odd Weibull random variable. They estimated its parameters using ML method in two different ways. Cordeiro et al. (2012) proposed Kumaraswamy Gumbel (KumGum) distribution, which is a generalization of Gumbel distribution and an important alternative to the BG distribution. Pinheiro and Ferrari (2016) presented a comprehensive comparative review of distributions that contain the Gumbel distribution as a special or limiting case. Demirhan (2018) considered generalized Gumbel distribution (GGD) which is a member of the generalized Gompertz-Verhulst family. They established its link to the marginal distributions of the generalized multivariate Gumbel model and demonstrated that the GGD distribution is effective for modeling block maxima arising from leptokurtic data. Gómez et al. (2019) developed an extension of Gumbel distribution using the construction principle of the ordinary slash distribution. Liu et al. (2024) proposed flexible Gumbel (FG) distribution as a mixture of two Gumbel

distributions, one for maximum and one for minimum. They used frequentist and Bayesian methods for estimating the parameters of the FG distribution.

Our motivation for this study comes from wave height (m) data set previously modeled by GG distribution. Wave height analysis constitutes an important issue in coastal and offshore engineering applications. In analyzing mentioned data set, we realized that it can also be effectively modeled by using KumGum distribution. Among the various generalizations of the Gumbel distribution proposed in the literature, the reason for focusing on the widely used and popular KumGum distribution is its flexibility in modeling data sets encountered in flood frequency analysis, risk-based engineering, software reliability engineering, offshore engineering etc., see Cordeiro et. al. (2012) for details.
Following this modeling framework, modified maximum likelihood (MML) methodology proposed by Tiku (1967, 1968) is used to estimate the unknown location and scale parameters of KumGum distribution under the assumption that the shape parameters are known. To provide more comprehensive comparisons, a Monte Carlo simulation study is conducted to compare the efficiencies of the MML estimators with those of the widely used ML, least squares (LS) and weighted least squares (WLS) estimators reported in the literature. Finally, the wave height data set is analyzed to illustrate the modeling performance of KumGum distribution.

2 KumGum Distribution

KumGum distribution is obtained by substituting the baseline cdf $F_{Gum}(x)$ and pdf $f_{Gum}(x)$ given in (1) and (2) into the cdf and pdf of the Kum-G distribution, see Kumaraswamy (1980) and Jones (2009) for details of the Kum-G distribution.
If X follows a KumGum distribution with parameters a, b, μ and σ, then it is denoted by $X \sim \text{KumGum}(a, b, \mu, \sigma)$. The cdf and pdf of the KumGum distribution are

$$F(x) = 1 - \left\{1 - \exp\left(-a \exp\left\{-\frac{(x-\mu)}{\sigma}\right\}\right)\right\}^b, \quad -\infty < x, \mu < \infty, \sigma > 0 \tag{3}$$

and

$$f(x) = \frac{ab}{\sigma} \exp\left\{-\frac{(x-\mu)}{\sigma}\right\} \exp\left\{-a \exp\left\{-\frac{(x-\mu)}{\sigma}\right\}\right\} \left\{1 - \exp\left(-a \exp\left\{-\frac{(x-\mu)}{\sigma}\right\}\right)\right\}^{b-1}, \tag{4}$$

respectively. Here, a and b are the shape parameters and μ and σ are the location and scale parameters, respectively.

To provide further information about the KumGum distribution, pdf plots and the table reporting the corresponding skewness and kurtosis values are given in Figure 1 and Table 1, respectively.

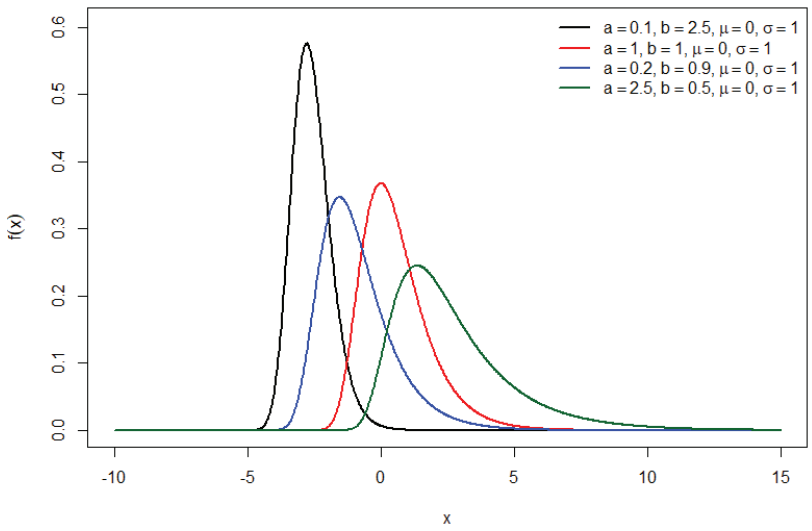


Figure 1. The pdf plots of the KumGum distribution for some selected parameter values.

Table 1. The simulated skewness and kurtosis values of the KumGum distribution for some selected parameter values.

$(a, b, \mu, \sigma) =$	(0.1,2.5,0,1)	(1,1,0,1)	(0.2,0.9,0,1)	(2.5,0.5,0,1)
Skewness	0.60	1.14	1.21	1.56
Kurtosis	3.77	5.40	5.65	7.09

It should be noted that Gumbel is the most important sub-model of the KumGum distribution and it is called as the standard Gumbel for the parameter setting $(a, b, \mu, \sigma) = (1,1,0,1)$. Also realize that alternative parameter settings generate distributions with either reduced or increased asymmetry and tail thickness as compared to the standard Gumbel distribution, see Figure 1 and Table 1.

3 Parameter Estimation Methods

Let $X_1, X_2, \dots, X_n$ be independent random variables such that $X_i \sim \text{KumGum}(a, b, \mu, \sigma)$ $i = 1, 2, \dots, n$. In the following subsections, ML, MML, LS and WLS methodologies are introduced to obtain estimators of unknown location parameter μ and scale parameter σ under the assumption that the shape parameters a and b are known.

3.1 ML Estimation

The ML estimators of μ and σ are obtained using the methodology provided below.

Step 1. Construct the likelihood (L) function based on $X_i \sim \text{KumGum}(a, b, \mu, \sigma)$

$$L = \prod_{i=1}^{n} \frac{ab}{\sigma} \exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\} \exp\left\{-a \exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\}\right\} \left\{1 - \exp\left(-a \exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\}\right)\right\}^{b-1} \quad (5)$$

Step 2. Obtain the log-likelihood $(\ln L)$ function by taking the logarithm of the L function

$$\ln L = n\ln(ab) - n\ln(\sigma) - \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{n} a\exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\} + (b-1)\sum_{i=1}^{n}\ln\left\{1-\exp\left(-a\exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\}\right)\right\}. \quad (6)$$

Step 3. Obtain the score functions by differentiating $\ln L$ with respect to the parameters μ and σ and then equate them to zero to obtain the likelihood equations

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n}\frac{1}{\sigma} - \frac{a}{\sigma}\sum_{i=1}^{n}\exp\{-z_i\} + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n}\frac{\exp\{-z_i\}}{\exp(a\exp\{-z_i\})-1} = 0 \quad (7)$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} z_i - \frac{a}{\sigma}\sum_{i=1}^{n} z_i\exp\{-z_i\} + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n} z_i\frac{\exp\{-z_i\}}{\exp(a\exp\{-z_i\})-1} = 0\,, \quad (8)$$

respectively, where $z_i = \frac{x_i-\mu}{\sigma}$.

Step 4. Solve the Equations (7) and (8) simultaneously to obtain the ML estimators $\hat{\mu}_{ML}$ and $\hat{\sigma}_{ML}$ for parameters μ and σ, respectively.

Remark. It should be noted that explicit solutions of the Equations in (7) and (8) are not obtained, therefore iterative numerical algorithms such as Newton Raphson, Nelder-Mead, bisection etc. are often used for solving this non-linear system of equations. However, some convergence problems may arise at this stage, see Barnett (1966), Puthenpura and Sinha (1986), Vaughan (1992). To overcome the difficulties encountered in the numerical methods, we use MML methodology having the closed-form solutions for the unknown KumGum parameters, see Tiku (1967,1968) and the following subsection for the details.

3.2 MML Estimation

The MML estimators of μ and σ are obtained using the methodology provided below.

Step 1. Express the likelihood Equations (7) and (8) in terms of ordered statistics

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n}\frac{1}{\sigma} - \frac{a}{\sigma}\sum_{i=1}^{n} g_1\left(z_{(i)}\right) + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n} g_2\left(z_{(i)}\right) = 0 \quad (9)$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} z_{(i)} - \frac{a}{\sigma}\sum_{i=1}^{n} z_{(i)}g_1\left(z_{(i)}\right) + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n} z_{(i)}g_2\left(z_{(i)}\right) = 0, \quad (10)$$

respectively. Here, $g_1\left(z_{(i)}\right) = \exp\left\{-z_{(i)}\right\}$, $g_2\left(z_{(i)}\right) = \frac{\exp\{-z_{(i)}\}}{\exp(a\exp\{-z_{(i)}\})-1}$ and $z_{(i)} = \frac{x_{(i)}-\mu}{\sigma}$. Note that ordering the observations does not change their sum, i.e., $\sum_{i=1}^{n} z_{(i)} = \sum_{i=1}^{n} z_i$.

Step 2. Linearize the functions $g_1(.)$ and $g_2(.)$ around the expected values of the standardized ordered statistics $t_{(i)} = E\left(z_{(i)}\right)$ using the first two terms of the Taylor series expansion. The linearized forms are given by

$$g_1\left(z_{(i)}\right) \cong \alpha_{1i} - \beta_{1i}z_{(i)} \quad (11)$$

and

$$g_2\left(z_{(i)}\right) \cong \alpha_{2i} + \beta_{2i} z_{(i)}, \tag{12}$$

respectively. Here,

$$\alpha_{1i} = \exp\{-t_{(i)}\} + t_{(i)} \exp\{-t_{(i)}\},$$
$$\beta_{1i} = \exp\{-t_{(i)}\},$$
$$\alpha_{2i} = \frac{(\exp\{-t_{(i)}\}[\exp(a \exp\{-t_{(i)}\})-1])+t_{(i)}\{\exp\{-2t_{(i)}\}[\exp(a \exp\{-t_{(i)}\}+t_{(i)})-a \exp\{a \exp\{-t_{(i)}\}\}-\exp\{t_{(i)}\}]\}}{(\exp(a \exp\{-t_{(i)}\})-1)^2},$$
$$\beta_{2i} = \frac{-\left[\exp\{-2t_{(i)}\}[\exp(a \exp\{-t_{(i)}\}+t_{(i)})-a \exp\{a \exp\{-t_{(i)}\}\}-\exp\{t_{(i)}\}]\right]}{(\exp(a \exp\{-t_{(i)}\})-1)^2},$$

and

$$t_{(i)} = -\ln\left\{-\frac{1}{a}\ln\left[1-\left(1-\left(\frac{i}{n+1}\right)\right)^{\frac{1}{b}}\right]\right\}, i = 1, \dots, n.$$

Note that approximate $t_{(i)}$ values are obtained from the following equality

$$\int_{-\infty}^{t_{(i)}} f(z)dz = \frac{i}{n+1},$$

where $f(z)$ is the standardized KumGum distribution.

Step 3. Incorporate (11) and (12) into the Equations (9) and (10) to obtain modified likelihood equations

$$\frac{\partial \ln L^*}{\partial \mu} = \sum_{i=1}^{n} \frac{1}{\sigma} - \frac{a}{\sigma}\sum_{i=1}^{n}\left(\alpha_{1i} - \beta_{1i} z_{(i)}\right) + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n}\left(\alpha_{2i} + \beta_{2i} z_{(i)}\right) = 0 \tag{13}$$

and

$$\frac{\partial \ln L^*}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} z_{(i)} - \frac{a}{\sigma}\sum_{i=1}^{n} z_{(i)}\left(\alpha_{1i} - \beta_{1i} z_{(i)}\right) + \frac{a(b-1)}{\sigma}\sum_{i=1}^{n} z_{(i)}\left(\alpha_{2i} + \beta_{2i} z_{(i)}\right) = 0. \tag{14}$$

Step 4. Solve the Equations (13) and (14) simultaneously to obtain the MML estimators $\hat{\mu}_{MML}$ and $\hat{\sigma}_{MML}$ for parameters μ and σ, respectively.

Following Steps 1-4, MML estimators of the parameters μ and σ are obtained as follows

$$\hat{\mu}_{MML} = K + L\hat{\sigma}_{MML} \text{ and } \hat{\sigma}_{MML} = \frac{B+\sqrt{B^2+4AC}}{2\sqrt{A(A-1)}}.$$

Here, $m_i = a\beta_{1i} + a(b-1)\beta_{2i}$, $m = \sum_{i=1}^{n} m_i$, $K_i = a\beta_{1i} x_{(i)} + a(b-1)\beta_{2i} x_{(i)}$, $K_1 = \sum_{i=1}^{n} K_i$ $\Delta_i = 1 - a\alpha_{1i} + a(b-1)\alpha_{2i}$, $\Delta = \sum_{i=1}^{n} \Delta_i$, $K = \frac{K_1}{m}$, $L = \frac{\Delta}{m}$, $A = n$, $B = \sum_{i=1}^{n}\left(x_{(i)} - K\right)\Delta_i$ and $C = \sum_{i=1}^{n} m_i\left(x_{(i)} - K\right)^2$.

Remark. MML estimators are functions of the sample observations and have closed forms. They are asymptotically equivalent to the ML estimators.

3.3 LS Estimation

The LS estimators of μ and σ are obtained using the methodology provided below.

Step 1. Define the LS objective function

$$S_{LS} = \sum_{i=1}^{n} \left(F\left(x_{(i)}\right) - i/(n+1) \right)^2 \tag{15}$$

where $x_{(i)}$ denotes the ith order statistics.

Step 2. Incorporate the cdf of the KumGum distribution into Equation (15)

$$S_{LS(KumGum)} = \sum_{i=1}^{n} \left(1 - \left(1 - \exp\left(-a \exp\left(-\left(x_{(i)} - \mu \right)/\sigma \right) \right) \right)^b - i/(n+1) \right)^2 . \tag{16}$$

Step 3. Differentiate $S_{LS(KumGum)}$ in (16) with respect to the parameters μ and σ and obtain the following nonlinear equations

$$\frac{\partial S_{LS(KumGum)}}{\partial \mu} = \sum_{i=1}^{n} \left[1 - \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \right)^b - \frac{i}{n+1} \right] F'_{\mu}\left(x_{(i)}; \mu, \sigma \right) = 0, \tag{17}$$

$$\frac{\partial S_{LS(KumGum)}}{\partial \sigma} = \sum_{i=1}^{n} \left[1 - \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \right)^b - \frac{i}{n+1} \right] F'_{\sigma}\left(x_{(i)}; \mu, \sigma \right) = 0. \tag{18}$$

Here,

$$F'_{\mu}\left(x_{(i)}; \mu, \sigma \right) = \left[-\frac{ab}{\sigma} \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \right)^{b-1} \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right] \tag{19}$$

and

$$F'_{\sigma}\left(x_{(i)}; \mu, \sigma \right) = \left[-\frac{ab}{\sigma} \left(\frac{x_{(i)} - \mu}{\sigma} \right) \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \right)^{b-1} \exp\left(-a \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right) \exp\left(-\left(\frac{x_{(i)} - \mu}{\sigma} \right) \right) \right] \tag{20}$$

are the partial derivatives of the cdf of KumGum distribution with respect to μ and σ, respectively.

Step 4. Solve the Equations (17) and (18) simultaneously to obtain the LS estimators $\hat{\mu}_{LS}$ and $\hat{\sigma}_{LS}$ for parameters μ and σ, respectively.

3.4 WLS Estimation

The WLS estimators of μ and σ are obtained using the methodology provided below.

Step 1. Define the WLS objective function

$$S_{WLS} = \sum_{i=1}^{n} w_i \left(F(x_{(i)}) - i/(n+1) \right)^2, \tag{21}$$

where $x_{(i)}$ denotes the ith order statistics and $w_i = (n+1)^2(n+2)/i(n-i+1)$.

Step 2. Incorporate the cdf of the KumGum distribution into Equation (21)

$$S_{WLS(KumGum)} = \sum_{i=1}^{n} w_i \left(1 - \left(1 - \exp\left(-a \exp\left(-\left(x_{(i)} - \mu \right)/\sigma \right) \right) \right)^b - i/(n+1) \right)^2 . \tag{22}$$

Step 3. Differentiate $S_{WLS(KumGum)}$ in (22) with respect to the parameters μ and σ and obtain the following nonlinear equations

$$\frac{\partial S_{WLS(KumGum)}}{\partial \mu} = \sum_{i=1}^{n} w_i \left[1 - \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)}-\mu}{\sigma}\right)\right)\right)\right)^b - \frac{i}{n+1}\right] F'_{\mu}(x_{(i)};\mu,\sigma) = 0, \quad (23)$$

$$\frac{\partial S_{WLS(KumGum)}}{\partial \sigma} = \sum_{i=1}^{n} w_i \left[1 - \left(1 - \exp\left(-a \exp\left(-\left(\frac{x_{(i)}-\mu}{\sigma}\right)\right)\right)\right)^b - \frac{i}{n+1}\right] F'_{\sigma}(x_{(i)};\mu,\sigma) = 0, \quad (24)$$

where $F'_{\mu}(x_{(i)};\mu,\sigma)$ and $F'_{\sigma}(x_{(i)};\mu,\sigma)$ are defined as in (19) and (20), respectively.

Step 4. Solve the Equations (23) and (24) simultaneously to obtain the WLS estimators $\hat{\mu}_{WLS}$ and $\hat{\sigma}_{WLS}$ for parameters μ and σ, respectively.

Remark. As in the ML framework, there are no closed form solutions for LS and WLS estimators due to the nonlinear structure of the estimating equations. Therefore, we use iterative numerical algorithms to obtain the LS and WLS estimates of the parameters μ and σ.

4 Monte Carlo Simulation

In this section, efficiencies of the ML, MML, LS, and WLS estimators of unknown parameters μ and σ are compared via an extensive Monte Carlo simulation study using R software program. The comparisons are evaluated by using the bias and mean-squared error (MSE) criteria. Simulated bias and MSE for the estimator of the parameter of interest θ are calculated by using the following equalities

$$\text{Bias}(\hat{\theta}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\theta}_i - \theta)$$
$$\text{MSE}(\hat{\theta}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\theta}_i - \theta)^2.$$

Here, $\hat{\theta}_i$ represents the estimate of the parameter θ obtained in the i-th replication of the simulation. Moreover, $M = \left[\!\left[\frac{100{,}000}{n}\right]\!\right]$ denotes the Monte Carlo replication number, where $[\![.]\!]$ represents the greatest integer function.

The joint efficiencies of the estimators $\hat{\mu}$ and $\hat{\sigma}$ are also evaluated by using deficiency (Def) criterion

$$\text{Def} = \text{MSE}(\hat{\mu}) + \text{MSE}(\hat{\sigma}).$$

To provide a more compact comparison between ML, MML, LS and WLS estimators, relative efficiencies (RE) defined below are also given

$$\text{RE}_1 = \frac{\text{MSE}(\hat{\mu}_{ML})}{\text{MSE}(\hat{\mu}_j)}, \qquad \text{RE}_2 = \frac{\text{MSE}(\hat{\sigma}_{ML})}{\text{MSE}(\hat{\sigma}_j)}.$$

Here, j represents the estimation method of interest, i.e., MML, LS or WLS.

In the simulation setup, sample sizes are taken as $n = 10, 20, 50, 100, 300, 500$. M independent samples of size n are generated from the KumGum distribution with parameter settings $(a, b, \mu, \sigma) =$ (0.1,2.5,0,1), (1,1,0,1), (0.2,0.9,0,1) and (2.5,0.5,0,1). Simulated bias, variance, MSE for the

estimators of parameters μ and σ, together with the corresponding Def values are presented in Tables 2-5.

Table 2. Simulated bias, variance, MSE, Def and RE values for the estimators of parameters μ and σ when $(a, b, \mu, \sigma) = (0.1,2.5,0,1)$.

	$\hat{\mu}$			$\hat{\sigma}$					
Estimation Methods	Bias	Var	MSE	Bias	Var	MSE	Def	RE_1	RE_2
				$\boldsymbol{n = 10}$					
ML	-0.1135	0.4331	0.4460	-0.0464	0.0452	0.0474	0.4934	1.0000	1.0000
MML	-0.0183	0.5740	0.5743	-0.0202	0.0606	0.0610	0.6354	0.7766	0.7770
LS	0.2792	0.9207	0.9987	0.1014	0.1091	0.1194	1.1180	0.4466	0.3970
WLS	0.2392	0.8260	0.8832	0.0863	0.0960	0.1034	0.9867	0.5050	0.4584
				$\boldsymbol{n = 20}$					
ML	-0.0629	0.2341	0.2380	-0.0255	0.0245	0.0251	0.2632	1.0000	1.0000
MML	-0.0081	0.2702	0.2703	-0.0105	0.0285	0.0286	0.2988	0.8806	0.8776
LS	0.1282	0.4009	0.4173	0.0459	0.0466	0.0487	0.4660	0.5699	0.5154
WLS	0.0875	0.3495	0.3571	0.0308	0.0397	0.0406	0.3977	0.6666	0.6182
				$\boldsymbol{n = 50}$					
ML	-0.0249	0.0970	0.0976	-0.0101	0.0100	0.0101	0.1077	1.0000	1.0000
MML	-0.0075	0.1072	0.1072	-0.0058	0.0111	0.0112	0.1184	0.9104	0.9018
LS	0.0543	0.1547	0.1577	0.0202	0.0175	0.0179	0.1756	0.6190	0.5642
WLS	0.0240	0.1278	0.1284	0.0087	0.0139	0.0140	0.1424	0.7601	0.7214
				$\boldsymbol{n = 100}$					
ML	-0.0143	0.0516	0.0518	-0.0047	0.0053	0.0053	0.0571	1.0000	1.0000
MML	-0.0041	0.0541	0.0542	-0.0020	0.0056	0.0056	0.0598	0.9557	0.9464
LS	0.0316	0.0759	0.0769	0.0127	0.0086	0.0088	0.0857	0.6736	0.6023
WLS	0.0092	0.0627	0.0628	0.0043	0.0068	0.0068	0.0696	0.8248	0.7794
				$\boldsymbol{n = 300}$					
ML	0.0000	0.0162	0.0162	-0,0009	0,0016	0.0016	0,0178	1.0000	1.0000
MML	0.0021	0.0170	0.0170	-0,0005	0,0017	0.0017	0,0187	0.9529	0.9412
LS	0.0132	0.0244	0.0246	0,0044	0,0026	0.0027	0,0273	0.6585	0.5926
WLS	0.0040	0.0197	0.0197	0,0009	0,002	0.0020	0,0218	0.8228	0.8000
				$\boldsymbol{n = 500}$					
ML	-0.0083	0.0097	0.0097	-0.0024	0.0010	0.0010	0.0107	1.0000	1.0000
MML	-0.0094	0.0104	0.0105	-0.0030	0.0011	0.0011	0.0116	0.9238	0.9091
LS	0.0028	0.0132	0.0132	0.0020	0.0015	0.0015	0.0147	0.7348	0.6667
WLS	-0.0039	0.0113	0.0113	-0.0006	0.0012	0.0012	0.0125	0.8584	0.8333

Table 3. Simulated bias, variance, MSE, Def and RE values for the estimators of parameters μ and σ when $(a, b, \mu, \sigma) = (1, 1, 0, 1)$.

	$\hat{\mu}$			$\hat{\sigma}$					
Estimation Methods	Bias	Var	MSE	Bias	Var	MSE	Def	RE_1	RE_2
$n = 10$									
ML	0.0396	0.1181	0.1197	-0.0418	0.0494	0.0512	0.1708	1.0000	1.0000
MML	0.0984	0.1217	0.1314	-0.0221	0.0700	0.0705	0.2019	0.9110	0.7262
LS	-0.0088	0.1248	0.1248	0.0988	0.1216	0.1313	0.2562	0.9591	0.3899
WLS	-0.0035	0.1212	0.1212	0.0843	0.1080	0.1151	0.2363	0.9876	0.4448
$n = 20$									
ML	0.0292	0.0566	0.0574	-0.0254	0.0295	0.0301	0.0876	1.0000	1.0000
MML	0.0601	0.0575	0.0612	-0.0095	0.0336	0.0337	0.0949	0.9379	0.8932
LS	0.0033	0.0613	0.0614	0.0512	0.0539	0.0565	0.1178	0.9349	0.5327
WLS	0.0084	0.0589	0.0590	0.0356	0.0461	0.0474	0.1064	0.9729	0.6350
$n = 50$									
ML	0.0038	0.0225	0.0225	-0.0143	0.0117	0.0119	0.0344	1.0000	1.0000
MML	0.0169	0.0228	0.0231	-0.0078	0.0123	0.0124	0.0354	0.9740	0.9597
LS	-0.0051	0.0249	0.0249	0.0186	0.0200	0.0203	0.0452	0.9036	0.5862
WLS	-0.0024	0.0236	0.0236	0.0071	0.0162	0.0162	0.0398	0.9534	0.7346
$n = 100$									
ML	0.0053	0.0116	0.0116	-0.0101	0.0057	0.0058	0.0174	1.0000	1.0000
MML	0.0119	0.0117	0.0118	-0.0079	0.0060	0.0061	0.0179	0.9831	0.9508
LS	0.0008	0.0125	0.0125	0.0056	0.0096	0.0097	0.0221	0.9280	0.5979
WLS	0.0026	0.0119	0.0119	-0.0020	0.0076	0.0076	0.0195	0.9748	0.7632
$n = 300$									
ML	-0.0020	0.0034	0.0034	0.0025	0.0016	0.0016	0.0050	1.0000	1.0000
MML	0.0002	0.0034	0.0034	0.0026	0.0017	0.0017	0.0051	1.0000	0.9412
LS	-0.0018	0.0039	0.0039	0.0023	0.0027	0.0028	0.0066	0.8718	0.5714
WLS	-0.0022	0.0036	0.0036	0.0013	0.0021	0.0021	0.0056	0.9444	0.7619
$n = 500$									
ML	0.0010	0.0022	0.0022	-0.0006	0.0011	0.0011	0.0033	1.0000	1.0000
MML	0.0024	0.0022	0.0022	-0.0004	0.0011	0.0011	0.0033	1.0000	1.0000
LS	0.0012	0.0026	0.0026	0.0008	0.0017	0.0017	0.0042	0.8462	0.6471
WLS	0.0011	0.0024	0.0024	-0.0005	0.0013	0.0013	0.0036	0.9167	0.8462

Table 4. Simulated bias, variance, MSE, Def and RE values for the estimators of parameters μ and σ when $(a, b, \mu, \sigma) = (0.2, 0.9, 0, 1)$.

	$\hat{\mu}$			$\hat{\sigma}$					
Estimation Methods	Bias	Var	MSE	Bias	Var	MSE	Def	RE_1	RE_2
$n = 10$									
ML	-0.0123	0.3156	0.3157	-0.0360	0.0515	0.0528	0.3685	1.0000	1.0000
MML	0.0796	0.4070	0.4133	-0.0186	0.0708	0.0712	0.4845	0.7639	0.7416
LS	0.1565	0.4887	0.5132	0.1065	0.1254	0.1367	0.6499	0.6152	0.3862
WLS	0.1386	0.4637	0.4829	0.0912	0.1113	0.1196	0.6025	0.6538	0.4415
$n = 20$									
ML	-0.0093	0.1681	0.1682	-0.0204	0.0273	0.0277	0.1959	1.0000	1.0000
MML	0.0354	0.1939	0.1952	-0.0125	0.0330	0.0332	0.2284	0.8617	0.8343
LS	0.0732	0.2319	0.2373	0.0520	0.0552	0.0579	0.2952	0.7088	0.4784
WLS	0.0536	0.2143	0.2172	0.0354	0.0463	0.0475	0.2647	0.7744	0.5832
$n = 50$									
ML	-0.0039	0.0650	0.0650	-0.0090	0.0107	0.0107	0.0757	1.0000	1.0000
MML	0.0140	0.0702	0.0704	-0.0064	0.0118	0.0119	0.0822	0.9233	0.8992
LS	0.0250	0.0819	0.0826	0.0164	0.0186	0.0189	0.1014	0.7869	0.5661
WLS	0.0116	0.0751	0.0752	0.0058	0.0151	0.0151	0.0903	0.8644	0.7086
$n = 100$									
ML	0.0024	0.0338	0.0338	-0.0055	0.0056	0.0056	0.0394	1.0000	1.0000
MML	0.0095	0.0357	0.0358	-0.0051	0.0060	0.0060	0.0418	0.9441	0.9333
LS	0.0194	0.0429	0.0432	0.0092	0.0094	0.0095	0.0528	0.7824	0.5895
WLS	0.0096	0.0388	0.0389	0.0015	0.0074	0.0074	0.0463	0.8689	0.7568
$n = 300$									
ML	-0.0001	0.0119	0.0119	0.0003	0.0020	0.0020	0.0139	1.0000	1.0000
MML	0.0009	0.0123	0.0123	-0.0004	0.0022	0.0022	0.0145	0.9675	0.9091
LS	0.0071	0.0143	0.0144	0.0065	0.0032	0.0033	0.0176	0.8264	0.6061
WLS	0.0016	0.0132	0.0132	0.0025	0.0026	0.0026	0.0158	0.9015	0.7692
$n = 500$									
ML	-0.0038	0.0068	0.0068	-0.0041	0.0011	0.0012	0.0079	1.0000	1.0000
MML	-0.0023	0.0069	0.0069	-0.0040	0.0012	0.0012	0.0081	0.9855	1.0000
LS	0.0000	0.0086	0.0086	0.0003	0.0018	0.0018	0.0104	0.7907	0.6667
WLS	-0.0024	0.0076	0.0076	-0.0022	0.0014	0.0014	0.0090	0.8947	0.8571

Table 5. Simulated bias, variance, MSE, Def and RE values for the estimators of parameters μ and σ when $(a, b, \mu, \sigma) = (2.5, 0.5, 0, 1)$.

	$\hat{\mu}$			$\hat{\sigma}$					
Estimation Methods	Bias	Var	MSE	Bias	Var	MSE	Def	RE_1	RE_2
$n = 10$									
ML	0.1678	0.2216	0.2497	-0.0680	0.0627	0.0673	0.3171	1.0000	1.0000
MML	0.2337	0.2441	0.2987	-0.0253	0.0763	0.0769	0.3756	0.8360	0.8752
LS	-0.1657	0.4595	0.4870	0.1123	0.1401	0.1527	0.6397	0.5127	0.4407
WLS	-0.1332	0.3921	0.4099	0.0962	0.1236	0.1328	0.5427	0.6092	0.5068
$n = 20$									
ML	0.0891	0.1106	0.1185	-0.0315	0.0319	0.0329	0.1513	1.0000	1.0000
MML	0.1253	0.1161	0.1318	-0.0140	0.0365	0.0367	0.1685	0.8991	0.8965
LS	-0.0805	0.2096	0.2161	0.0570	0.0614	0.0647	0.2808	0.5484	0.5085
WLS	-0.0452	0.1726	0.1746	0.0396	0.0521	0.0537	0.2283	0.6787	0.6127
$n = 50$									
ML	0.0352	0.0463	0.0475	-0.0101	0.0141	0.0142	0.0618	1.0000	1.0000
MML	0.0504	0.0466	0.0491	-0.0038	0.0148	0.0148	0.0640	0.9674	0.9595
LS	-0.0367	0.0758	0.0771	0.0278	0.0235	0.0243	0.1014	0.6161	0.5844
WLS	-0.0134	0.0608	0.0610	0.0153	0.0191	0.0194	0.0804	0.7787	0.7320
$n = 100$									
ML	0.0156	0.0225	0.0227	-0.0060	0.0068	0.0069	0.0296	1.0000	1.0000
MML	0.0243	0.0224	0.0230	-0.0031	0.0070	0.0070	0.0300	0.9870	0.9857
LS	-0.0174	0.0380	0.0383	0.0112	0.0110	0.0112	0.0494	0.5927	0.6161
WLS	-0.0039	0.0290	0.0290	0.0043	0.0089	0.0089	0.0379	0.7828	0.7753
$n = 300$									
ML	0.0078	0.0063	0.0063	-0.0037	0.0018	0.0018	0.0082	1.0000	1.0000
MML	0.0116	0.0063	0.0065	-0.0035	0.0019	0.0019	0.0084	0.9692	0.9474
LS	0.0011	0.0115	0.0115	-0.0018	0.0033	0.0033	0.0148	0.5478	0.5455
WLS	0.0050	0.0085	0.0085	-0.0035	0.0025	0.0025	0.0110	0.7412	0.7200
$n = 500$									
ML	0.0016	0.0041	0.0041	0.0007	0.0013	0.0013	0.0055	1.0000	1.0000
MML	0.0043	0.0042	0.0042	0.0003	0.0014	0.0014	0.0056	0.9762	0.9286
LS	-0.0087	0.0061	0.0061	0.0054	0.0025	0.0025	0.0087	0.6721	0.5200
WLS	-0.0034	0.0045	0.0046	0.0030	0.0019	0.0019	0.0065	0.8913	0.6842

The following comments can be made regarding simulation results.

In the context of* bias*: ML, MML, LS and WLS estimators for μ and σ have negligible biases for moderate and large sample sizes ($n = 50, 100, 300, 500$) across all parameter settings. For small sample sizes ($n = 10, 20$), there is no estimator which is the best across all parameter settings.

***In the context of* MSE, Def *and* RE:** ML estimators for both μ and σ have the best overall performance in terms of MSE and Def across all parameter settings and sample sizes. They are closely followed by MML estimators, especially for moderate and large sample sizes ($n = 50, 100, 300, 500$) as expected. LS estimators exhibit the worst performance among the others. The WLS estimators improve upon the LS estimators but are not as efficient as the ML and MML estimators.
When the REs for μ and σ are examined, RE_1 and RE_2 values for the MML estimators exceed 0.90 for moderate and large sample sizes ($n = 50, 100, 300, 500$). This is another indication of the superiority of MML estimators over the corresponding LS and WLS estimators.

5 Application

In this section, we analyze the wave height data set consisting of 2,894 observations on wave heights (in meters) measured at a single offshore location in south-west England to illustrate the modeling performance of KumGum distribution. This data set was originally given by Coles et. al. (2001) and it is available in the ismev package in R software.

As an initial step, Q-Q (Quantile-Quantile) plot is drawn in order to visually assess whether KumGum distribution provides adequate fits to the wave height data set, see Figure 2. The theoretical quantiles are computed using ML estimates of the parameters of the KumGum distribution. The choice of the ML estimates is motivated by the simulation results since the ML estimators have the highest efficiencies compared to other estimators. Simulation results also show that efficiencies of the ML and MML estimators are close to each other for moderate and large sample sizes as the theory says.

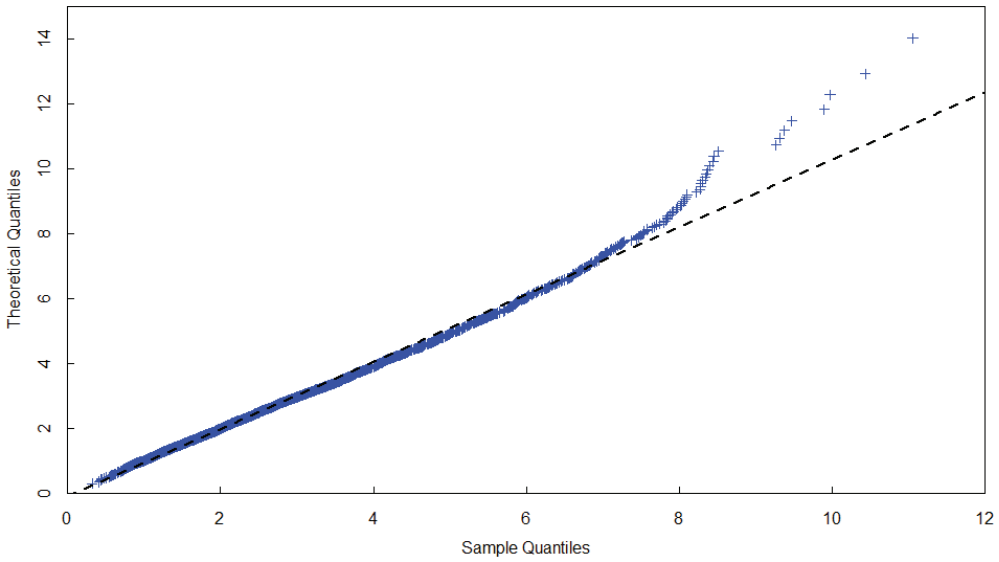


Figure 2. KumGum Q-Q plot for wave height data set.

KumGum Q-Q plot in Figure 2 shows a good match between the sample quantiles and theoretical quantiles therefore, KumGum distribution provides adequate fit to the wave height data set.
After showing the adequacy of the KumGum distribution for this data set, the histogram of the wave height data set together with the fitted pdfs of the KumGum distribution based on ML, MML, LS and WLS estimates are given in Figure 3.

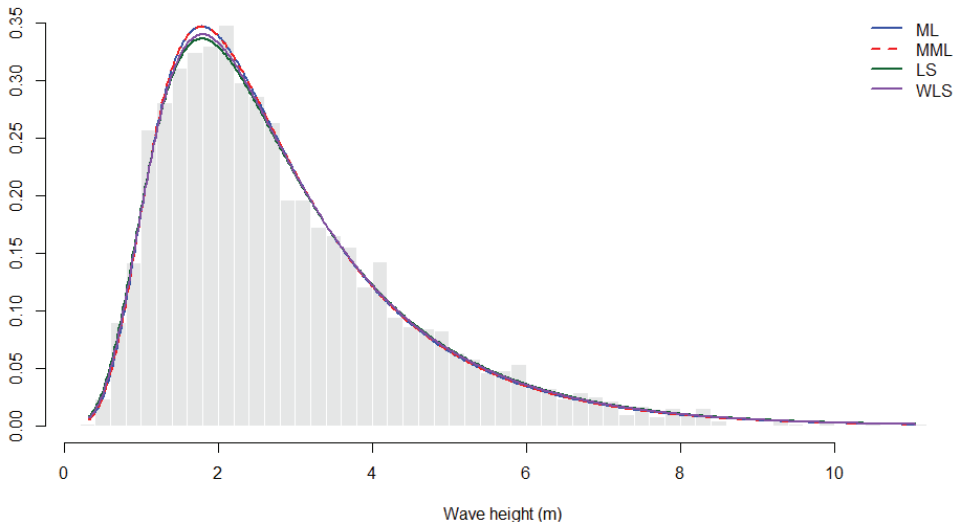


Figure 3. Histogram of the wave height data set together with the pdfs of the KumGum distribution based on ML, MML, LS and WLS estimates.

Figure 3 shows that ML and MML based fitted pdfs exhibit superior agreement with the histogram of the data set compared to LS and WLS based fitted pdfs. To validate these visual findings Akaike information criterion (AIC) and Bayesian information criterion (BIC) are also used. These criteria are defined as

$$\text{AIC} = -2\ln L + 2k \quad \text{and} \quad \text{BIC} = -2\ln L + k\ln n,$$

where k is the number of the unknown parameters. Parameter estimates for parameters μ and σ obtained using the ML, MML, LS, and WLS methods, AIC and BIC values are given in Table 6.

Table 6. Estimate values for the parameters μ and σ, AIC and BIC values.

Estimation Methods	$\hat{\mu}$	$\hat{\sigma}$	AIC	BIC
ML	2.484110	0.574250	10009.82	10021.77
MML	2.484658	0.574361	10009.83	10021.77
LS	2.506926	0.592154	10014.21	10026.15
WLS	2.501643	0.585629	10011.47	10023.41

Table 6 indicates that AIC and BIC values for ML and MML methods are almost the same with each other and they are smaller than those of the LS and WLS methods. LS method exhibits the poorest performance among the others. These findings agree with the simulation results. Therefore, ML and MML methods outperform the LS and WLS methods.

For benchmarking purposes and assessing the flexibility of the KumGum distribution, the pdfs of the KumGum and Gumbel distributions based on ML estimates are also fitted to the wave height data set.

It should be remembered that ML and MML methods yield nearly identical parameter estimates for large n values (in our case n=2,894). Therefore, ML-based fitted densities are used in the graphical presentation to avoid visual redundancy and confusion, see Figure 4.

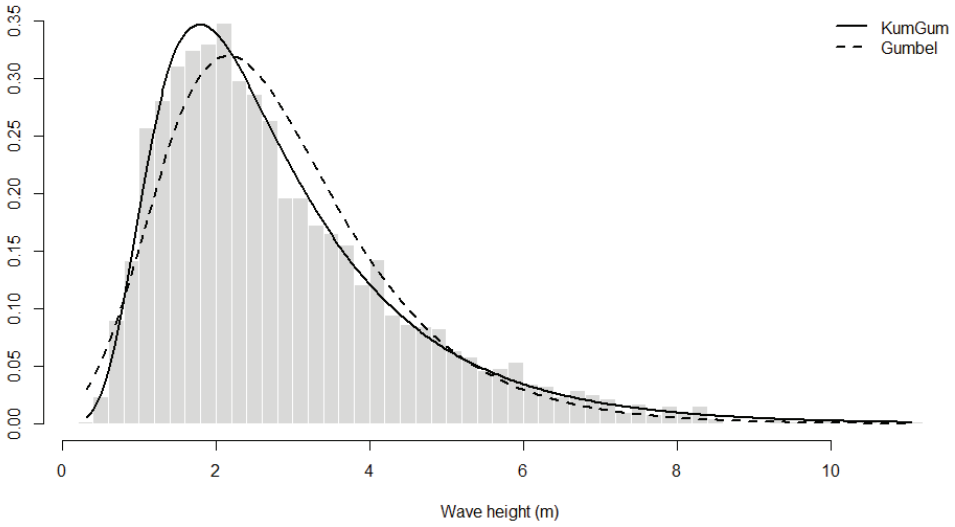


Figure 4. Histogram of the wave height data set together with the pdfs of the KumGum and Gumbel distributions based on ML estimates.

Figure 4 shows that the KumGum distribution provides a noticeably better fit to the wave height data set than the traditional Gumbel distribution. Moreover, ML estimates, $\ln L$, AIC and BIC values and calculated values of the Kolmogorov-Smirnov (KS) statistics for the KumGum and Gumbel distributions are reported in Table 7.

Table 7. ML estimates for the parameters μ and σ and values of the $\ln L$, AIC, BIC and KS statistics for the KumGum and Gumbel distributions.

	$\hat{\mu}$	$\hat{\sigma}$	$\ln L$	AIC	BIC	KS
KumGum	2.484110	0.57425	-5002.912	10013.82	10037.71	0.018585
Gumbel	2.162453	1.14946	-5068.683	10141.37	10153.31	0.044266

Table 7 shows that the KumGum distribution yields substantially higher log likelihood and lower AIC and BIC values compared to Gumbel distribution. In addition, calculated value of the KS statistics for KumGum distribution is considerably smaller than that of Gumbel distribution. These findings indicate that the wave height data set is more adequately modeled by the KumGum distribution compared to Gumbel distribution.

6 Concluding Remarks

In this study, ML, MML, LS and WLS estimators for the unknown location and scale parameters of the KumGum distribution are obtained. Efficiencies of the ML, MML, LS and WLS estimators are compared under different sample sizes and parameter settings via an extensive Monte Carlo simulation study. To provide a comprehensive comparison of the efficiencies of the estimators, sample sizes are chosen to represent small, moderate and large samples and parameter settings are

determined to produce KumGum distributions with gradually increasing positive skewness and leptokurtosis. According to simulation results, ML estimators have the highest performance, and they are followed by MML estimators among the methods considered in this study. As expected, ML and MML estimators have similar performance in terms of MSE and Def due to their asymptotic properties, especially for moderate and large sample sizes. WLS estimators are more efficient than the LS estimators. MSE and Def values decrease for all estimators of parameters μ and σ as the sample size increases.

Modeling performances of the KumGum and Gumbel distributions are also compared using wave height data set, given the fundamental importance of wave height in offshore engineering. Based on the results of the real data analysis, KumGum distribution provides a better fit to the wave height data set when compared to Gumbel distribution, which indicates that KumGum distribution offers greater flexibility compared to the baseline distribution. Also, KumGum distribution is well suited for the data sets with positive skewness and leptokurtosis as observed in the wave height data set.

References

Aydın, D., & Şenoğlu, B. (2015). Monte Carlo comparison of the parameter estimation methods for the two-parameter Gumbel distribution. *Journal of Modern Applied Statistical Methods*, *14*(2), 123–140.

Barnett, V. D. (1966). Order statistics estimators of the location of the Cauchy distribution. *Journal of the American Statistical Association*, *61*(316), 1205–1218.

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208). London, UK: Springer.

Cooray, K. (2010). Generalized Gumbel distribution. *Journal of Applied Statistics*, *37*(1), 171–179.
Cordeiro, G. M., Nadarajah, S., & Ortega, E. M. (2012). The Kumaraswamy Gumbel distribution. *Statistical Methods & Applications*, *21*(2), 139–168.

Demirhan, H. (2018). A generalized Gumbel distribution and its parameter estimation. *Communications in Statistics—Simulation and Computation*, *47*(10), 2829–2848.

Gómez, Y. M., Bolfarine, H., & Gómez, H. W. (2019). Gumbel distribution with heavy tails and applications to environmental data. *Mathematics and Computers in Simulation*, *157*, 115–129.

Gumbel, E. J. (1941). The return period of flood flows. *The Annals of Mathematical Statistics*, *12*(2), 163–190.

Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, *6*(1), 70–81.

Kang, D., Ko, K., & Huh, J. (2015). Determination of extreme wind values using the Gumbel distribution. *Energy*, *86*, 51–58.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, *46*(1–2), 79–88.

Liu, Q., Huang, X., & Zhou, H. (2024). The flexible Gumbel distribution: A new model for inference about the mode. *Stats*, *7*(1), 317–332.

Nadarajah, S. (2006). The exponentiated Gumbel distribution with climate application. *Environmetrics*, *17*(1), 13–23.

Nadarajah, S., & Kotz, S. (2004). The beta Gumbel distribution. *Mathematical Problems in Engineering*, *2004*(4), 323–332.

Pérez-Rodríguez, P., Vaquera-Huerta, H., & Villaseñor-Alva, J. A. (2009). A goodness-of-fit test for the Gumbel distribution based on Kullback–Leibler information. *Communications in Statistics—Theory and Methods*, *38*(6), 842–855.

Pinheiro, E. C., & Ferrari, S. L. (2016). A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data. *Journal of Statistical Computation and Simulation*, *86*(11), 2241–2261.

Purohit, S. U., & Lalit, P. N. (2022). European option pricing using Gumbel distribution. *International Journal of Financial Engineering*, *9*(1), 2141002.

Puthenpura, S., & Sinha, N. K. (1986). Modified maximum likelihood method for the robust estimation of system parameters from very noisy data. *Automatica*, *22*(2), 231–235.

Solomon, O., & Prince, O. (2013). Flood frequency analysis of Osse River using Gumbel's distribution. *Civil and Environmental Research*, *3*(10), 55–59.

Tiku, M. L. (1967). Estimating the mean and standard deviation from a censored normal sample. *Biometrika*, *54*(1–2), 155–165.

Tiku, M. L. (1968). Estimating the parameters of log-normal distribution from censored samples. *Journal of the American Statistical Association*, *63*(321), 134–140.

Vaughan, D. C. (1992). On the Tiku–Suresh method of estimation. *Communications in Statistics—Theory and Methods*, *21*(2), 451–469.

Chapter 6

DETERMINATION OF PHOTON INTERACTIONS AND TOTAL MASS ATTENUATION COEFFICIENTS FOR SOME TRANSITION ELEMENTS

Kadir GÜNOĞLU[1]

1 Doç. Dr., Isparta Uygulamalı Bilimler Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, kadirgunoglu@isparta.edu.tr, ORCID: 0000-0002-9008-9162.

1. INTRODUCTION

The increasing application of gamma radiation across various domains—including industrial processes, medical diagnostics and therapy, agricultural practices, nuclear reactors, and particle accelerator technologies—has raised concerns regarding the potential health hazards associated with long-term exposure to high-energy photons (Durante and Cucinotta, 2011; Zou, 2017; Gunoglu et al., 2021) For this reason, the development and application of effective radiation shielding strategies are of paramount importance to ensure safe operational conditions. While numerous materials can be utilized for gamma-ray attenuation, the selection of an appropriate shielding material is strongly dependent on the energy of the incident photons (Dur´an et al., 2013; Lakhwani et al., 2019; Abu Bakar et al., 2019; Gunoglu et al., 2024). The attenuation capability of a given material is primarily determined by the nature of the interactions between gamma photons and the atoms within the shielding medium. These interaction processes are governed not only by photon energy but also by the material's elemental composition and density (Içelli and Erzeneoglu, 2004; Abu Al Roos et al., 2019).. Accordingly, a comprehensive understanding of photon–matter interaction mechanisms is essential, as it forms the theoretical basis for assessing shielding performance and optimizing the design of radiation protection systems.

Understanding how photons interact with matter constitutes a fundamental aspect of radiation physics and is essential for applications ranging from nuclear engineering and medical diagnostics to high-energy particle technologies. When X-rays or gamma photons penetrate a material, they may undergo three primary interaction mechanisms—photoelectric absorption (PE), Compton scattering (CS), and pair production (PP)—each of which is governed largely by the photon's energy and the atomic characteristics of the absorber, such as atomic number, electron configuration, and nuclear field properties (Singh and Badiger, 2016).

PE predominates in the low-energy photon region and involves the complete annihilation of the incident photon within the atomic structure of the absorber. During this interaction, the photon transfers its entire energy to a tightly bound inner-shell electron, causing the ejection of that electron from the atom as a photoelectron. For this mechanism to occur, the incident photon energy must exceed the binding energy of the targeted atomic shell. The likelihood of PE exhibits a strong dependence on the atomic number of the absorbing material, approximately scaling with Z^4 to Z^5, which results in significantly enhanced absorption in high-Z materials. In contrast, substances composed of low-Z elements show a considerably reduced probability for this interaction. Additionally, the photoelectric cross-section is characterized by abrupt increases, known as absorption edges, at photon energies matching the binding energies of the K, L, and M electron shells. These shell-specific features give rise to distinct and material-dependent interaction behaviors (Fern`andez and Scot 2007; Singh and Badiger, 2016).

Scattering interactions represent processes in which the photon is not fully absorbed but is deflected from its original trajectory, either with or without a change in energy. Two principal scattering mechanisms can occur depending on the photon energy and the electronic characteristics of the medium:

1. Coherent (Rayleigh) scattering is an elastic interaction in which the photon is deflected by the collective electric field of the atomic electron cloud without losing energy. This mechanism is more significant for low-energy photons and in atoms with high atomic numbers, where the dense electron cloud and stronger nuclear field increase the probability of elastic deflection (Fern`andez and Scot 2007).

2. Incoherent scattering, commonly referred to as Compton scattering, arises from the interaction between an incident photon and a weakly bound or outer-shell electron. During this process, only a fraction of the photon's energy is imparted to the electron, while the scattered photon emerges with reduced energy and is deflected from its original trajectory. This interaction mechanism is predominant in the intermediate photon energy range, extending approximately from 100 keV to several MeV. Materials with moderate atomic numbers exhibit significant contributions from Compton scattering over wide energy intervals. The probability of this interaction is primarily governed by the electron density of the material, playing a critical role in the attenuation behavior, energy degradation, and angular redistribution of photons as they traverse the medium (Khan, 2003; Fern`andez and Scot 2007).

At sufficiently high photon energies, PP becomes an important interaction mechanism. When a photon with energy exceeding 1.022 MeV interacts with the intense electromagnetic field near an atomic nucleus, it can spontaneously convert into an electron–positron pair. The probability of this mechanism rises rapidly with both photon energy and atomic number, since the strength of the nuclear field plays a crucial role in facilitating the conversion process. As a result, high-Z materials exhibit substantial PP cross-sections at elevated energies, whereas low-Z elements require significantly higher photon energies before this process becomes relevant. PP thus provides a powerful means of distinguishing photon interaction behavior among elements, particularly in high-energy radiation environments (Khan, 2003).

A comprehensive understanding of these fundamental mechanisms enables the systematic analysis of how different elements respond to photons across varying energy regimes. In this study, the photon interaction characteristics of pure Fe, Cu, Ti, Mo, W, and Pb are examined with respect to PE, CS, and PP. The objective is to elucidate the energy-dependent transition between these mechanisms and highlight the underlying physical relationships associated with atomic number. Such an approach not only enriches the fundamental scientific understanding of photon–matter interactions but also provides valuable insights for selecting appropriate elements in the design of advanced composite materials for radiation shielding applications (Khan, 2003; Fern`andez and Scot 2007).

2. MATERIALS AND METHODS

This work presents a detailed investigation of photon–matter interaction processes for the elements Fe, Cu, Ti, Mo, W, and Pb. As a foundation for the analysis, mass attenuation coefficient data covering the full photon energy spectrum were retrieved from the XCOM database (Berger and Hubbell, 1987), which is internationally recognized as a standard reference for X-ray and gamma-ray interactions. The XCOM database provides comprehensive information on the total photon interaction probability, along with the individual contributions of the underlying interaction mechanisms, namely PE, CS, and PP, as functions of photon energy. Accordingly, the calculations in this study employed both the total mass attenuation coefficients (μ/ρ) and the fractional contributions of each interaction process for the selected elements within the specified energy range. The relative significance of these mechanisms is strongly influenced by photon energy, atomic electron configuration, atomic number (Z), and material density. The fundamental physical properties of the investigated elements, including atomic number, atomic weight, and density, are summarized in Table 1.

Table 1. The physical properties of the elements

Element	Atomic Number (Z)	Atomic Weight (g/mol)	Density (g/cm³)
Ti	22	47,867	4,51
Fe	26	55,845	7,87
Cu	29	63,546	8,96
Mo	42	95,95	10,28
W	74	183,84	19,25
Pb	82	207,2	11,34

Each mechanism produces different physical results that determine the energy transfer characteristics of the photon and the shielding properties of the material. Therefore, a detailed examination of the mathematical and physical foundations of these interactions is essential for understanding radiation transport.

2.2. The photoelectric absorption

The PE describes an interaction in which an incident photon is completely absorbed by an atom through its interaction with an inner-shell electron. In this process, the photon transfers its entire energy to the bound electron, leading to the emission of a photoelectron and the disappearance of the incident photon. This mechanism is most significant at low photon energies and may extend into the lower portion of the intermediate energy range, typically below approximately 100–200 keV (Patra et al., 2014). For the emission of a photoelectron to occur, the incident photon must possess an energy exceeding the binding energy (E_b) of the electron in the corresponding atomic shell. Based on the conservation of energy, the incident photon energy (E_γ) is equal to the sum of the kinetic energy of the emitted electron (E_e) and its binding energy (Fern`andez and Scot, 2007). This energy balance can be expressed by the following relation:

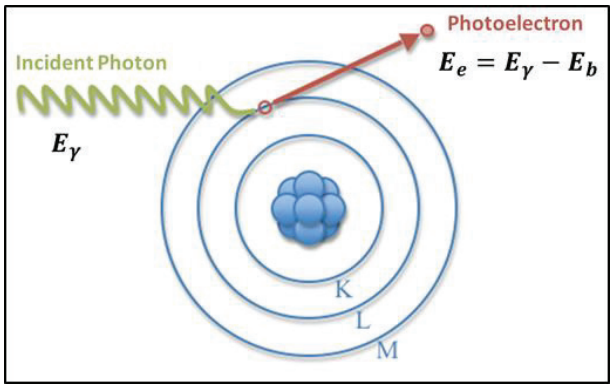
$$E_\gamma = E_b + E_e \tag{1}$$

Therefore, the kinetic energy of the photoelectron is obtained as follows:

$$E_e = E_\gamma - E_b \tag{2}$$

Following the ejection of an inner-shell electron, the resulting vacancy is occupied by an electron transitioning from a higher-energy shell (Mayles et al., 2008). This atomic relaxation process is accompanied by the emission of either characteristic X-ray photons or Auger electrons.

Figure 1. Schematic view of the photoelectric effect.



The likelihood of PE is strongly influenced by the atomic number of the absorbing material. An approximate expression describing the dependence of the PE coefficient can be written as follows: (Podgorsak 2005)

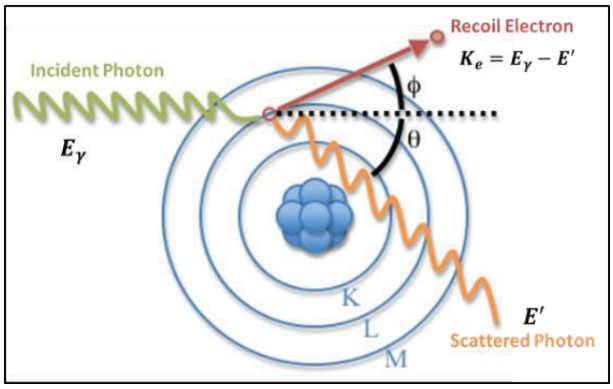
$$\left(\frac{\mu}{\rho}\right)_{pe} \propto \frac{Z^n}{E_\gamma^3} \tag{3}$$

and the value of n usually varies between 4 and 5. Due to this strong dependence on Z, materials with high atomic numbers such as lead (Z = 82) and tungsten (Z = 74) are extremely effective for shielding low-energy photons. The dependence on energy shows that the photoelectric effect decreases rapidly as the photon energy increases; therefore, this mechanism becomes negligible for high-energy photons (Khan 2003).

2.2. The Compton scattering (CS)

CS is an inelastic interaction in which an incident photon interacts with a free or weakly bound electron within an atom. During this process, a portion of the photon's energy is transferred to the electron, and the photon is deflected from its original path, emerging with reduced energy. A defining characteristic of CS is that the magnitude of energy transfer depends on both the initial photon energy and the scattering angle (Khan 2003).

Figure 2. Schematic view of the compton scattering.



The energy of the scattered photon is governed by the CS relation, commonly expressed as E′ (Podgorsak 2005).

$$E' = \frac{E_\gamma}{1+\frac{E_\gamma}{m_e c^2}(1-COS\theta)} \tag{4}$$

here, θ is the scattering angle of the photon, and $m_e c^2$ = 0.511 MeV is the rest mass energy of the electron. According to this relationship, the energy of the photon decreases as the scattering angle increases and reaches its minimum value for θ = 180°. The kinetic energy transferred to the electron is calculated as follows, based on the energy difference (Patra et al. 2014).

$$K_e = E_\gamma - E' \tag{5}$$

CS is the dominant type of interaction at photon energies between approximately 100 keV and a few MeV. The dependence of the absorption coefficient on the atomic number is weak and is approximately expressed as follows.

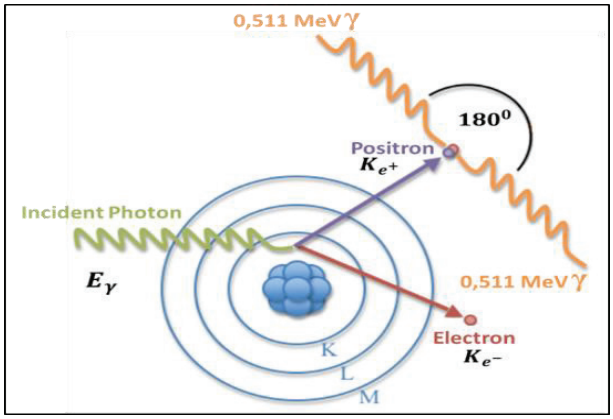
$$\left(\frac{\mu}{\rho}\right)_{cs} \propto \frac{Z}{E_\gamma} \tag{6}$$

Therefore, it has a significant effect even in materials with low and medium Z values. In particular, a large proportion of the scattering seen in medical imaging in low Z environments such as human tissue is due to CS.

2.3. The pair production (PP)

PP refers to an interaction in which a high-energy photon converts into an electron–positron pair in the vicinity of an atomic nucleus, provided that the photon energy exceeds a critical threshold. This threshold energy is 1.022 MeV, corresponding to twice the rest-mass energy of an electron (Khan, 2003).

Figure 3. Schematic view of the pair production.



The presence of the nuclear Coulomb field is essential in this process, as it enables the conservation of momentum during the interaction, which is why PP is classified as a nuclear interaction. Once the incident photon energy surpasses the threshold, the excess energy beyond 1.022 MeV is converted into the combined kinetic energy of the produced electron and positron, as described by the following relation.

$$K_{e^-} + K_{e^+} = E_\gamma - 1.022\ MeV \tag{7}$$

The resulting positron then slows down within the matter and undergoes annihilation upon encountering an electron; this process typically results in the production of two 511 keV gamma photons. PP is particularly prevalent in atoms with high Z values, and the absorption coefficient shows a dependence on photon energy approximately as follows (Mayles et al. 2008).

$$\left(\frac{\mu}{\rho}\right)_{pp} \propto Z\ln(E_\gamma) \tag{8}$$

2.4. The mass attenuation coefficient (MAC)

The attenuation behavior of photons within matter is characterized by the MAC (μ/ρ), which is defined as the sum of the total contributions of photon-matter interaction mechanisms. This quantity expresses the probability of interaction per unit mass and quantitatively describes the energy loss or reduction in photon flux that occurs during the passage of photons through matter. The MAC is directly related to the cross-sections of physical interaction processes and varies depending on photon energy, atomic number, material density, and elemental composition. Attenuation is a process in which photons interact with matter, losing all or part of their energy or changing direction, and consists primarily of contributions from three independent mechanisms: PE, CS, and PP. The contribution of each mechanism to attenuation becomes dominant in specific energy regions

depending on the photon energy. Therefore, the total MAC is expressed as the sum of the cross-sections of these three basic mechanisms (Khan 2003). The formula for the total MAC is as follows:

$$\left(\frac{\mu}{\rho}\right)_{total} = \left(\frac{\mu}{\rho}\right)_{pe} + \left(\frac{\mu}{\rho}\right)_{cs} + \left(\frac{\mu}{\rho}\right)_{pp} \quad (9)$$

here, $(\mu/\rho)_{pe}$ represents the PE mass coefficient, $(\mu/\rho)_{cs}$ the CS (incoherent scattering) mass coefficient, and $(\mu/\rho)_{pp}$ the PP mass coefficient. Since each of these three terms shows a different dependence on photon energy, the total attenuation coefficient exhibits a characteristic variation with energy. At low energies, photoelectric absorption is dominant, while at medium energies, CS, and at high energies, pair formation are the primary determinants of attenuation behavior. This behavior forms the fundamental link between microscopic cross-sections of photon interaction mechanisms and macroscopic attenuation parameters.

3. RESULTS AND DISCUSSION

Figures 4-9 present the variation of the partial mass attenuation coefficients of Fe, Ti, Cu, Mo, W, and Pb as a function of photon energy, together with the dominant photon–matter interaction mechanisms. The results clearly demonstrate a systematic transition of the governing interaction processes with increasing photon energy, as well as a strong dependence on the atomic number (Z) of the investigated transition elements.

Figure 4. The partial mass attenuation coefficient for iron

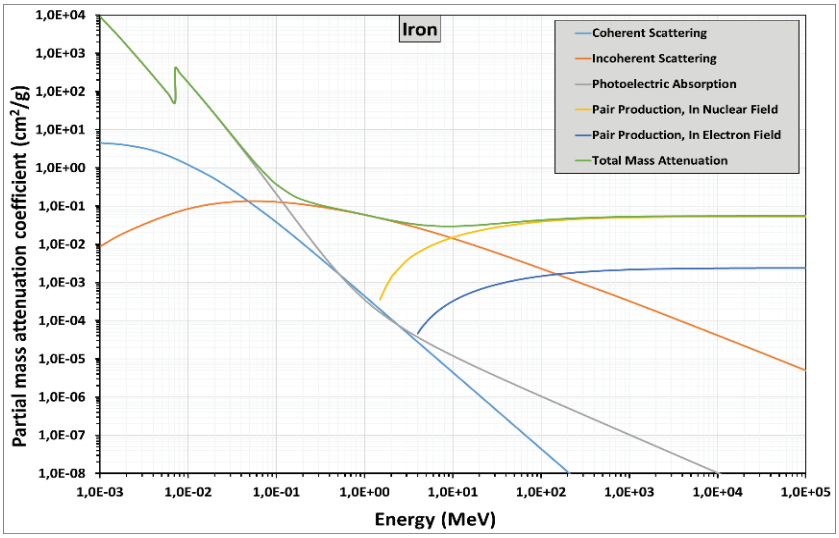


Figure 5. The partial mass attenuation coefficient for copper

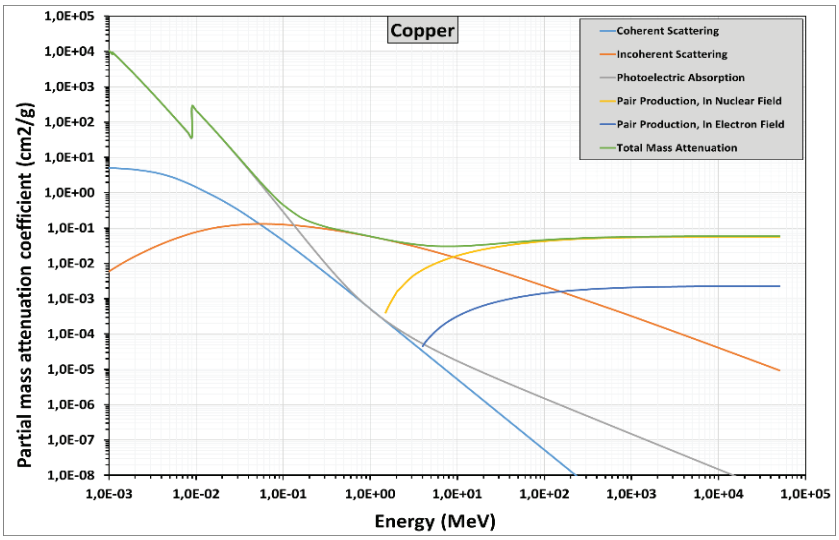


Figure 6.The partial mass attenuation coefficient for tungsten

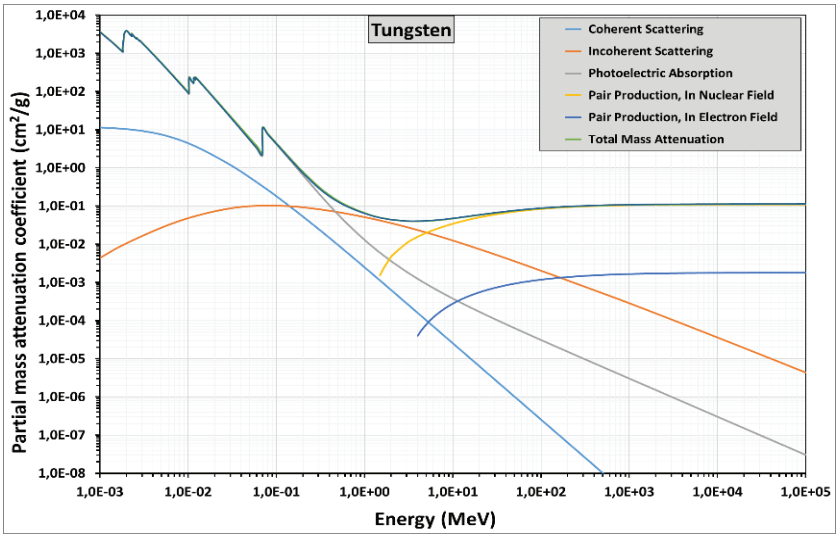


Figure 7. The partial mass attenuation coefficient for molybdenum

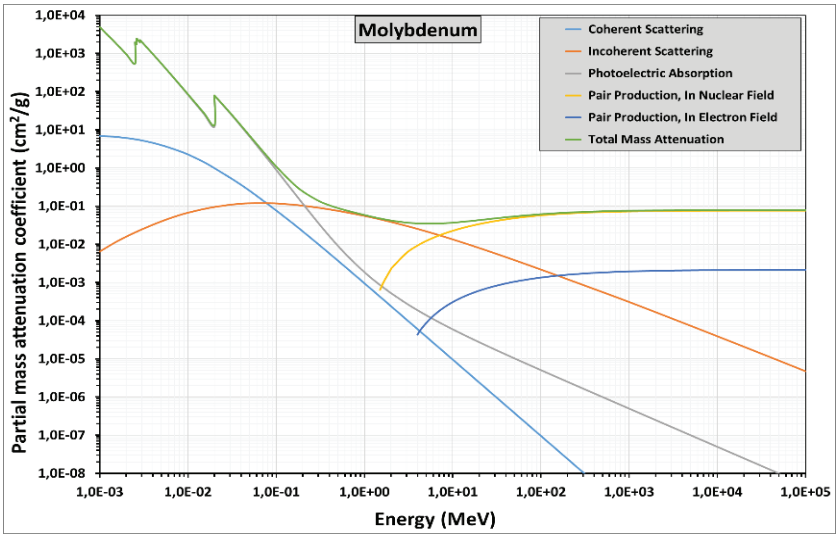


Figure 8. The partial mass attenuation coefficient for titanium

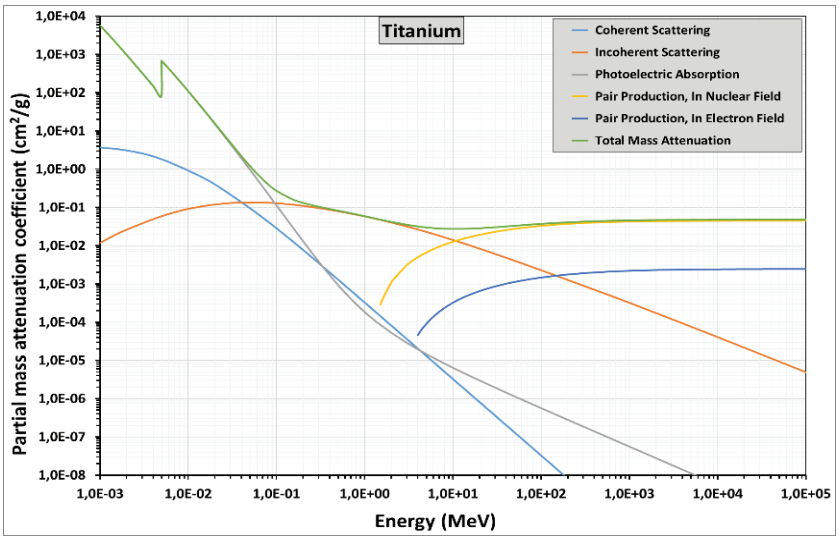
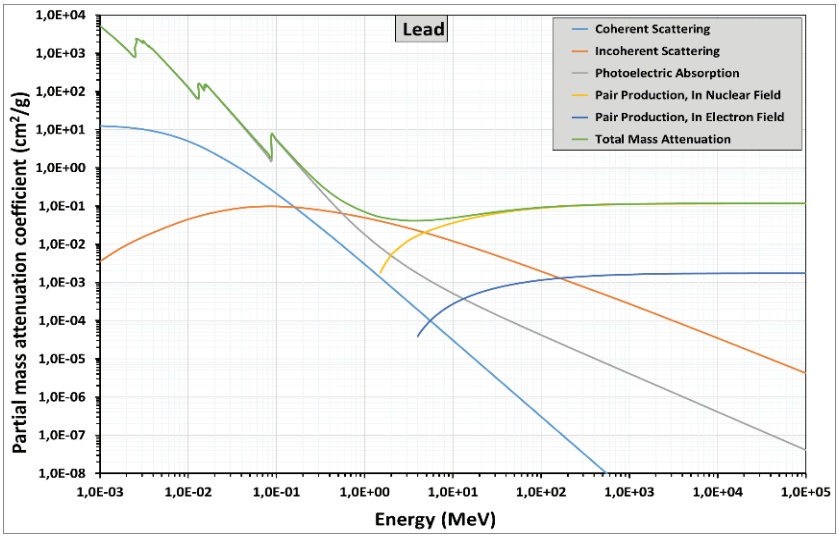


Figure 9. The partial mass attenuation coefficient for lead



When the overall structure of the graphs is examined, it becomes evident that, for all elements, the dominant photon–matter interaction mechanisms shift across distinct energy regions as the photon energy increases. In the low-energy photon regime (on the order of keV), the total mass attenuation coefficient reaches very high values for all elements, a behavior that can be primarily attributed to PE. The rapid decrease of the photoelectric contribution with increasing energy leads to a steep decline in the total attenuation curve at low energies. This decline becomes progressively less abrupt with increasing atomic number, and in high-Z elements such as W and Pb, the photoelectric effect remains dominant over a much broader energy range.

For elements with relatively low to intermediate atomic numbers, such as Fe, Ti, and Cu, PE is significant only in the very low-energy region and loses its importance rapidly beyond a few tens of keV. In contrast, for Mo, W, and especially Pb, the photoelectric contribution is not only larger in magnitude but also exhibits pronounced irregularities in the graphs in the form of abrupt jumps associated with K- and L-absorption edges. This behavior is directly related to the higher inner-shell binding energies in high-Z elements and to the strong dependence of the PE on the atomic number.

As the photon energy enters the intermediate energy range (approximately 0.1–1 MeV), incoherent (Compton) scattering becomes the dominant interaction mechanism for all elements. In the graphs, this transition is clearly reflected by the Compton component reaching a maximum and then gradually decreasing with increasing energy. Within this energy interval, the differences among elements diminish substantially, and the Compton contributions for Fe, Cu, Mo, W, and Pb converge toward similar values. This convergence arises from the weak dependence of CS on atomic number and its stronger dependence on electron density. Consequently, in the intermediate energy region, the total MAC shows a markedly reduced sensitivity to the choice of element.

Rayleigh (coherent) scattering appears in the graphs as a secondary contribution at low energies alongside the photoelectric effect. Its influence decreases very rapidly with increasing photon energy for all elements. Although Rayleigh scattering is somewhat more pronounced in high-Z materials, its contribution to the total attenuation becomes negligible in the intermediate and high-energy regions. For Fe, Ti, and Cu in particular, Rayleigh scattering plays only a minor role confined to the lowest energy range.

Once the photon energy exceeds the threshold of 1.022 MeV, the onset of PP is clearly observed in the graphs. This process is represented by PP occurring in both the nuclear and electron fields. With increasing energy, the contribution of PP rises continuously and begins to make a significant contribution to the total MAC. This increase becomes much more pronounced with increasing atomic number. For Pb and W, the pair production contribution becomes comparable to, or even exceeds, that of CS at energies of several MeV, whereas for low-Z elements such as Fe and Ti, the contribution of pair production remains relatively limited.

A comparison among the elements clearly indicates that Pb and W exhibit the highest total mass attenuation coefficients over the entire energy range. This behavior results from the combined effects of strong PE at low energies, CS with similar characteristics at intermediate energies, and the rapidly increasing contribution of PP at high energies. Mo displays an intermediate behavior between these heavy elements and the Fe–Cu group, reflecting the characteristic features of medium-to-high atomic number materials. Fe, Cu, and Ti, on the other hand, show similar trends particularly in the intermediate energy region, but exhibit significantly lower total attenuation values at both low and high energies compared to heavier elements.

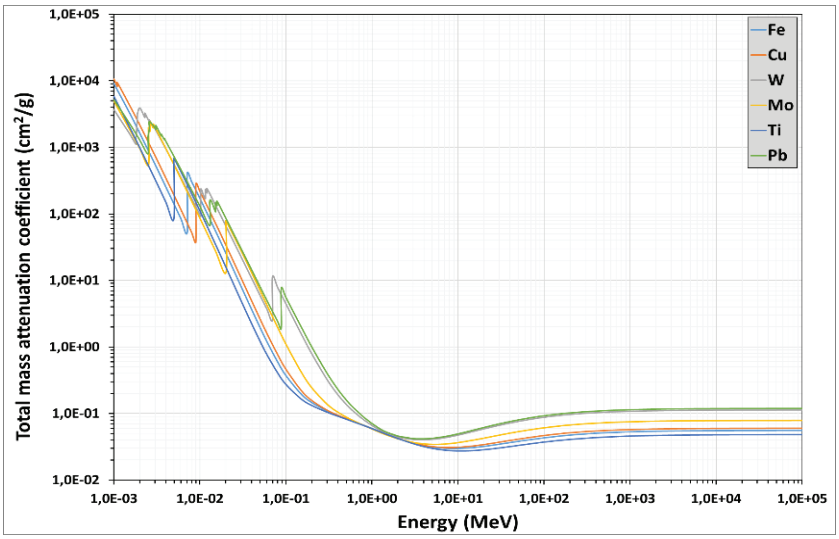
In summary, the graphs clearly demonstrate that photon–matter interactions are strongly energy dependent and that the atomic number plays a decisive role, particularly in PE, and PP processes. The pronounced superiority of heavy elements at low energies, the near mechanism-independent behavior at intermediate energies, and the renewed dominance of heavy elements at high energies together provide a coherent physical framework for interpreting the radiation interaction properties of these materials. This framework serves as a fundamental reference for predicting the behavior of different elements in photon fields spanning a wide range of energies.

The presented graph (Figure 10) illustrates the dependence of the total MAC on photon energy for selected transition and heavy metals, including Fe, Ti, Cu, Mo, W, and Pb. The overall trends in the curves reflect the prevailing photon–matter interaction mechanisms across different energy intervals, as well as their strong dependence on the atomic number (Z) of the interacting material.

At low photon energies (approximately 10^{-3}–10^{-2} MeV), all elements exhibit relatively high MAC values. In this energy range, PE is the dominant interaction process. Because the photoelectric cross-section follows an approximate proportionality of Z^n/E^3 (where n is typically between 4 and 5),

the absorption coefficient increases markedly with increasing atomic number. Consequently, high-Z elements such as tungsten and lead demonstrate substantially larger MAC values than lower-Z transition metals like iron and titanium. The sharp discontinuities observed in this region correspond to K- and L-shell absorption edges, which are particularly evident in heavier elements and arise from the sudden enhancement of photoelectric absorption once the photon energy exceeds the binding energies of inner-shell electrons.

Figure 10 The total mass attenuation coefficient for all elements



In the intermediate energy region (roughly 10^{-2}–1 MeV), the MAC values decrease rapidly with increasing photon energy. CS becomes the predominant interaction mechanism in this interval. Since the Compton cross-section depends primarily on electron density and exhibits only a weak dependence on atomic number, the variation in MAC values among different elements is significantly reduced. This behavior is reflected in the convergence of the curves, where materials such as Fe, Ti, and Cu display nearly overlapping MAC values, while W and Pb retain slightly higher coefficients, though with much smaller differences than those observed at lower energies.

Within the high-energy range (approximately 1–10 MeV), the MAC values reach a minimum and then show a modest increasing trend. Although CS remains relevant, pair production begins to contribute increasingly, particularly for elements with high atomic numbers. Given that the pair production cross-section scales with Z^2, a gradual rise in MAC values is observed for heavy elements such as lead and tungsten, whereas this effect remains minimal for lower-Z materials like iron and titanium.

At very high photon energies (well above 10 MeV), PP becomes the dominant interaction process. As a result, the dependence of MAC on atomic number becomes pronounced once again. The graph indicates that lead exhibits the highest MAC values in this region, while lighter transition metals display considerably lower absorption coefficients, underscoring the superior shielding capability of high-Z materials against high-energy gamma radiation.

In conclusion, the graph clearly demonstrates the sequential transition of dominant photon–matter interaction mechanisms—from PE to CS and ultimately to PP—as photon energy increases. Moreover, it highlights that the influence of atomic number is substantial at low and very high energies, while it is comparatively limited in the intermediate energy range. These findings emphasize the importance of selecting appropriate elemental constituents or fillers in the design of shielding and composite materials tailored to specific photon energy regimes.

4. CONCLUSION

In this study, the photon–matter interaction mechanisms of selected transition elements (Fe, Ti, Cu, Mo, W, and Pb) were systematically examined over a broad photon energy range using partial and total mass attenuation coefficients. The results demonstrated that the dominant interaction processes vary markedly with photon energy and that these variations are strongly dependent on the atomic number of the elements.

At low photon energies, PE was identified as the prevailing interaction mechanism. Elements with high atomic numbers, particularly W and Pb, exhibited significantly enhanced shielding effectiveness due to the strong dependence on atomic number and the presence of absorption edges. In the intermediate energy region, incoherent (Compton) scattering became the dominant process, leading to reduced differences in attenuation performance among the elements and a diminished influence of atomic number. At high photon energies, the onset of the PP mechanism resulted in an increased contribution from high-Z elements, causing a noticeable rise in the total attenuation coefficients.

Comparative analyses indicated that Pb and W provide superior shielding performance in both low- and high-energy photon regions. In contrast, transition elements with intermediate atomic numbers, such as Cu and Mo, showed more uniform and reliable attenuation behavior across a wide energy range. Elements with relatively low atomic numbers were found to be less effective, particularly in the photoelectric absorption and pair production regions.

In conclusion, the findings indicate that no single element can offer optimal radiation shielding over the entire photon energy spectrum. The study highlights the importance of selecting shielding materials according to the targeted energy range and demonstrates that composite or multilayer shielding structures incorporating elements with different atomic numbers can provide more efficient attenuation across a broad range of photon energies. The results are expected to serve as a scientific foundation for the development of next-generation radiation shielding materials based on transition elements.

REFERENCES

Abu Al Roos, N.J., Baharul Amin, N.A., Zainon, R., (2019). Conventional and new lead-free radiation shielding materials for radiation protection in nuclear medicine: A review, *Radiat. Phys. Chem.* 165, 108439,

Abu Bakar, N.F., Amira Othman, S., Amirah Nor Azman, N.F., Saqinah Jasrin, N., (2019). Effect of ionizing radiation towards human health: a review. *IOP Conf. Series: Earth Environ. Sci.* 268 (1) 012005,

Berger, M.J., Hubbell, J.H., NBSIR 87-3597, (1987). Photon cross sections on a personal computer. *National Institute of Standards*, Gaithersburg, MD, USA.

Dur´an, A., Hian, S.K., Miller, D.L., Le Heron, J., Padovani, R., Vano, E., (2013). Recommendations for occupational radiation protection in interventional cardiology, Catheter. *Cardiovasc. Interv.* 82 (1), 29–42,

Durante, M., Cucinotta, F.A., (2011). Physical basis of radiation protection in space travel, *Rev. Mod. Phys.* 83 (4), 1245–1281,

Fern`andez, J.E., and Scot, V., (2007). Deterministic and Monte Carlo codes for multiple scattering photon transport with arbitrary polarization states. *Spectrochimica Acta Part B: Atomic Spectroscopy,* 62(6), pp.517-528.

Gunoglu, K., Akkurt, I., Sayyed, M. I., (2024). Radiation shielding properties of some igneous rocks in isparta province at different gamma energies: Experimental and theoretical study. *Journal of Radiation Research and Applied Sciences* 17 (2024) 100796, 1-9

Gunoglu, K., Varol Ozkavak, H., Akkurt, I., (2021). Evaluation of gamma ray attenuation properties of boron carbide (B4C) doped AISI 316 stainless steel: Experimental, XCOM and Phy-X/PSD database software. *Materials Today Communications* 29, 102793, 1-9.

Içelli, O., Erzeneoglu, S. (2004). Effective atomic numbers of some vanadium and nickel compounds for total photon interactions using transmission experiments. *Journal of Quantitative Spectroscopy and Radiative Transfer* 85: 115-24.

Khan, F.M., (2003). The physics of radiation therapy, Third edition. *Lippincott Williams and Wilkins*, Philadelphia.

Lakhwani, O.P., Dalal, V., Jindal, M., Nagala, A., (2019). Radiation protection and standardization, *J. Clin. Orthopaedics Trauma* 10 (4) 738–743

Mayles, P., Nahum, A., Rosenwald, J. and Papanikolaou, N., (2008). Handbook of Radiotherapy Physics: Theory and Practice. *Medical Physics,* 35(9), pp.4281-4281.

Patra, S.R., Sinha, S.P., Mishra, S.C. and Mallick, B., (2014). X-Ray Fibre Diffraction Analysis of 6 MV Photon-Induced Mimosa pudica. *Advanced Science Letters*, 20(3-4), pp.733-736.

Podgorsak, E.B., (2005). Radiation oncology physics. A handbook for teachers and students/EB *Podgorsak.–Vienna: International Atomic Energy Agency*, 657.

Singh, V.P., Badiger, N.M., (2016). Photon interaction with semiconductor and scintillation detectors. *Nuclear Science and Techniques* 27: 72. (12 pp.)

Zou, Y., Wei, Y., G. Wang, Meng, F., Gao, M., Storm, G., Zhong, Z., (2017). Nanopolymersomes with an ultrahigh iodine content for high-performance X-ray computed tomography imaging in vivo. *Adv. Mater.* 29 (10) 1603997

Chapter 7

MOLECULAR DOCKING ANALYSIS OF ZAVEGEPANT INTERACTION WITH THE AMYLIN-1 (AMY1) RECEPTOR

Beyza DAĞDELENER[1], Sertan AYTAÇ[2],

Özlem GÜNDOĞDU AYTAÇ[3]

1 Associate Degree Student, Kırşehir Ahi Evran University, Kaman Vocational School, Food Processing Department, Kaman-Kırşehir, Türkiye. ORCID ID: 0009-0007-7952-7815
2 Assist. Prof. Dr., Kırşehir Ahi Evran University, Kaman Vocational School, Food Processing Department, Kaman-Kırşehir, Türkiye. ORCID ID: 0000-0002-3196-4545
3 Assoc. Prof. Dr., Kırşehir Ahi Evran University, Kaman Vocational School, Food Processing Department, Kaman-Kırşehir, Türkiye. ORCID ID: 0000-0002-6943-9674

INTRODUCTION

Pain is defined by the International Association for the Study of Pain (IASP) as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage" (Öngel, 2016; Mukui et al., 2025). When not adequately controlled, pain may lead to significant physiological and psychological consequences, thereby markedly reducing individuals' quality of life (Soyuer & Kepenek-Varol, 2019). At the same time, pain is considered a component of the body's protective mechanisms and represents an important clinical sign indicating the presence of an underlying problem within the organism (Duman & Melek, 2010). Pain can be classified as acute or chronic based on its onset and duration, and as nociceptive or neuropathic according to its underlying mechanism. In terms of its anatomical origin, pain is further categorized as somatic, visceral, sympathetic, or peripheral (Aydın, 2002; Yağcı & Saygın, 2018).

Among neurologically mediated pain conditions, migraine (Chen et al., 2020) is currently regarded as a major public health problem (Steiner & Stovner, 2023). Migraine is a chronic disorder characterized by recurrent attacks of moderate to severe headache, often accompanied by nausea, vomiting, and increased sensitivity to light (photophobia) and sound (phonophobia) (Bernstein et al., 2013). Worldwide, migraine is one of the most common causes of disability among individuals under the age of 50, affecting more than one billion people, with a higher prevalence observed in women than in men (Gawde et al., 2023; Lechowicz et al., 2025).

Due to its complex neurobiology involving both the central and peripheral nervous systems, migraine is described as a highly disabling and prevalent neurological disorder (Lucas, 2021; Puledda et al., 2023). In this respect, migraine is not merely a type of headache (Yavuz et al., 2013), but rather a significant health problem that substantially restricts daily activities and negatively affects quality of life (Akyüz, 2021).

In migraine pathophysiology, the classical vascular theory that dominated for many years has gradually been replaced by neural theories (Christensen et al., 2025). It is well established that neurotransmitters released as a result of activation in cortical, subcortical, and brainstem regions play an important role in the initiation of migraine attacks. However, the precise triggers of this neurological activation and the exact mechanisms responsible for the onset of migraine attacks have not yet been fully elucidated (Hasırcı Bayır, 2021).

Various triggering factors have been reported to contribute to the development of migraine. These include stress (Berktaş et al., 2019), disturbances in sleep patterns (Vgontzas & Pavlović, 2018), hormonal fluctuations—particularly during menstruation, menopause, and pregnancy in women (Kim & Park, 2024)—as well as environmental factors such as bright light, loud noise, and strong odors (Bol, 2023). Certain foods, including chocolate, milk and dairy products, nuts, citrus fruits, and alcohol, have also been identified as potential triggers (Ekinoğlu & Semerci, 2022). In addition, irregular eating habits, sudden changes in weather conditions, and physical fatigue are among other factors that may precipitate migraine attacks (Yaman et al., 2007).

Migraine is classified into three main subtypes: migraine without aura, migraine with aura, and chronic migraine (Grodzka et al., 2025). Migraine without aura is characterized by unilateral, pulsating, and severe headaches lasting between 4 and 72 hours, frequently accompanied by nausea and sensitivity to light and sound. These headaches tend to worsen with physical activity (Yücel, 2008). Migraine with aura typically begins with attacks lasting several minutes and is associated with transient neurological symptoms affecting visual, auditory, or motor functions, followed by the characteristic migraine headache and related symptoms (Lucas, 2021). Chronic migraine is defined by the presence of headache on at least 15 days per month for a duration of three months or longer (Lipton et al., 2023).

Migraine attacks are generally described as occurring in four distinct phases. The first phase, known as the prodromal phase, may begin hours or even days before the onset of headache and is characterized by symptoms such as mood changes, impaired concentration, fatigue, alterations in appetite, gastrointestinal disturbances, and increased urinary frequency (Ekizoğlu, 2012). The second phase, the aura phase, is marked by transient and focal neurological symptoms. Visual disturbances are the most common manifestations, although temporary sensory, speech/language, and motor impairments may also occur during this phase (Dodick, 2018). The third phase, the headache phase, typically involves unilateral, pulsating, and severe pain, often accompanied by nausea, vomiting, fatigue, irritability, and depressive mood (Ekizoğlu, 2012; Ünal-Cevik & Arslan, 2023). The final phase, known as the postdromal phase, follows the resolution of headache and is characterized by fatigue, slowed cognitive function, difficulty concentrating, and a general sense of weakness. Nevertheless, some individuals report feeling unexpectedly refreshed and restored during this period (Rizzoli et al., 2018).

Conventional medications used in migraine treatment often show limited efficacy and may cause various adverse effects, thereby increasing the need for more effective and safer therapeutic approaches (Bentivegna et al., 2024). Migraine management is generally divided into non-pharmacological and pharmacological strategies (Bol, 2023). Non-pharmacological approaches include regular sleep habits, balanced nutrition, mild to moderate physical exercise, and stress management, all of which may help reduce the frequency and severity of migraine attacks. Avoidance of caffeine, processed foods, and sleep deprivation also contributes to non-pharmacological migraine control (Tarlacı, 2006). Pharmacological treatment constitutes the second main pillar of migraine management and is subdivided into acute attack treatment and preventive (prophylactic) therapy (Aygün et al., 2018). Acute treatment aims to rapidly alleviate symptoms during a migraine attack (Ong et al., 2018), whereas preventive therapy is defined as a regular and long-term approach intended to reduce the frequency, duration, and severity of attacks (Tarlacı, 2006; Silberstein, 2011) (Figure 1).

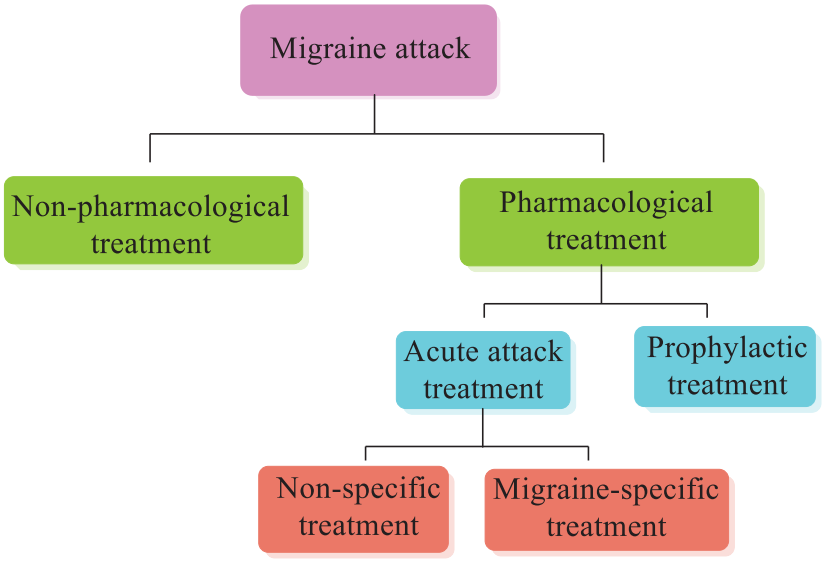


Figure 1. Therapeutic approaches in migraine management (Tarlacı, 2006).

Acute treatment of migraine involves the use of non-specific and specific therapeutic approaches, depending on the severity of the disease and individual patient characteristics (Öztürk, 2013). In non-specific migraine treatment, particularly for the management of mild to moderate acute migraine attacks, non-steroidal anti-inflammatory drugs (NSAIDs) such as paracetamol, aspirin, ibuprofen, naproxen, and diclofenac are widely used. Although these agents are effective in some patients, they

may cause various adverse effects, especially involving the gastrointestinal system (Láinez et al., 2013; Roceanu et al., 2015).

In addition, antiemetic medications (e.g., metoclopramide and domperidone) may be administered concomitantly with NSAIDs to alleviate vegetative symptoms frequently associated with migraine attacks, such as nausea and vomiting, and to enhance the gastrointestinal absorption of analgesics (Láinez et al., 2013) (Figure 2).

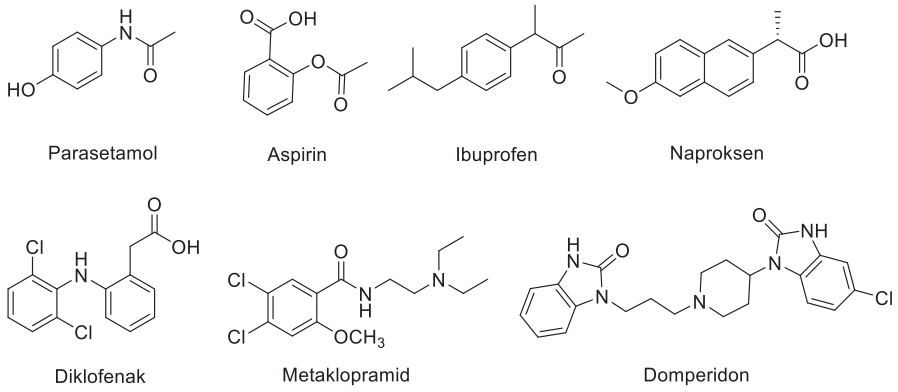


Figure 2. Pharmacological agents used in non-specific migraine management.

Triptans used in the specific treatment of migraine (sumatriptan, almotriptan, zolmitriptan, eletriptan, naratriptan, rizatriptan, and frovatriptan) are 5-HT1B and 5-HT1D receptor agonists and are widely used as migraine-specific abortive medications (Durgaprasad et al., 2022). Lasmiditan, on the other hand, is a more recently developed agent for the acute treatment of migraine and is classified as a ditan, acting as a selective 5-HT1F receptor agonist. Both triptans and lasmiditan are considered first-line treatment options, particularly in patients experiencing moderate to severe migraine attacks (Shibata et al., 2024).

In contrast, ergotamine derivatives (Tfelt-Hansen et al., 2000) are currently less preferred in clinical practice due to their low oral bioavailability, relatively limited efficacy compared with triptans, and the potential to cause prominent adverse effects such as nausea (Bol, 2023) (Figure 3).

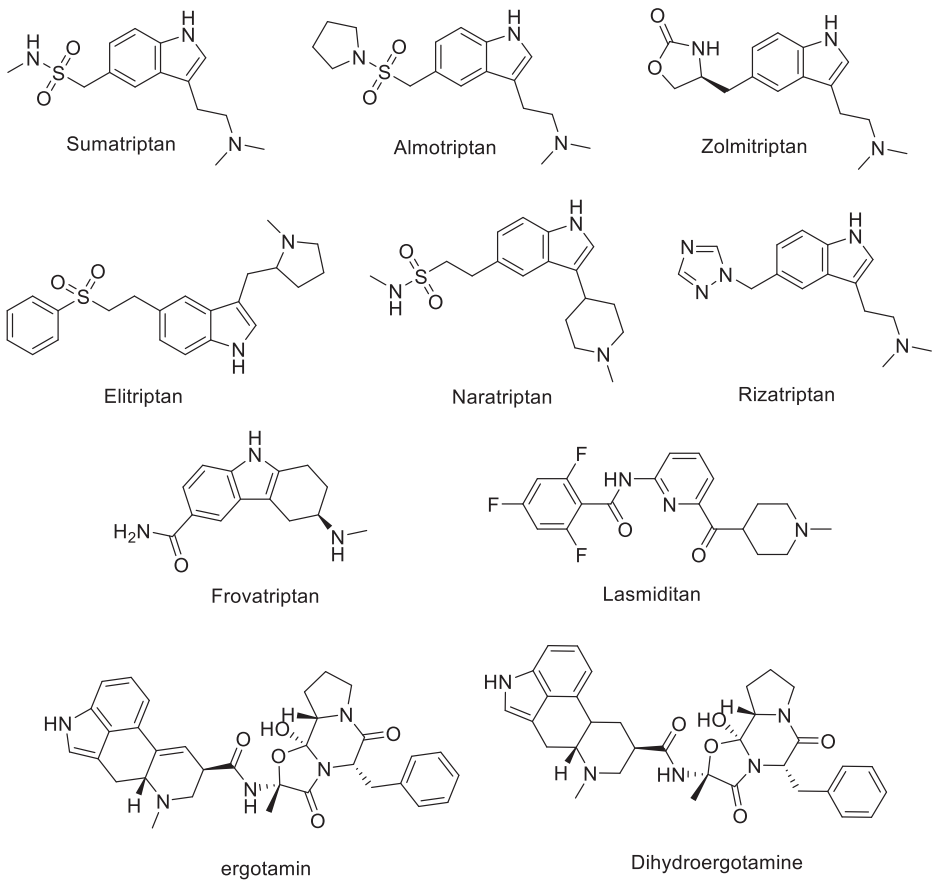


Figure 3. Commonly used medications in the specific treatment of migraine.

Individuals who experience frequent migraine attacks may also receive preventive (prophylactic) treatments aimed at reducing the frequency, intensity, and duration of attacks. Medications used in migraine prophylaxis act through various receptors or ion channels. For instance, commonly used preventive agents such as propranolol (Szeleszczuk & Fraczkowski, 2022), topiramate (Edvinsson & Linde, 2010), amitriptyline (Khalid & Ahmad, 2022), and flunarizine have been shown to suppress cortical spreading depression (CSD) (Sprenger et al., 2018) (Figure 4).

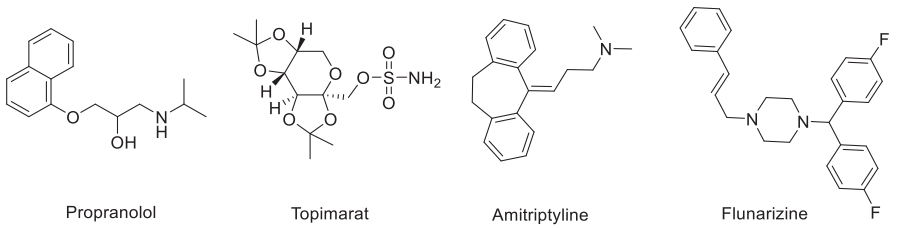


Figure 4. Medications commonly employed in the prophylactic treatment of migraine.

Following the demonstration of the important role of calcitonin gene-related peptide (CGRP) in the development of migraine, research efforts aimed at developing agents targeting CGRP have accelerated (Moriarty & Barch, 2025). Calcitonin gene-related peptide (CGRP) is a neuropeptide consisting of 37 amino acids and is widely distributed in both the central and peripheral sensory nervous systems (Poyner et al., 2009). CGRP receptors are transmembrane G protein–coupled receptors formed through the interaction between the calcitonin receptor-like receptor (CLR) and receptor activity-modifying protein 1 (RAMP1) (dos Santos et al., 2022).

Novel migraine therapies targeting CGRP are broadly classified into two main groups: CGRP monoclonal antibodies (mAbs) and CGRP receptor antagonists, known as gepants. These agents exert

their effects by targeting the CGRP signaling pathway, which plays a critical role in migraine pathogenesis (Berger et al., 2022). Therapeutic strategies that block either the CGRP ligand or its receptor are considered a major advancement in the prevention and management of migraine (Orlando et al., 2024).

CGRP monoclonal antibodies used in migraine treatment offer significant advantages over traditional preventive therapies because they directly target the underlying pathophysiology of the disease (Choi et al., 2025). In patients with episodic and chronic migraine, these biological agents represent an effective option when conventional treatments fail. Despite their large molecular size and limited ability to cross the blood–brain barrier, they can exert clinical effects due to the presence of CGRP and its receptors outside the barrier (Hong et al., 2023; Karlı et al., 2022). Clinical evidence indicates that these treatments reduce attack frequency, decrease the need for acute medication, alleviate functional impairment, and improve quality of life (Karlı et al., 2022; Arzalluz-Luque et al., 2024). Anti-CGRP monoclonal antibodies are divided into two groups: those targeting the CGRP receptor (e.g., erenumab) and those targeting the CGRP peptide itself (e.g., fremanezumab, galcanezumab, and eptinezumab). These agents differ in antibody structure, routes of administration (subcutaneous or intravenous), and pharmacokinetic properties (Dolinkiewicz et al., 2025).

Gepants are a newly developed class of small-molecule drugs for migraine treatment that are administered orally (Rissardo & Caprara, 2022). By blocking CGRP receptors, these agents primarily exert their effects in regions located outside the blood–brain barrier, such as the trigeminal ganglion and the meninges. In this way, they inhibit CGRP-induced vasodilation and neurogenic inflammation, thereby contributing to a reduction in the occurrence and severity of migraine attacks (Anukoolwittaya et al., 2025). Randomized controlled clinical trials have demonstrated that gepants are effective in reducing both the severity and duration of migraine attacks (Messina et al., 2019).

Among the first gepants developed, olcegepant (Durham, 2024) and telcagepant (MacGregor, 2011) had limited clinical use due to pharmacokinetic constraints and the risk of hepatotoxicity (Yao et al., 2013). In contrast, ubrogepant (Yin et al., 2020), rimegepant (Ocheretyaner et al., 2022), and zavegepant (Dhillon, 2023) have demonstrated efficacy in the acute treatment of migraine, while atogepant (Baraldi et al., 2024) and rimegepant have shown successful clinical outcomes in migraine prevention (Labastida-Ramírez et al., 2023) (Figure 5).

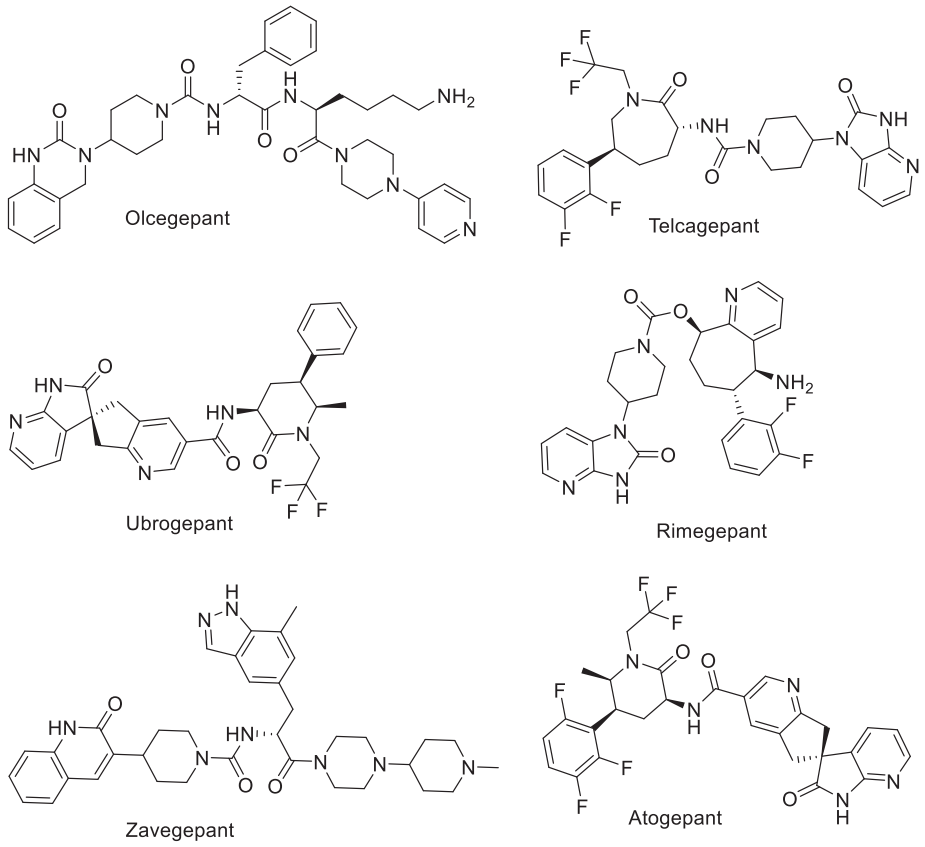


Figure 5. Some gepants developed for the acute and prophylactic treatment of migraine.

The calcitonin gene-related peptide (CGRP) system plays a critical role in migraine pathogenesis, and therapies targeting this system have led to significant advances in the pharmacological management of migraine (Garelja et al., 2022). The peptide hormone amylin is a member of the calcitonin (CT)/CGRP peptide family and shares the amylin-1 (AMY1) receptor, which can be activated by both CGRP and amylin (Gostynska et al., 2024). The AMY1 receptor is expressed in the trigeminovascular system and is considered a potential target for migraine treatment (Elbendary et al., 2025).

Small-molecule CGRP receptor antagonists (gepants) can affect both CGRP and AMY1 receptors (Bolay, 2018). CGRP activates two distinct receptors in the trigeminovascular system: the CGRP receptor and the AMY1 receptor (Rees et al., 2022). The function of these receptors depends on the co-expression of receptor activity-modifying protein 1 (RAMP1) with different G protein-coupled receptors (Russo & Hay, 2023). The CGRP receptor is formed by RAMP1 and the calcitonin receptor-like receptor (CLR), whereas the AMY1 receptor is formed by RAMP1 and the calcitonin receptor (CTR) (Walker et al., 2018). The identification of both receptors in the trigeminovascular system highlights the central role of CGRP-mediated signaling in migraine development (Eftekhari & Edvinsson, 2011).

Clinical studies have demonstrated the efficacy of small-molecule antagonists such as olcegepant (BIBN4096BS) and telcagepant (MK-0974), which bind to the CGRP receptor and block its activation; however, the development of these agents was discontinued for various reasons. Although these drugs are primarily considered specific to the CGRP receptor, they are reported to potentially have antagonistic effects on the AMY1 receptor. Next-generation gepants, including ubrogepant (Haghdoost et al., 2023) and zavegepant (Song et al., 2025), used in the acute treatment of migraine, exert their effects by antagonizing CGRP receptors (Chiang & VanderPluym, 2021). Nevertheless, data on the potential effects of these drugs on the AMY1 receptor remain limited.

AIM OF THE STUDY

The process of transforming pharmaceutical compounds into effective drugs is highly time-consuming and costly, as it requires extensive in vitro and in vivo testing (Özgen & Ünlü, 2022). In this process, a detailed examination of the structure-activity relationships of compounds contributes to the design of more effective and selective drug candidates, while understanding pharmacological properties enables the identification of new therapeutic applications (Tarı & Arpacı, 2024).

In this context, molecular docking studies play a crucial role in predicting the therapeutic potential of compounds by analyzing how they interact with target proteins (Chaudhary & Tyagi, 2024). This approach, which reveals the binding tendencies and interaction mechanisms of ligands with target proteins, offers significant advantages for enhancing the efficacy of potential drug candidates and reducing possible side effects (Gündoğdu, 2023; Şenol, 2023; Muhammed & Aki-Yalçın, 2024). Molecular docking, widely used in early-stage drug discovery, provides rapid, reliable, and detailed information on protein-ligand interactions and potential binding sites, thereby supporting the development of targeted therapeutic strategies (Mahmudov et al., 2022).

As a third-generation CGRP receptor antagonist developed in a nasal spray formulation, Zavegepant has been shown in Phase 2/3 clinical trials to be significantly superior to placebo in both pain freedom and relief of the most bothersome symptom (Croop et al., 2022). In addition, the results of meta-analyses on clinical applications indicate that Zavegepant provides pain relief within 2 hours of administration and reduces migraine-associated bothersome symptoms (Zhu et al., 2025).

Although structural and pharmacological studies have elucidated the interactions of the AMY1 receptor with CGRP and other ligands, no peer-reviewed study in the literature has investigated the molecular docking of Zavegepant with the AMY1 receptor. Based on this gap, this study focuses on evaluating the interaction potential of the AMY1 receptor with gepants and aims to better understand its role in migraine treatment. For this purpose, Zavegepant, a third-generation gepant, was selected as the compound of interest and forms the foundation of the research conducted in this study.

MATERIAL AND METHODS

Molecular Docking Studies

Molecular docking is a widely used computational method in drug discovery and pharmacology that enables the prediction of protein–ligand interactions and the identification of potential binding sites with high accuracy and efficiency (Gündoğdu, 2023; Mahmudov et al., 2022). In this study, the Amylin1 Receptor in complex with Gs and rat amylin peptide (rAmy-AMY1R:Gs) (PDB ID: 7TYF, G) was selected as the target protein for in silico docking experiments (Cao et al., 2024). Zavegepant was evaluated as the test compound, while Ibuprofen was used as the reference drug for comparative purposes.

The 3D structures of Zavegepant and Ibuprofen were generated by first drawing their 2D chemical structures, followed by geometry optimization using the Avogadro software. Protein structures were obtained from the Protein Data Bank (https://www.rcsb.org/). Ligand and protein preparations were carried out using MGLTools, where water molecules were removed, polar hydrogens were added, and Kollman charges were assigned. The torsional flexibility of the ligands was defined, and all molecules were saved in PDBQT format for docking.

To ensure accurate binding predictions, the active sites of the proteins were defined, and grid box parameters were carefully set. For rAmy-AMY1R:Gs (PDB ID: 7TYF), the grid dimensions were 60 × 60 × 60 Å^3, with a spacing of 0.375 Å, and center coordinates at x = 117.938, y = 104.417, and z = 107.316. Molecular docking simulations were carried out using AutoDock version 1.5.7 employing

the Lamarckian Genetic Algorithm (LGA). This approach enabled a detailed assessment of the principal non-covalent interactions, such as hydrogen bonding, van der Waals interactions, and hydrophobic contacts, between the ligands and the corresponding binding pockets.

The predicted protein–ligand complexes were subsequently visualized and analyzed with BIOVIA Discovery Studio in order to evaluate binding conformations and interaction patterns. Docking scores and interaction diagrams are presented in Figures 6–7, providing insights into the potential activity of Zavegepant compared to the reference drug (Ibuprofen).

RESULTS AND DISCUSSION

In this study, a comparative molecular docking analysis was performed to evaluate the analgesic properties of Zavegepant and the reference drug Ibuprofen on migraine. The target protein selected was rAmy-AMY1R:Gs (PDB ID: 7TYF). Examination of Zavegepant's binding interactions within the active site of the target protein (7TYF) revealed that the compound exhibited a total binding energy of −7.97 kcal/mol, indicating a strong binding affinity of Zavegepant for the active site. The stability of the ligand–protein complex was primarily supported by a combination of hydrogen bonds and hydrophobic interactions.

Evaluation of Zavegepant's binding interactions showed that the ligand formed a π–π stacking interaction with the Tyr263 residue at a distance of 5.44 Å, and multiple π–π stacking interactions with the Trp267 residue at distances of 4.24, 4.68, 3.61, and 3.85 Å. Additionally, various alkyl interactions were observed between the aliphatic regions of the ligand and hydrophobic amino acid residues in the active site. For instance, alkyl interactions were detected with Leu264 at 5.06 and 4.47 Å, Met229 at 5.23 Å, and Leu193 at 5.06 Å. Moreover, π–alkyl interactions were also observed between the ligand and certain amino acid residues: Tyr263 at 4.90 Å, Trp267 at 4.93 and 4.94 Å, Leu189 at 5.20 and 5.13 Å, and Lys185 at 5.05 Å.

Van der Waals interactions were additionally identified between the ligand and Arg268, Arg260, Tyr228, Thr182, and Leu167 residues. All molecular interactions between Zavegepant and the target protein are detailed in Figure 6.

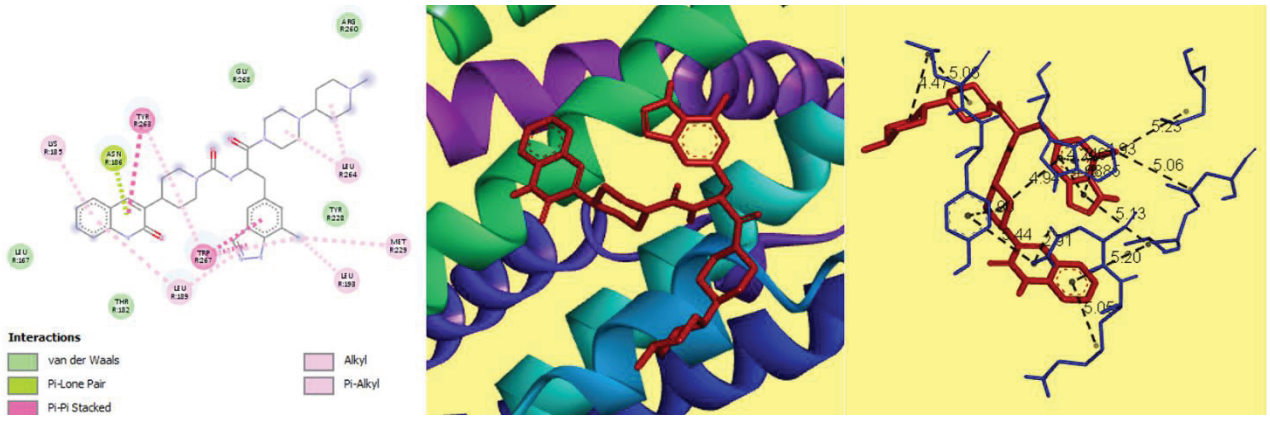


Figure 6. Docking interactions between the protein 7TYF and the Zavegepant

Evaluation of Ibuprofen's binding mode within the active site of the target protein (PDB ID: 7TYF) revealed that the compound exhibited a total binding energy of −4.95 kcal/mol, indicating a weak binding affinity. Analysis of the molecule–protein interactions showed that the carbonyl oxygen of Ibuprofen formed a conventional hydrogen bond with the Thr182 residue at a distance of 2.09 Å.

Furthermore, hydrophobic interactions were observed to be dominant. The alkyl groups of Ibuprofen established alkyl–alkyl interactions with Leu189 (4.17 Å) and Lys185 (5.11 Å) residues. In addition,

a π–alkyl interaction was identified between Ibuprofen and Lys185 at a distance of 4.86 Å. Van der Waals interactions were also detected with Asn186, Leu167, and Tyr263 residues.

Overall, the limited number of hydrogen bonds combined with predominantly weak hydrophobic interactions supports the conclusion that Ibuprofen binds to the 7TYF active site with low binding strength. All molecular interactions between Ibuprofen and the target protein are detailed in Figure 7.

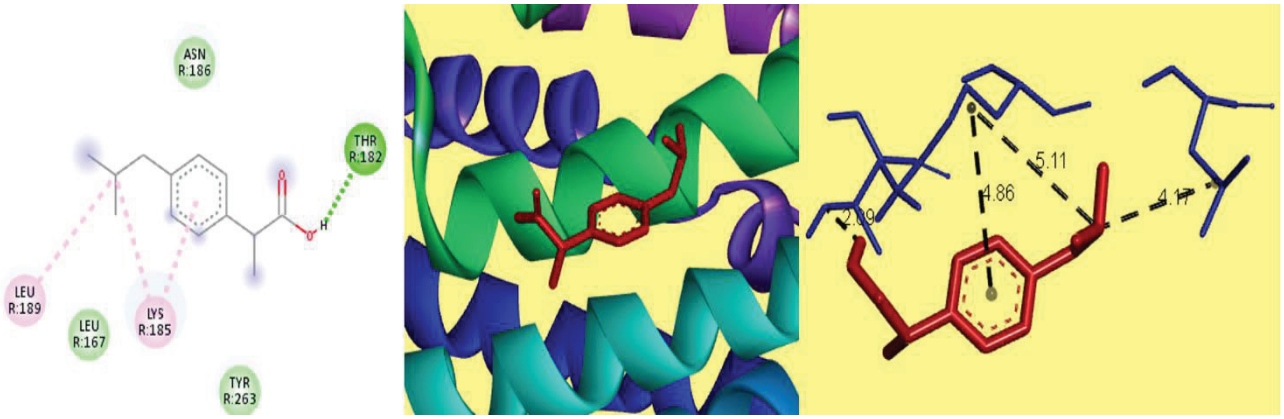


Figure 7. Docking interactions between the 7TYF protein and Ibuprofen

CONCLUSION

In this study, molecular docking analyses were conducted to comparatively evaluate the interaction profiles of Zavegepant and Ibuprofen with the 7TYF target protein. The results indicated that Zavegepant binds to the active site more strongly, stably, and specifically than Ibuprofen. The calculated binding free energies were −7.97 kcal/mol for Zavegepant and −4.95 kcal/mol for Ibuprofen, demonstrating that Zavegepant has a higher binding affinity for the target protein. This difference in binding energy suggests that Zavegepant can form more favorable intermolecular interactions, resulting in a more stable ligand–protein complex. In contrast, Ibuprofen formed more limited and transient interactions.

These findings confirm at the molecular level that Zavegepant can establish strong and stable interactions with the 7TYF target protein. Clinical applications have also shown that the drug is effective in achieving pain freedom and reducing migraine-associated bothersome symptoms; this indicates that the theoretical data are consistent with clinical outcomes and supports Zavegepant's therapeutic potential.

REFERENCES

Akyüz, H. Ö. (2021). Migrende görülen belirtilerin yaşam kalitesi ve ağrı şiddeti üzerine etkisinin incelenmesi. *Van Sağlık Bilimleri Dergisi, 14*(2), 209–215. https://doi.org/10.52976/vansaglik.859024

Anukoolwittaya, P., Rattanawong, W., Vongvaivanich, K., Pongpitakmetha, T., Thanprasertsuk, S., Poonpedpun, T., Soontrapa, P., Suwanlaong, K., Kongbunkiat, K., Komonchan, S., Dusitanond, P., Teekaput, C., Sirimaharaj, N., Yuvasilp, N., & Tanprawate, S. (2025). Expert consensus on gepants for acute and preventive treatment of migraine in Thailand. *The Journal of Headache and Pain, 26*(131). https://doi.org/10.1186/s10194-025-02074-4

Arzalluz-Luque, J., Millán Vázquez, M., Lamas Pérez, R., Sánchez Rodríguez, N., Gómez López, P., Gómez Fernández, F. J., Viguera Romero, J., Jurado Cobo, C., Fernández Recio, M., & González Oria, C. (2024). Switching of monoclonal antibodies against the calcitonin gene-related peptide or its receptor in migraine: Results from a Spanish cohort. *Neurology Perspectives, 4*, 100168. https://doi.org/10.1016/j.neurop.2024.100168

Aydın, O. N. (2002). Ağrı ve ağrı mekanizmalarına güncel bakış. *ADÜ Tıp Fakültesi Dergisi*, *3*(2), 37-48.

Aygün, R. B., Genç, F., & Şahin, E. (2018). Migren tedavisinde eczacının rolü. *Lectio Scientific Journal of Health and Natural Sciences*, *2*(1), 8–20.

Baraldi, C., Beier, D., Martelletti, P., & Pellesi, L. (2024). The preclinical discovery and development of atogepant for migraine prophylaxis. *Expert Opinion on Drug Discovery, 19*(7), 783–788. https://doi.org/10.1080/17460441.2024.2365379

Bentivegna, E., Galastri, S., Onan, D., & Martelletti, P. (2024). Unmet needs in the acute treatment of migraine. *Advances in Therapy, 41*(1), 1–13. https://doi.org/10.1007/s12325-023-02650-7

Berger, A. A., Winnick, A., Carroll, A. H., Welschmeyer, A., Li, N., Colon, M., Paladini, A., Ramírez, G. F., Hasoon, J., Cornett, E. M., Song, J., Varrassi, G., Kaye, A. M., Kaye, A. D., & Ganti, L. (2022). Rimegepant for the treatment of migraine. *Health Psychology Research, 10*(5). https://doi.org/10.52965/001c.38534

Berktaş, F., Ekemen, E., Kıroğlu, O., & Aksu, F. (2019). Migren tedavisinde antidepresan ilaçların rolü. *Çukurova Medical Journal, 44*(Suppl 1), 555–566. https://doi.org/10.17826/cumj.570837

Bernstein, J. A., Fox, R. W., Martin, V. T., & Lockey, R. F. (2013). Headache and facial pain: Differential diagnosis and treatment. *Journal of Allergy and Clinical Immunology: In Practice*, *1*(3), 242-251. https://doi.org/10.1016/j.jaip.2013.03.014

Bol, C. (2023). *Migren tipi baş ağrısı olan bireylerin duygusal özyeterliklerinin anksiyete duyarlılığı, yeti yitimi ve ağrı düzeyi ile ilişkisi* (Yayımlanmamış yüksek lisans tezi). Ordu Üniversitesi Sağlık Bilimleri Enstitüsü, Ordu.

Bolay, H. (2018). Migren tedavisinde ufukta görünen yeni seçenekler. In Ş. Bıçakcı, M. Öztürk, S. Üçler, N. Karli, & A. Siva (Eds.), *Baş ağrısı: Tanı ve tedavi, güncel yaklaşımlar* (ss. 79–83). İstanbul: Galenos Yayınevi. ISBN: 978-605-89294-7-0

Chaudhary, M., & Tyagi, K. (2024). A review on molecular docking and its application. *International Journal of Advanced Research, 12*(3), 1141–1153. https://doi.org/10.21474/IJAR01/18505

Chen, D., Willis-Parker, M., & Lundberg, G. P. (2020). Migraine headache: Is it only a neurological disorder? Links between migraine and cardiovascular disorders. *Trends in Cardiovascular Medicine, 30*(7), 424-430. https://doi.org/10.1016/j.tcm.2019.10.005

Chiang, C. C., & VanderPluym, J. H. (2021). Ubrogepant in the acute management of migraine: A narrative review. *Journal of Pain Research, 14*, 1185–1192. https://doi.org/10.2147/JPR.S244249

Choi, S., Hong, Y., Kang, M.-K., Song, T.-J., & Cho, S.-J. (2025). Calcitonin gene-related peptide monoclonal antibody treatment in nine cases of persistent headache following COVID-19 infection. *Journal of Korean Medical Science, 40*(25), e127. https://doi.org/10.3346/jkms.2025.40.e127

Christensen, R. H., Ashina, H., & Ashina, M. (2025). The vessel-to-neuron trigeminovascular hypothesis of migraine pathogenesis – the 'pro' argument. *The Journal of Headache and Pain, 26*, 248. https://doi.org/10.1186/s10194-025-02130-z

Croop, R., Madonia, J., Stock, D. A., Thiry, A., Forshaw, M., Murphy, A., Coric, V., & Lipton, R. B. (2022). Zavegepant nasal spray for the acute treatment of migraine: A phase 2/3 double-blind, randomized, placebo-controlled, dose-ranging trial. *Headache, 62*(9), 1153–1163. https://doi.org/10.1111/head.14389

Dhillon, S. (2023). Zavegepant: First approval. *Drugs, 83*(8), 825–831. https://doi.org/10.1007/s40265-023-01885-6

Dodick, D. W. (2018). A phase-by-phase review of migraine pathophysiology: Supplement article. *Headache: The Journal of Head and Face Pain, 58*(S1), 4–16. https://doi.org/10.1111/head.13300

Dolinkiewicz, J., Szczerbińska, A., Kamiński, J., Rurka, S., & Błoniecka, J. (2025). Monoclonal antibodies targeting CGRP in migraine: Current perspectives and clinical guidelines. *International Journal of Innovative Technologies in Social Science, 3*(47). https://doi.org/10.31435/ijitss.3(47).2025.3641

dos Santos, J. B. R., & da Silva, M. R. R. (2022). Small molecule CGRP receptor antagonists for the preventive treatment of migraine: A review. *European Journal of Pharmacology, 922*, 174902. https://doi.org/10.1016/j.ejphar.2022.174902

Duman, T., & Melek, İ. (2010). Ağrı mekanizmaları ve ağrı tedavisine farmakolojik yaklaşım. *Türkiye Klinikleri Journal of Neurology - Special Topics*, *3*(1), 51-64.

Durgaprasad, K., Gowri, S., & Mishra, K. (2022). Formulation and evaluation of triptans: A critical overview. *High Technology Letters, 28*(6), 181–192. https://doi.org/10.37896/HTL28.06/5914

Edvinsson, L., & Linde, M. (2010). New drugs in migraine treatment and prophylaxis: Telcagepant and topiramate. *The Lancet, 376*(9741), 645–655. https://doi.org/10.1016/S0140-6736(10)60323-6

Eftekhari, S., & Edvinsson, L. (2011). Calcitonin gene-related peptide (CGRP) and its receptor components in human and rat spinal trigeminal nucleus and spinal cord at C1-level. *BMC Neuroscience, 12*, 112. https://doi.org/10.1186/1471-2202-12-112.

Ekinoğlu, E., & Semerci, C. (2022). Migren ve güncel diyet yaklaşımları. *Uluslararası Hakemli Akademik Spor Sağlık ve Tıp Bilimleri Dergisi, 45*, 161–170. https://doi.org/10.17363/SSTB.2022/ABCD89/.45.11.161-170

Ekizoğlu, E. (2012). *Migren ve küme tipi başağrısı olan hastalarda allodini ve kortikal eksitabilite ilişkisinin araştırılması* (Uzmanlık tezi). İstanbul Üniversitesi İstanbul Tıp Fakültesi, Nöroloji Anabilim Dalı, İstanbul.

Elbendary, L. A., Abdel-Latif, G. A., Khattab, M. A., El-Sahar, A. E., & Sayed, R. H. (2025). The role of the AMY1 receptor signaling cascade in the protective effect of sulforaphane against nitroglycerin-induced migraine in mice. *Archives of Pharmacal Research, 358*(9), e70107. https://doi.org/10.1002/ardp.70107

Garelja, M. L., Walker, C. S., & Hay, D. L. (2022). CGRP receptor antagonists for migraine: Are they also AMY1 receptor antagonists? *British Journal of Pharmacology, 179*(3), 454–459. https://doi.org/10.1111/bph.15585

Gawde, P., Shah, H., Patel, H., Bharathi, K. S., Patel, N., Sethi, Y., & Kaka, N. (2023). Revisiting migraine: The evolving pathophysiology and the expanding management armamentarium. *Cureus, 15*(2), e34553. https://doi.org/10.7759/cureus.34553

Gostynska, S. E., Karim, J. A., Ford, B. E., Gordon, P. H., Babin, K. M., Inoue, A., Lambert, N. A., & Pioszak, A. A. (2024). *Amylin receptor subunit interactions are modulated by agonists and determine signaling*. bioRxiv. https://doi.org/10.1101/2024.10.09.617487 *Update in: Sci. Signal., 18*(900), eadt8127. https://doi.org/10.1126/scisignal.adt8127

Grodzka, O., Dzagoevi, K., Rees, T., Cabral, G., Chądzyński, P., Di Antonio, S., Sochań, P., MaassenVanDenBrink, A., & Lampl, C.; European Headache Federation School of Advanced Studies (EHF-SAS). (2025). Migraine with and without aura—Two distinct entities? A narrative review. *The Journal of Headache and Pain, 26*(1), 77. https://doi.org/10.1186/s10194-025-01998-1

Gündoğdu, O. (2023). Molecular docking studies and ADME predictions on synthesized chalcone compounds targeting EGFR. *Hittite Journal of Science and Engineering,* 10(2), 167–175.

Haghdoost, F., Puledda, F., Garcia-Azorin, D., Huessler, E. M., Messina, R., & Pozo-Rosich, P. (2023). Evaluating the efficacy of CGRP mAbs and gepants for the preventive treatment of migraine: A systematic review and network meta-analysis of phase 3 randomised controlled trials. *Cephalalgia, 43*(4), 3331024231159366. https://doi.org/10.1177/03331024231159366

Hasırcı Bayır, B. R. (2021). *Menstrüel migren hastalarında profilakside transkraniyal doğru akım stimülasyonunun etkinliğinin ve CGRP düzeyleri ile bağlantısının değerlendirilmesi* (Yayımlanmamış yüksek lisans tezi). İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Sinirbilim Anabilim Dalı, Elektronörofizyoloji Programı, İstanbul.

Hong, J. B., Lange, K. S., Overeem, L. H., Triller, P., Raffaelli, B., & Reuter, U. (2023). A scoping review and meta-analysis of anti-CGRP monoclonal antibodies: Predicting response. *Pharmaceuticals, 16*(7), 934. https://doi.org/10.3390/ph16070934

Khalid, M., & Ahmad, S. (2022). Box–Behnken design used to optimize the simultaneous quantification of amitriptyline and propranolol in tablet dosages by RP-HPLC-DAD method and their stability tests. *Separations, 9*(12), 421. https://doi.org/10.3390/separations9120421

Kim, S., & Park, J. W. (2024). Migraines in women: A focus on reproductive events and hormonal milestones. *Headache Pain Research, 25*(1), 3–15. https://doi.org/10.62087/hpr.2024.000

Labastida-Ramírez, A., Caronna, E., Gollion, C., Stanyer, E., Dapkute, A., Braniste, D., Naghshineh, H., Meksa, L., Chkhitunidze, N., Gudadze, T., Pozo-Rosich, P., Burstein, R., & Hoffmann, J. (2023). Mode and site of action of therapies targeting CGRP signaling. *The Journal of Headache and Pain, 24*(1), 125. https://doi.org/10.1186/s10194-023-01644-8

Láinez, M. J., García-Casado, A., & Gascón, F. (2013). Optimal management of severe nausea and vomiting in migraine: Improving patient outcomes. *Patient Related Outcome Measures, 4*, 61–73. https://doi.org/10.2147/PROM.S31392

Lechowicz, E., Łaciński, A., Smulska, A., Grodzka, O., & Domitrz, I. (2025). A correlation between migraine and endometriosis and its clinical implications—A systematic literature review. *J. Clin. Med.*, *14*, 2744. https://doi.org/10.3390/jcm14082744

Lipton, R. B., Buse, D. C., Nahas, S. J., Tietjen, G. E., Martin, V. T., Löf, E., Brevig, T., Cady, R., & Diener, H.-C. (2023). Risk factors for migraine disease progression: A narrative review for a patient-centered approach. *Journal of Neurology, 270*(12), 5692–5710. https://doi.org/10.1007/s00415-023-11880-2

Lucas, C. (2021). Migraine with aura. *Revue Neurologique*, 177(7), 779-784. https://doi.org/10.1016/j.neurol.2021.07.010

MacGregor, E. A. (2011). Telcagepant: A new therapeutic option for acute migraine? *Clinical Medicine Insights: Therapeutics, 3*, 301–314. https://doi.org/10.4137/CMT.S7477

Mahmudov, I., Demir, Y., Sert, Y., Abdullayev, Y., Sujayev, A., Alwasel, S. H., & Gülçin, İ. (2022). Synthesis and inhibition profiles of *N*-benzyl- and *N*-allyl aniline derivatives against carbonic anhydrase and acetylcholinesterase – A molecular docking study. *Arabian Journal of Chemistry, 15*(3), 103645. https://doi.org/10.1016/j.arabjc.2021.103645

Messina, R., & Goadsby, P. J. (2019). CGRP—a target for acute therapy in migraine: Clinical data. *Cephalalgia, 39*(3), 420–427. https://doi.org/10.1177/0333102418768095.

Moriarty, M. A., & Barch, C. A. (2025). Gepants in primary care: A targeted approach to acute and preventive treatment of migraine. *Pain and Therapy, 14*, 1263–1278. https://doi.org/10.1007/s40122-025-00757-z

Muhammed, T. M., & Aki-Yalçın, E. (2024). Molecular docking: Principles, advances, and its applications in drug discovery. *Letters in Drug Design & Discovery, 21*(3), 480–495.

Mukui, A. M., Bishi, A. A., Ghazwani, B. H. A., Alnasyan, T. M., Alkhaibari, B. N., Sharahili, A. Y., Almassari Aldosari, N. A. A., Al-Anzi, S. K., Al-Anzi, M. A., Al-Anzi, F. A., & Alkhybari, N. A. (2025). An updated review about pain management medications and biochemical mechanisms—Overview for healthcare professionals. *Egyptian Journal of Chemistry, 68*(2), 135–144. https://doi.org/10.21608/ejchem.2025.340884.10925

Ocheretyaner, E. R., Kofman, M., & Quattrocchi, E. (2022). Calcitonin gene-related peptide (CGRP) receptor antagonists for the acute treatment of migraines in adults. *Drugs in Context, 11*, 2022-3-5. https://doi.org/10.7573/dic.2022-3-5

Ong, J. J. Y., & De Felice, M. (2018). Migraine treatment: Current acute medications and their potential mechanisms of action. *Neurotherapeutics, 15*(2), 274–290.

Orlando, B., Egeo, G., Aurilia, C., Fiorentini, G., & Barbanti, P. (2024). Calcitonin Gene-Related Peptide Monoclonal Antibodies: Key Lessons from Real-World Evidence. *Brain Sciences*, *14*(9), 948. https://doi.org/10.3390/brainsci14090948

Öngel, K. (2017). Ağrı tanımı ve sınıflaması. *Klinik Tıp Aile Hekimliği Dergisi*, *9*(1), 12-14.

Özgen, A., & Ünlü, N. (2022). Moleküler kenetleme (docking) yöntemi ile bakteriyel bir sinyal kompleksi ve peonidin molekülü etkileşiminin incelenmesi. *Fırat Üniversitesi Fen Bilimleri Dergisi, 34*(2), 201–206.

Öztürk, V. (2013). Migren Akut Tedavisi. *Nöropsikiyatri Arşivi*, 50 (Özel Sayı 1), 26–29. https://doi.org/10.4274/npa.y7299

Poyner, D. R., Hay, D. L., & Conner, A. C. (2009). CGRP receptor antagonists: Design and screening. *Expert Opinion on Drug Discovery, 4*(12), 1253–1265. https://doi.org/10.1517/17460440903413496

Puledda, F., Silva, E. M., Suwanlaong, K., & Goadsby, P. J. (2023). Migraine: From pathophysiology to treatment. *Journal of Neurology, 270*(10), 3654–3666. https://doi.org/10.1007/s00415-023-11706-1

Rees, T. A., Russo, A. F., O'Carroll, S. J., Hay, D. L., & Walker, C. S. (2022). CGRP and the calcitonin receptor are co-expressed in mouse, rat and human trigeminal ganglia neurons. *Frontiers in Physiology, 13,* 860037. https://doi.org/10.3389/fphys.2022.860037

Rissardo, J. P., & Caprara, A. L. F. (2022). Gepants for acute and preventive migraine treatment: A narrative review. *Brain Sciences, 12*(1612). https://doi.org/10.3390/brainsci12121612

Rizzoli, P., & Mullally, W. J. (2018). Headache. *The American Journal of Medicine, 131*(1), 17–24. https://doi.org/10.1016/j.amjmed.2017.09.005

Roceanu, A., Antochi, F., & Băjenaru, O. (2015). New molecules in migraine treatment. *Farmacia, 63*(4), 475.

Russo, A. F., & Hay, D. L. (2023). CGRP physiology, pharmacology, and therapeutic targets: Migraine and beyond. *Physiological Reviews, 103*(2), 1565–1644. https://doi.org/10.1152/physrev.00059.202

Shibata, Y., Sato, H., Sato, A., & Harada, Y. (2024). Efficacy of lasmiditan as a secondary treatment for migraine attacks after unsuccessful treatment with a triptan. *Neurology International, 16*(3), 643–652. https://doi.org/10.3390/neurolint16030048

Silberstein, S. D. (2010). Migraine: Preventive treatment. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of clinical neurology* (Vol. 97, pp. 337–354). Elsevier. https://doi.org/10.1016/S0072-9752(10)97027-9

Song, Q., Gao, S., & Tan, Y. (2025). Adverse events associated with gepants: A pharmacovigilance analysis based on the FDA Adverse Event Reporting System. *The Journal of Headache and Pain, 26*, 147. https://doi.org/10.1186/s10194-025-02091-3

Soyuer, F., & Kepenek-Varol, B. (2019). Quality of life and pain. *International Journal of Family & Community Medicine*, *3*(2), 110–114. https://doi.org/10.15406/ijfcm.2019.03.00140

Sprenger, T., Viana, M., & Tassorelli, C. (2018). Current prophylactic medications for migraine and their potential mechanisms of action. *Neurotherapeutics, 15*(2), 313–323. https://doi.org/10.1007/s13311-018-0621-8

Steiner, T. J., & Stovner, L. J. (2023). Global epidemiology of migraine and its implications for public health and health policy. *Nature Reviews Neurology, 19*(2), 109–117. https://doi.org/10.1038/s41582-022-00763-1

Szeleszczuk, Ł., & Frączkowski, D. (2022). Propranolol versus other selected drugs in the treatment of various types of anxiety or stress, with particular reference to stage fright and post-traumatic stress disorder. *International Journal of Molecular Sciences, 23*(17), 10099. https://doi.org/10.3390/ijms231710099

Tarı, Ö., & Arpacı, N. (2024). İlaç tasarımında yapay zekâ uygulamaları. *Ankara Eczacılık Fakültesi Dergisi, 48*(1), 327–365.

Tarlacı, S. (2006). Akut migren atağı tedavisi. *Nobel Med, 2*(3), 4–14.

Tfelt-Hansen, P., Saxena, P. R., Dahlöf, C., Pascual, J., Láinez, M., Henry, P., Diener, H.-C., Schoenen, J., Ferrari, M. D., & Goadsby, P. J. (2000). Ergotamine in the acute treatment of migraine: A review and European consensus. *Brain, 123*(1), 9–18.

Ünal-Çevik, I., & Arslan, D. (2023). Similarities and differences between migraine and other types of headaches: Migraine mimics. *Neurology Perspectives, 3*(2), 100122. https://doi.org/10.1016/j.neurop.2023.100122

Vgontzas, A., & Pavlović, J. M. (2018). Pathophysiology mechanisms. *Headache, 58*(7), 1030–1039. https://doi.org/10.1111/head.13358

Walker, C. S., Raddant, A. C., Woolley, M. J., Russo, A. F., & Hay, D. L. (2018). CGRP receptor antagonist activity of olcegepant depends on the signalling pathway measured. *Cephalalgia, 38*(3), 437–451. https://doi.org/10.1177/0333102417691762

Yağcı, Ü., & Saygın, M. (2019). Ağrı fizyopatolojisi. *Medical Journal of Süleyman Demirel University*, *26*(2), 209-220. https://doi.org/10.17343/sdutfd.444237

Yaman, M., Demirkıran, M. K., & Oruç, S. (2007). Migrende baş ağrısını tetikleyici ve kötüleştirici faktörler. *Düzce Tıp Fakültesi Dergisi, 3*, 9–13.

Yao, G., Yu, T., Han, X., Mao, X., & Li, B. (2013). Therapeutic effects and safety of olcegepant and telcagepant for migraine: A meta-analysis. *Neural Regeneration Research, 8*(10), 938–947. https://doi.org/10.3969/j.issn.1673-5374.2013.10.009

Yavuz, K. F., Yavuz, N., Ulusoy, S., Alnıak, İ., & Gökçe Güneş, H. N. (2013). Gerilim ve migren tipi baş ağrılarına eşlik eden işlevsel olmayan bilişsel içerikler ve tutumlar. *Düşünen Adam The Journal of Psychiatry and Neurological Sciences, 26*(1), 12-21. https://doi.org/10.5350/DAJPN2013260101

Yin, Z., Hu, W., Zhang, W., Konno, H., Moriwaki, H., Izawa, K., Han, J., & Soloshonok, V. A. (2020). Tailor-made amino acid-derived pharmaceuticals approved by the FDA in 2019. *Amino Acids, 52*, 1289–1305. https://doi.org/10.1007/s00726-020-02887-4

Yücel, Y. (2008). Migren baş ağrısında tanı ve tedavi yaklaşımları. *Dicle Tıp Dergisi*, *35*(4), 281-286.

Zhu, Z., Tang, Y., Li, L., et al. (2025). The efficacy and safety of zavegepant nasal inhalation versus oral calcitonin-gene related peptide receptor antagonists in the acute treatment of migraine: A systematic review and network meta-analysis of the literature. *Journal of Headache and Pain, 26,* 48. https://doi.org/10.1186/s10194-025-01984-7

Chapter 8

STATISTICAL ANALYSIS OF THE RELATIONSHIP BETWEEN ELECTRIC VEHICLE BATTERY COSTS AND CRITICAL MINERAL PRICES

Çağrı Kibar, Tuba Koç[1]

1 Doç. Dr. Tuba Koç, Çankırı Karatekin University, Faculty of Science, Department of Statistics, ORCID: 0000-0001-5204-0846

1. Introduction

Growing concerns about global climate change and environmental sustainability are accelerating the shift from fossil fuel-based vehicles to electric vehicles (EVs) in the transportation sector. While internal combustion engine vehicles are highly dependent on fossil fuels and cause significant greenhouse gas emissions, electric vehicles contribute to reducing carbon footprints by not producing direct exhaust emissions. Since road transport accounts for more than 15% of global energy-related emissions, this transformation is critical, particularly in large cities, for improving air quality and achieving climate goals (IEA, 2024).

In line with the carbon neutrality targets set out in the Paris Climate Agreement, many countries have made the electrification of transport a strategic priority. The European Union, for example, plans to stop selling new petrol and diesel cars from 2035 onwards, considering this a vital part of its aim to be climate neutral by 2050. The literature emphasises that this process requires emission reductions and structural transformations in energy production, infrastructure and supply chains. Globally, electric vehicle sales surpassed 17 million in 2024, accounting for over 20% of total vehicle sales, and this figure is projected to reach 65% by 2030 (IEA, 2024).

Electric vehicles offer significant environmental and economic advantages. While price fluctuations in fossil fuels can have a negative impact on drivers and fleets, the energy costs of electric vehicles are more predictable and can be more economical in the long term. Furthermore, electric vehicles have fewer moving parts than internal combustion engines and therefore require less maintenance and have a lower risk of failure, which contributes to reduced long-term ownership costs. These factors make electric vehicles competitive with, or even more advantageous than, internal combustion engine vehicles in terms of total cost of ownership (TCO).

Lithium-ion batteries are at the heart of this transformation and constitute one of the most critical components in terms of both technological development and cost structure. They are the preferred energy storage technology for electric vehicles due to their high energy density, long life and relatively low weight. These batteries consist of four basic components: the cathode, anode, electrolyte and separator. The cathode material is the most critical factor in determining the battery's performance and cost. Cathode chemistries commonly used in electric vehicles today include NMC (nickel-manganese-cobalt), NCA (nickel-cobalt-aluminium) and LFP (lithium-iron-phosphate). According to the International Energy Agency (IEA, 2021), lithium, nickel, cobalt, manganese, and graphite are minerals of critical importance in terms of battery performance, lifespan, and energy density.

Among the new generation of battery technologies, solid-state batteries are a particularly important area of research. They have the potential to offer higher energy density and improved safety features compared to traditional liquid electrolyte batteries (Janek and Zeier, 2016). While studies in the literature predict that solid-state batteries could significantly increase the range of electric vehicles, technical challenges such as interfacial resistance, low ionic conductivity and production scalability are hindering the commercialisation of this technology (Janek and Zeier, 2023).

Battery packs account for around 30–40% of the total cost of electric vehicles and are considered one of the most significant economic barriers to their widespread adoption. A comprehensive study by Ziegler and Trancik (2021) found that the price of lithium-ion battery cells has fallen by around 97%

since 1991, decreasing by an average of 20% for every doubling of cumulative production capacity. According to BloombergNEF data, the price of lithium-ion battery packs fell from $684/kWh in 2013 to $139/kWh in 2023. In 2024, they fell by a further 20% to $115/kWh, marking the largest annual decline since 2017 (BloombergNEF, 2024). However, due to increases in raw material prices in 2022, battery prices rose for the first time since 2010 in BloombergNEF's surveys. This clearly demonstrates the importance of mineral prices on battery costs.

Raw material costs currently account for 50–70% of battery pack costs, which is higher than the 40–50% recorded five years ago (IEA, 2021). This means that changes in mineral prices play a more decisive role in battery costs. According to the IEA's analysis, doubling the price of either lithium or nickel would result in a 6% increase in battery costs. If the prices of both minerals were to double simultaneously, the expected cost reductions from doubling production capacity could be completely offset (IEA, 2021). This highlights the strategic importance of mineral price dynamics for the battery industry.

The geographic concentration of critical mineral production has significant implications for supply chain risk. The three countries that produce the most lithium, cobalt and rare earth elements control more than three-quarters of global production. For example, the Democratic Republic of Congo accounts for around 70% of global cobalt production, while China accounts for around 60% of rare earth element production (IEA, 2021). In 2023, Australia, Chile and China extracted around 85% of the world's lithium, with around 65% of the refining taking place in China (IEA, 2024). This concentration makes the supply chain vulnerable to geopolitical risks and supply disruptions.

Various studies in the literature examine the relationship between battery costs and mineral prices. Nykvist and Nilsson (2015) systematically analyzed over 80 cost estimates between 2007 and 2014, revealing that battery pack costs across the industry declined by approximately 14% annually. Schmidt et al. (2017) created experience curves for 11 energy storage technologies and estimated that battery costs would fall to $175±25/kWh when 1 TWh of capacity was installed. Mauler et al. (2021) comprehensively evaluated battery cost estimation methods by reviewing 53 academic studies and predicted that pack costs could fall to approximately $70/kWh by 2050. Ziegler and Trancik (2021) examined the decline in lithium-ion battery prices between 1990 and 2018, highlighting the role of economies of scale and technological advances in cost reduction. In 2023, the supply of cobalt and nickel exceeded demand by 6.5% and 8%, respectively, while the supply of lithium exceeded demand by over 10% (IEA, 2024). This surplus lowered mineral prices and consequently battery costs. However, low mineral prices also have a negative impact on the cash flow and profit margins of mining companies, which could deter future investments and potentially lead to shortages of lithium and nickel by 2030 (IEA, 2024). Therefore, it is crucial to understand the impact of mineral prices on battery costs in both upward and downward price scenarios.

The aim of this study is to statistically examine the prices of critical minerals affecting the cost of lithium-ion battery packs used in electric vehicles. The impact of cobalt, lithium, nickel, manganese, and phosphoric acid price indices on battery costs is analyzed by comparing classical regression methods with penaltyed regression and dimensionality reduction methods that are robust against multicollinearity. Multicollinearity is a significant statistical problem arising from high correlation between independent variables and negatively affecting the estimation of regression coefficients (Hoerl and Kennard, 1970). Penalized methods such as Ridge regression and Lasso regression are commonly used regularization techniques to address this problem (Tibshirani, 2011; Zou and Hastie,

2005). The study demonstrates the applicability of alternative regression techniques in datasets with a limited number of observations and reveals which critical minerals have the most significant impact on battery costs.

2. MATERIALS AND METHODS

2.1. Data Set

The dataset used in this study consists of two main elements. The first element is index data on critical mineral prices. The International Energy Agency (IEA) publishes comprehensive datasets to monitor the market dynamics of minerals of strategic importance in the energy transition process (IEA, 2023; 2024). These datasets are created using raw material price series compiled by Bloomberg and Bloomberg New Energy Finance (BNEF) (IEA, 2024). These indices cover five key battery raw materials: lithium carbonate, cobalt sulfate, nickel sulfate, manganese flakes, and phosphoric acid, and are presented quarterly. In the IEA methodology, mineral prices are normalized using an index structure where the January 2017 value is taken as 100 (IEA, 2023). Index-based price measurement makes it possible to compare commodities traded in different units on a common scale and is widely preferred in the commodity markets literature (Baffes and Etienne, 2016). The second element of the dataset is lithium-ion battery pack prices. BloombergNEF has been tracking global battery price trends through its comprehensive Battery Price Survey, conducted annually since 2010. This research reports volume-weighted average battery pack prices across a wide range of applications, including electric passenger vehicles, buses, commercial vehicles, and stationary energy storage systems. According to the BNEF methodology, prices are presented in real US dollars (USD 2023) and in units of $/kWh.

This study used annual battery pack prices for the period 2015–2023 as the dependent variable. The analysis period was chosen based on data availability and the starting date of the mineral indices. The model included cobalt sulfate, lithium carbonate, nickel sulfate, manganese flake and phosphoric acid price indices as independent variables. The impact of critical mineral prices on battery costs is a significant topic of research in energy economics and clean energy technologies (IEA, 2023).

Using the independent variables in index format and the dependent variable in nominal price format does not create any methodological inconsistency in the model. In econometric modelling studies, modelling cost determinants with index series and the output variable with real prices is a common approach (Krichene, 2002; Chevillon and Rifflart, 2009). Numerous studies in the field of energy economics examine the effects of raw material price indices on final product costs (Zagaglia, 2010; Lang and Auer, 2020). Since the indices consistently reflect relative changes over time, the regression coefficients can be interpreted directly and meaningfully. Consequently, the impact of a one-unit change in the mineral index on battery cost can be evaluated in a statistically consistent manner.

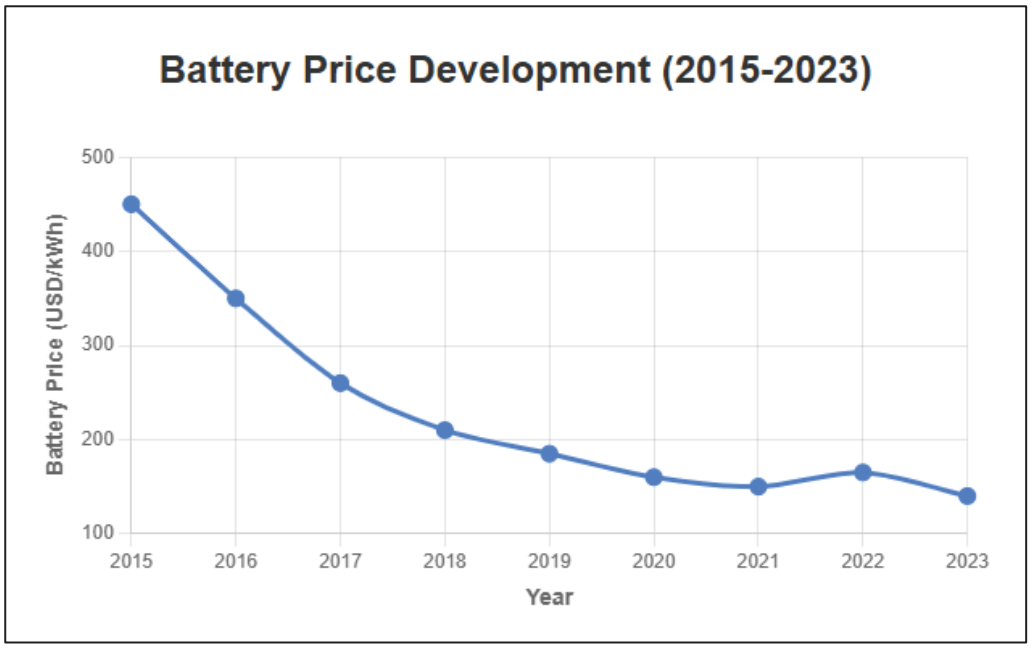


Figure 1. Global Lithium-Ion Battery Pack Price Trends (2015-2023)

As shown in Figure 1, lithium-ion battery pack prices have shown a significant downward trend between 2015 and 2023. The average price, which was $448/kWh in 2015, fell to $139/kWh by 2023, representing a cost reduction of approximately 69% during this period. However, the temporary increase observed in 2022 clearly demonstrates the impact of global supply chain constraints and the sudden rise in critical mineral prices on battery costs.

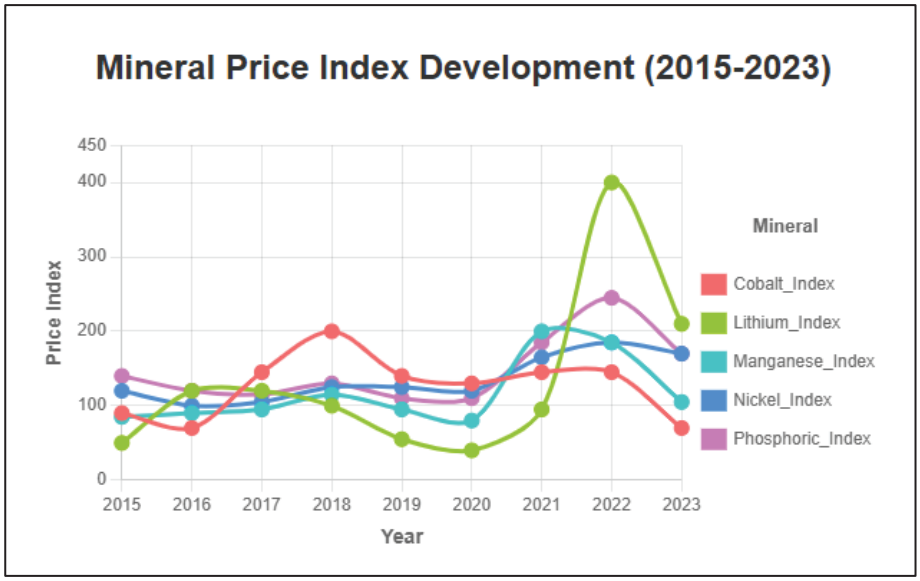


Figure 2. Development of Critical Mineral Price Indices (2015-2023)

Figure 2 shows the trend of the critical mineral price indices used as the study's independent variables during the period from 2015 to 2023. Of these, the lithium carbonate index exhibited the highest volatility; falling to 56 in 2020 and rising to 371 in 2022, it registered an increase of approximately 6.6 times. Similarly, the cobalt sulphate index peaked at 202.5 in 2018 and then fell to 67.5 in 2023. The phosphoric acid and nickel sulphate indices showed significant increases during the period 2021–2022. The manganese flake index followed a trend that was relatively more stable compared to that

of other minerals. These findings reveal significant price fluctuations in critical mineral markets and asymmetrical movements among minerals.

Table 1. Descriptive statistics of variables

Variable	N	Mean	Std. Deviation	Median	Min.	Max	Skewness	Kurtosis
Kobalt Sülfat	9	113.75	46.97	92.00	65.50	202.50	0.53	-1.24
Lityum Karbonat	9	128.89	107.32	100.00	34.50	377.75	1.27	0.38
Nikel Sülfat	9	135.31	30.00	125.25	96.25	185.50	0.39	-1.48
Mangan Flake	9	115.78	43.54	98.50	79.00	205.00	1.01	-0.63
Fosforik Asit	9	145.28	45.93	126.50	100.00	246.75	1.04	-0.15
Batarya Fiyatı ($/kWh)	9	228.33	105.13	183.00	139.00	448.00	0.98	-0.58

Note: Mineral price indices are compiled from the IEA (2024) database, and battery pack prices are compiled from the BloombergNEF (2023) Battery Price Survey. Index values are normalized to January 2017 = 100. Battery prices are expressed in real 2023 US dollars as $/kWh. Skewness and kurtosis values indicate the degree of deviation from normality in the distributions.

The descriptive statistics in Table 1 reveal the basic distribution characteristics of the variables used in the study. The dependent variable, battery pack prices, ranged from $139 to $448/kWh during the 2015-2023 period, with an average of $228.33/kWh. Among the independent variables, the lithium carbonate index exhibited the highest volatility, with a wide fluctuation range of 34.50-377.75 and a standard deviation of 107.32, significantly exceeding other minerals. This high volatility reflects supply-demand imbalances triggered by the rapid increase in electric vehicle demand, particularly during 2021-2022. While the cobalt sulfate index (s = 46.97) showed fluctuations related to geographical concentration and geopolitical risks, the nickel sulfate index (s = 30.00) followed a relatively stable trajectory. The high average phosphoric acid index (145.28) reflects increased demand driven by the widespread adoption of LFP battery chemistry.

2. Methods

2.1. Multiple Linear Regression (OLS)

Multiple linear regression is a classic regression method used to model the relationship between the dependent variable and multiple independent variables through a linear structure. The model is expressed as follows:

$$Y = \beta_0 + \sum_{j=1}^{n} \beta_j x_j + \varepsilon \quad (1)$$

Here, Y represents the dependent variable, x_j represents the independent variables, β_j represents the regression coefficients and ε represents the error term. Coefficient estimates are obtained by minimising the sum of squared residuals using the least squares method (Koç and Akın, 2021).

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2)$$

OLS estimators exhibit BLUE (best linear unbiased estimator) properties if the Gaussian-Markov assumptions are satisfied (Greene, 2020). However, the variance of the estimated coefficients may increase, causing the model to become unstable, if there is high correlation between the independent variables (i.e. multicollinearity) or if the sample size is limited (Kutner et al., 2005; Montgomery et al., 2021). In such cases, regression techniques based on regularisation are suggested as an alternative.

2.2. Ridge Regression

Ridge regression is a regularization technique used to reduce the variance of regression coefficients in datasets containing multicollinearity. In this method, a penalty term containing the squares of the coefficients is added to the OLS residual sum of squares function (Hoerl & Kennard, 1970). This reduces the regression coefficients towards zero but does not eliminate them entirely. The model is formulated as follows:

$$\hat{\beta}_{ridge} = arg_{\beta}^{min}\left\{\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}{\beta_j}^2\right\} \quad (3)$$

Where, $\lambda \geq 0$ is the regularization parameter and controls the severity of the penalty term. When $\lambda = 0$, the model reverts to classical OLS regression, while as λ increases, the coefficient magnitudes decrease. Ridge regression is particularly recommended as an alternative to the OLS method in situations involving highly correlated variables and a limited number of observations (Hastie et al., 2005).

2.3. Lasso Regression

Lasso regression (Least Absolute Shrinkage and Selection Operator) is a regularization-based method that simultaneously performs both the reduction of regression coefficients and the variable selection process (Tibshirani, 2011). Unlike ridge regression, the penalty term is based on the absolute values of the coefficients, allowing some coefficients to be reduced to zero. This simplifies the model and performs variable selection.

The Lasso regression model is expressed as follows:

$$\hat{\beta}_{lasso} = \arg\min_{\beta}\left\{\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\left|\beta_j\right|\right\} \quad (4)$$

Where, $\lambda \geq 0$ is the regularization coefficient and determines the effect of the penalty term. The Lasso method is recommended as an effective alternative, particularly when the number of variables is high relative to the number of observations and in datasets with multicollinearity (Hastie et al., 2005).

2.4. Elastic Net Regression

It is a hybrid regularization method that combines Lasso and Ridge regression methods. It both reduces the coefficients and can perform variable selection (Zou & Hastie, 2005). Elastic Net uses both the L1 norm (Lasso) and the L2 norm (Ridge) as penalty terms. This combines Lasso's ability to reduce coefficients to zero with Ridge's robustness against multicollinearity. The Elastic Net model is expressed as follows:

$$\hat{\beta}_{\text{EN}} = \arg\min_{\beta}\left\{\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\left(\alpha\sum_{j=1}^{p}\left|\beta_j\right| + \frac{1-\alpha}{2}\sum_{j=1}^{p}{\beta_j}^2\right)\right\} \quad (5)$$

Where $\lambda \geq 0$ is the regularization parameter and controls the overall strength of the penalty term, $\alpha \in [0,1]$ is the mixing parameter that determines the relative weight of the L1 and L2 penalties. When $\alpha = 1$, the model becomes Lasso regression (only L1 penalty); when $\alpha = 0$, the model becomes Ridge regression (only L2 penalty). The Elastic Net method is recommended as an effective alternative, particularly for high-dimensional data sets and situations where multicollinearity is strong (Zou and Hastie, 2005; Hastie et al., 2005).

2.5. Principal Component Regression (PCR)

Principal component regression (PCR) is a dimension reduction method used for regression analysis in datasets with high multicollinearity. First, PCR applies Principal Component Analysis (PCA) to the explanatory variables and obtains a small number of components (principal components) that combine the original variables linearly. These components are then used in the regression model to make predictions (Jolliffe, 2011).

The PCR model can be expressed in two stages:

i. PCA is applied to the explanatory variables X to obtain m≤p principal components Z=XW, where W is defined as the loading matrix.
ii. Regression is performed on the target variable y using these components:

$$\hat{y} = Z\hat{\gamma} = XW\hat{\gamma} \quad (6)$$

Here, $\hat{\gamma}$ are the PCR regression coefficients.

The PCR method is recommended as an effective regression method, especially when the number of variables exceeds the number of observations or when there is strong multicollinearity among the explanatory variables (Jolliffe, 2011; Hastie et al., 2005).

3. Results

This section presents the results of analyses conducted to examine the relationship between critical mineral price indices and lithium-ion battery pack prices. First, correlation analysis was performed to reveal the bivariate relationships between the variables. Subsequently, OLS, Ridge, Lasso, Elastic Net, and Principal Component Regression (PCR) methods were applied to develop prediction models. Finally, the performance of these models was compared to determine the most suitable model. Figure 3 shows the correlation matrix between the variables.

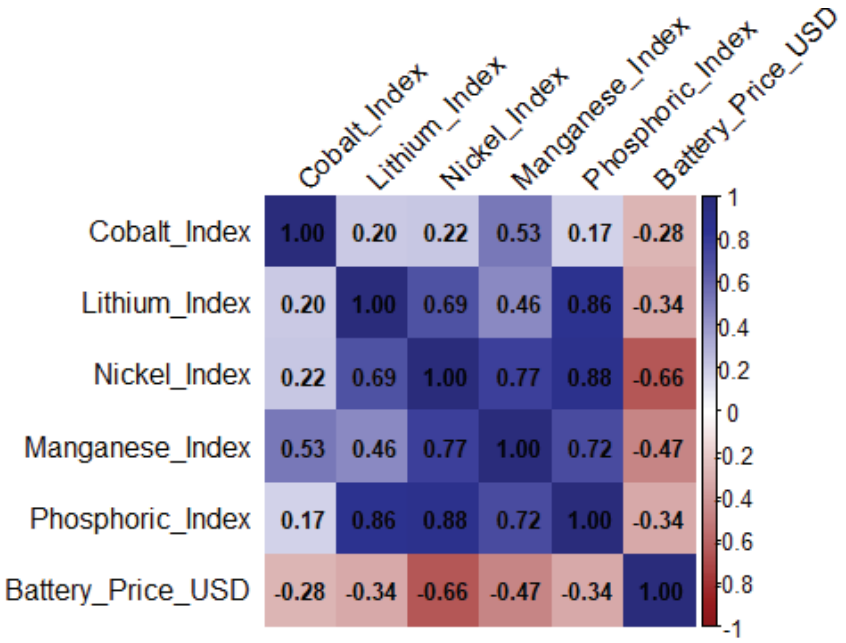


Figure 3. Correlation matrix between critical mineral price indices and lithium-ion battery pack prices

When examining the correlation matrix, it is observed that there are generally strong and positive relationships between critical mineral price indices. In particular, the high correlation between the Nickel_Index and the Phosphoric_Index (r=0.88) and the relationship between the Lithium_Index and the Phosphoric_Index (r=0.86) are noteworthy. However, a relatively strong inverse correlation was observed between battery pack prices (Battery_Price_USD) and, in particular, the Nickel_Index (r=–0.66) and Manganese_Index (r=–0.47). This result is consistent with findings in the literature (e.g., BloombergNEF, 2023; IEA, 2023) indicating that increases in mineral costs do not always directly increase battery costs due to technological developments and economies of scale.The correlation matrix shows strong linear relationships between mineral price indices. As this indicates a risk of multicollinearity in regression analyses, a Variance Inflation Factor (VIF) analysis was performed to examine the level of connection between variables in more detail.

Table 2. VIF values for the variables

Variables	**VIF Value**
Cobalt_Index	2.1474
Lithium_Index	6.5702
Nickel_Index	5.8253
Manganese_Index	5.6986
Phosphoric_Index	15.4873

When examining the VIF values in Table 2, it is observed that there is a very high risk of multicollinearity, particularly in the Phosphoric_Index variable (VIF = 15.49) (O'Brien, 2007).

Additionally, there is a moderate level of multicollinearity in the Lithium_Index (6.57) and Nickel_Index (5.83) variables. These findings are consistent with the strong linear relationships observed in the correlation analysis and indicate that the OLS method alone may not be sufficient during the modeling phase. Therefore, it was deemed appropriate to use regularization methods such as Ridge and Lasso regression to improve estimation performance and reduce the effects of multicollinearity. Furthermore, the PCR method was also used to reduce the effects of high correlation between variables through component transformation.

After identifying the presence and degree of multicollinearity, the performance of different regression methods was comparatively evaluated in light of these findings. Table 3 presents the estimation results for the OLS, Ridge, Lasso, and PCR methods.

Table 3. Model Comparison

Models	**AIC**	**BIC**	**RMSE**
OLS	104.27	105.65	161.29
Ridge	83.03	84.02	86.33
Lasso	83.58	84.77	87.35
Elastic Net	83.05	84.09	86.63
PCR	83.91	84.31	109.47

According to Table 3, the Ridge, Lasso and Elastic Net models fit the data much better than the OLS model, as measured by both AIC and BIC. The RMSE values also reveal that these models have strong generalisation capabilities for new observations (Ridge: 86.33; Elastic Net: 86.63; Lasso: 87.35). In datasets with high variability and potential multicollinearity, such as battery prices, penalised regression techniques (Ridge, Lasso and Elastic Net) provide much more reliable and stable estimates than the classic OLS model. Although the PCR model reduces model complexity through dimension reduction, its estimation performance is limited due to its high RMSE. To assess the ability of the best-performing Ridge regression model to capture time-dependent variation in battery prices, its observed and predicted values were compared.

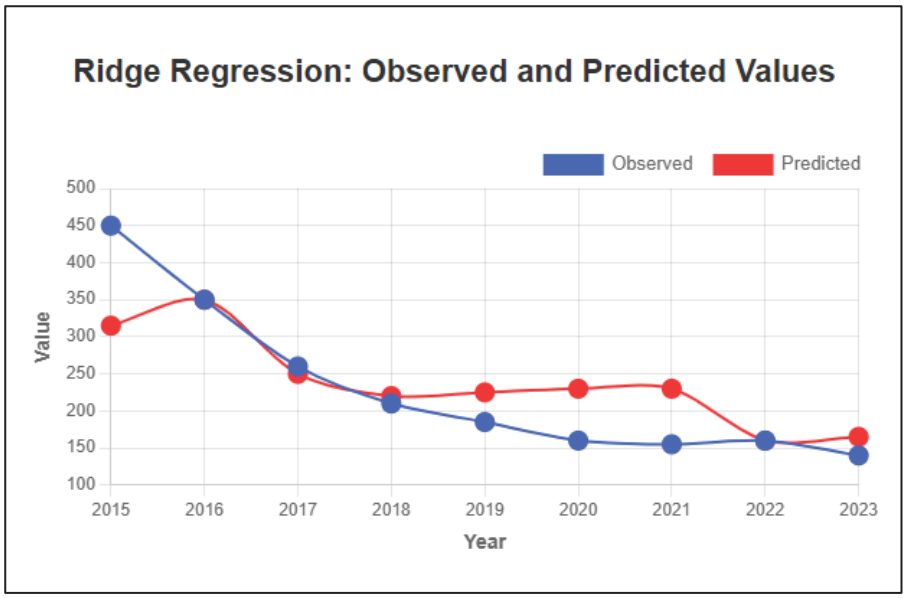


Figure 4. Ridge Regression Model: Observed and Predicted Values Over the Years

As shown in Figure 4, the Ridge regression model successfully captured the general downward trend in battery prices. Battery prices, which were at 448 in 2015, showed a steady decrease in the following years due to technological advancements and increased production scale, falling to 139 in 2023. The parallel trend between the model prediction (red line) and the observed values (blue line) reveals that Ridge regression accurately models the technology-driven downward trend in battery prices.

4. Conclusion and Discussion

This study investigates the impact of critical mineral prices on the cost of lithium-ion battery packs used in electric vehicles. Different regression methods were employed using data from 2015 to 2023. The effects of price indices for cobalt, lithium, nickel, manganese and phosphoric acid on battery costs were analysed by comparing classical ordinary least squares (OLS) regression with regression methods that are robust to multicollinearity, such as Ridge, Lasso and Elastic Net, as well as dimensionality reduction methods, such as principal component regression (PCR). The results of the analysis show that the price of battery packs declined by around 69% during the period under review, falling from $448/kWh to $139/kWh. This downward trend aligns with the global cost reduction trends reported by Ziegler and Trancik (2021) and BloombergNEF (2024). However, the temporary price increase observed in 2022 demonstrates the direct impact of global supply chain constraints, particularly the sharp rise in lithium prices, on battery costs. This finding empirically supports the critical role of mineral price fluctuations in battery economics, as emphasised in reports by the IEA (2021; 2024).

Correlation analysis and VIF values revealed strong linear relationships and the presence of multicollinearity among critical mineral price indices. In particular, the VIF value of the phosphoric acid index (15.49) indicates a very high risk of multicollinearity. Similarly, the lithium (VIF=6.57) and nickel (VIF=5.83) indices also show medium-high levels of correlation. This confirms the risk of multicollinearity negatively affecting regression estimates, as noted by O'Brien (2007) and Kutner et al. (2005). Model comparison results show that penalized regression techniques provide a clear advantage over the classical OLS method. Ridge regression (RMSE=86.33; AIC=83.03), Elastic Net (RMSE=86.63; AIC=83.05), and Lasso (RMSE=87.35; AIC=83.58) models performed much better

than OLS (RMSE=161.29; AIC=104.27) in terms of both goodness-of-fit criteria and prediction accuracy. These findings empirically confirm the ability of regularization techniques proposed by Hoerl and Kennard (1970), Tibshirani (2011), and Zou and Hastie (2005) to produce reliable estimates in datasets with multicollinear structures and limited observations.

Figure 4 shows that the ridge regression model successfully captures the general downward trend in battery prices. The parallel trend between the model estimates and the observed values accurately reflects the cost-reducing effects of technological advances, economies of scale and learning effects on battery costs. This is consistent with the experience curve dynamics predicted by Schmidt et al. (2017) and Mauler et al. (2021). However, the negative correlations observed between minerals and battery prices (particularly nickel: $r = -0.66$) suggest that increases in mineral prices do not necessarily lead to higher battery costs, due to technological improvements and alternative chemical compositions (e.g. the increased use of LFP batteries). This underlines the importance of battery chemistry diversity and cathode material optimisation, as emphasised in reports by BloombergNEF (2023) and the IEA (2024).

The relatively low performance of the PCR model (RMSE=109.47) demonstrates that dimension reduction techniques do not always provide the best solution in a regression context. As noted by Jolliffe (2011) and Hastie et al. (2005), since principal components are created to explain maximum variance, components with weak relationships to the dependent variable may be included in the model. This situation indicates that variable selection or penalty-based methods may be more effective, especially when modeling complex cost structures such as battery prices.

In terms of the practical implications of the study, important results emerge for electric vehicle manufacturers and policymakers. High volatility and geographic concentration in critical mineral prices (e.g., 65% of lithium refining occurs in China) point to strategic risks in terms of supply chain security and price stability. The risk that periods of oversupply, as highlighted by the IEA (2024), could deter future investments and lead to supply gaps after 2030 underscores the importance of long-term planning and diversification strategies. Furthermore, battery recycling and the development of secondary raw material sources hold the potential to reduce mineral dependency.

This study has several limitations. Firstly, the dataset contains nine years of observations (2015–2023) at an annual frequency, which limits its statistical power. Studies using longer time series and quarterly frequency data could provide stronger generalisations. Secondly, the model only used mineral price indices, and other potential determinants such as production volume, technological innovations and supply chain dynamics were not included. Future research could model complex interactions using machine learning approaches, such as Random Forest, Gradient Boosting or Neural Network, if the necessary data is available. In conclusion, this study demonstrates the effectiveness of penalised regression techniques on datasets with limited observations and high multicollinearity, and reveals the role of critical mineral prices in the battery economy. The superior performance of the ridge regression model indicates its potential as a reliable methodological tool for cost projection and risk management studies in the electric vehicle industry.

References

Baffes, J., and Etienne, X. (2015). Analyzing food price trends in the context of Engel? s law and the Prebisch-Singer hypothesis. S Law and the Prebisch-Singer Hypothesis (September 28, 2015). World Bank Policy Research Working Paper, (7424).

BloombergNEF. (2023). Battery price survey 2023. Bloomberg New Energy Finance.

BloombergNEF. (2024). Lithium-ion battery pack prices hit record low of $115/kWh. Bloomberg New Energy Finance.

Chevillon, G., and Rifflart, C. (2009). Physical market determinants of the price of crude oil and the market premium. Energy Economics, 31(4), 537-549.

Greene, W. H. (2008). The econometric approach to efficiency analysis. The measurement of productive efficiency and productivity growth, 1(1), 92-250.

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer, 27(2), 83-85.

Hoerl, A. E., and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67.

International Energy Agency. (2021). The role of critical minerals in clean energy transitions. IEA Publications.

International Energy Agency. (2023). Global critical minerals outlook 2023. IEA Publications.

International Energy Agency. (2024). Global EV outlook 2024. IEA Publications.

Janek, J., and Zeier, W. G. (2016). A solid future for battery development. Nature Energy, 1 (9), 1-4.

Janek, J., and Zeier, W. G. (2023). Challenges in speeding up solid-state battery development. Nature Energy, 8(3), 230-240.

Jolliffe, I. (2011). Principal component analysis. In International encyclopedia of statistical science (pp. 1094-1096). Springer, Berlin, Heidelberg.

Koç, T., and Akın, P. (2021). Comparison of machine learning methods in prediction Gini coefficient for OECD countries. *International Journal of Data Science and Applications*, *4*(1), 16-21.

Krichene, N. (2002). World crude oil and natural gas: a demand and supply model. Energy economics, 24(6), 557-576.

Kutner, M. H. (2005). Applied linear statistical models.

Lang, K., and Auer, B. R. (2020). The economic and financial properties of crude oil: A review. The North American Journal of Economics and Finance, 52, 100914.

Mauler, L., Duffner, F., Zeier, W. G., and Leker, J. (2021). Battery cost forecasting: a review of methods and results with an outlook to 2050. Energy & Environmental Science, 14(9), 4712-4739.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). Introduction to linear regression analysis. John Wiley & Sons.

Nykvist, B., and Nilsson, M. (2015). Rapidly falling costs of battery packs for electric vehicles. Nature climate change, 5(4), 329-332.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. Quality & quantity, 41(5), 673-690.

Schmidt, O., Hawkes, A., Gambhir, A., and Staffell, I. (2017). The future cost of electrical energy storage based on experience rates. Nature Energy, 2(8), 1-8.

Tibshirani, R. J. (2011). Regression shrinkage and selection via the lasso.

World Bank Group. (2024). Global Economic Prospects, June 2024. World Bank Publications.

Zagaglia, P. (2010). Macroeconomic factors and oil futures prices: a data-rich model. Energy Economics, 32(2), 409-417.

Ziegler, M. S., and Trancik, J. E. (2021). Re-examining rates of lithium-ion battery technology improvement and cost decline. Energy & Environmental Science, 14(4), 1635-1651.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2), 301-320.

Chapter 9

DEFINITIONS, TERMS AND CLASSIFICATION OF NANOMATERIALS

Burcu BOZKURT ÇIRAK[1]

1 Doç. Dr. Erzincan Binali Yıldırım Üniversitesi

Meslek Yüksekokulu Elektrik ve Enerji Bölümü

1. Introduction

Nanotechnology has quickly become a frontier of scientific research and technological development, attracting significant attention from a variety of disciplines and industrial sectors. The term ‘nano’ comes from the Greek word for ‘dwarf’ (Warheit, 2018). The Greek term “nano” means “dwarf,” while the Latin “nanos” means “nanus.” (Awan et al., 2025). The concept of nanotechnology was first proposed by Richard Feynman in 1959. In his lecture titled “There's Plenty of Room at the Bottom,” Feynman described the potential for developing a system to manipulate individual atoms and molecules. Before that, nanoparticles (NPs) had existed naturally for millions of years (Vi & Tang, 2014) The International Organization for Standardization says that the prefix 'nano' refers to a size. This size ranges approximately from 1 to 100 nm(Dolez, 2015). In its ground state, a single hydrogen atom has a diameter of around 1 Å, whereas atoms such as carbon, sulfur, or chlorine have a diameter of roughly 2 Å. Therefore, as shown schematically in Figure 1, one nanometer is equivalent to the breadth of ten hydrogen atoms placed side by side (Boustani, 2020). However, if the size becomes very small (e.g. less than 10 nm), the surface energy becomes very high. This results in agglomeration, and, in turn, the high efficiency of nanomaterials is decreased. To overcome this problem, encapsulation (Mukhopadhyaya, 2017) (capping) is required to reduce the surface energy. Materials at the nanometer scale (1 nm to 100 nm, or 10^{-9} m), electronic structures, and systems can be manipulated using nanotechnology processes (Bilal Tahir et al., 2021)attractive-shaped nanoparticles (Salem & Fouda, 2011) Materials specified at this scale are called nanomaterials and exhibit versatile properties that differ from bulk materials. The scientific term for nanotechnology is as follows. The definition of nanomaterials is not straightforward, and this is because they are a relatively new field of study and there is no consensus on a universal definition. Table 1 provides definitions and explanations of terms related to nanomaterials.

Term	Definition	Reference No.
Nanotechnology	Nanotechnology refers to the development of materials, devices, or systems at the nanoscale level by controlling matter to exploit unique properties that emerge at the nano dimension.	(Kreyling et al., 2010)
Nanomanufacturing	Nanomanufacturing refers to production at the nanoscale level and is achieved through bottom-up or top-down fabrication approaches.	(Bhushan, 2017)

Term	Definition	Reference No.
Nanoscale	A dimensional scale ranging from 1 to 100 nm.	(Nasrollahzadeh et al., 2019)
Nanomaterial	If a material has at least one exterior dimension that falls between 1 and 100 nm, it is classified as a nanomaterial.	(Trotta & Mele, 2019)
Nano-object	A distinct substance in the nanoscale range with one, two, or three exterior dimensions.	(Jeevanandam et al., 2018; Shtykov, 2018)
Nanoparticle	A nano-object in which all three dimensions are within the nanoscale (1–100 nm).	(Trotta & Mele, 2019)
Aspect Ratio	The ratio of a nano-object's major axis length to its minor axis width.	(Murphy & Jana, n.d.)
Nanosphere	A roughly spherical nanoparticle with an aspect ratio of 1.	(Murphy & Jana, n.d.)
Nanorod	A nano-object with one dimension significantly longer than the others, having a width of 1–100 nm and an aspect ratio greater than 1.	(Jeevanandam et al., 2018; Murphy & Jana, n.d.)
Nanofiber	Two dimensions of the nano-object fall within the nanoscale range, whereas a third dimension is much greater.	(Trotta & Mele, 2019)
Nanowire	Similar to a nanorod but with a higher aspect ratio, typically used in electronic and sensor applications.	(Murphy & Jana, n.d.)
Nanotube	A hollow nanofiber structure. Carbon nanotubes are the most common example.	(Trotta & Mele, 2019)
Nanostructured Material	A material that contains structural elements, crystallites, clusters, or molecular features with dimensions in the range of 1–100 nm.	(Moriarty, 2001)
Engineered Nanomaterials	Intentionally designed or manufactured materials that have one or more dimensions ≤100 nm.	(Kreyling et al., 2010)
Nanocomposite	A multicomponent material system in which at least one of the phases contains nanoscale domains.	(Baig et al., 2021)

Table 1. A rundown of the different names connected to nanomaterials. (Baig et al., 2021)

Organization	Term	Definition	Ref.
European Commission: Cosmetic Products Regulation	Nanomaterials	Insoluble, bio-persistent, and intentionally created material with one or more external dimensions or an interior structure that ranges in size from 1 to 100 nm.	(Atanassov et al., 2017) (Hischier & Walser, 2012)
American Chemistry Council	Engineered nanomaterial	Any purposely produced material with a size in 1, 2, or 3 dimensions between 1 and 100 nanometers. Materials without novel features compared to bulk materials are excluded.	(Rauscher et al., 2017)
European Commission	Nanomaterials	Any natural, incidental, or manufactured material that contains unbound, aggregated or agglomerated particles of which ≥50% are in the 1–100 nm size range.	(Sharma et al., 2012)
European Commission for novel foods	Nanomaterials	Purposely produced material with ≥1 dimension ≤100 nm or composed of discrete functional parts many with ≥1 dimension ≤100 nm.	(Sharma et al., 2012)
British Standards Institution	Nano-object	Material with one or more peripheral nanoscale dimensions.	(Slezakova et al., 2013)
British Standards Institution	Nanoparticle	An object with three dimensions at the nanoscale.	(Dahm et al., 2018)
British Standards Institution	Nanofiber	Nanomaterial with two similar exterior nanoscale dimensions and one larger dimension.	(Zidan et al., 2019)
British Standards Institution	Nanocomposite	A multiphase structure containing at least one nanoscale phase.	(Bathi et al., 2021)
British Standards Institution	Nanostructure	Structure of interconnected constituent parts in the nanoscale region.	(Jeevanandam et al., 2018)
British Standards Institution	Nanostructured materials	Materials with internal or surface nanostructures.	(Chen & Ke, 2016)
British Standards Institution	Nanoscale	The size is between 1 and 100 nanometres.	(Soltani & Pouypouy, 2019)
ISO Technical Committees	Nanotechnology	Scientific information used to manipulate and control matter at the nanoscale for unique structure-specific properties.	(Kallinger et al., n.d.)
European Patent Office	Nanotechnology	Entities with a controlled geometrical size of at least one functional component	(Hammani et al., 2020)

		of 100 nm or less are described, leading to size-dependent effects.	
ANSI Nanotechnology Standards Panel	Nanotechnology	Understanding and controlling matter of ~1–100 nm enabling unique applications and phenomena.	(Bowman et al., 2018)
British Standards Institution	Nanoscience	Study of nano-scaled matter to understand size- and structure-dependent features.	(Soltani & Pouypouy, 2019)
EU Scientific Committee on Consumer Products	Nanomaterials	Materials that have ≥1 external dimension or internal structure in the nanoscale and may display novel features.	(Kazakov et al., 2020)

Table 2. Nanotechnology definition by international organizations.(Barhoum et al., 2022)

1. Properties of Nanomaterials

The optical behavior of the nanomaterials is determined by their size and shape. The fundamental cause of different optical properties is the electronic distribution within the metal (Krishna Podagatlapalli, 2025a)

The reason for the change in the optical, electrical, thermal and magnetic properties of nanomaterials is the quantum confinement of electrons at the nanoscale. The properties of nanomaterials are significantly altered by size, shape and material composition (Jeevanandam et al., 2018; Krishna Podagatlapalli, 2025b).

As the size of the materials gets closer and closer to the nanoscale or the scales near the de Broglie wavelengths of electrons/holes, the periodic boundary conditions of the carriers change a lot. Additionally, the surface will be able to accommodate a greater number of atoms at the nanoscale. Consequently, compared to bulk materials, nanomaterials display highly distinctive properties(D. Guo et al., 2014; Schmid, 20052004).

Weak interaction forces between particles are known as the driving force behind all physical and mechanical properties of nanomaterials. These forces, also called van der Waals forces (vdW), act between molecules in nanomaterials and colloids(Krishna Podagatlapalli, 2025b).

Nanomaterials are better than conventional materials because of their special qualities. Nanomaterials possess excellent mechanical properties due to the volume, surface and quantum effects of nanoparticles. Adding nanoparticles to a material refines it, leading to the formation of an intra- or inter-particle structure. This strengthens the boundaries between particles and improves the material's mechanical properties (J. L. Wang & Meng, 2014; X. Wang et al., 2011; Wu et al., 2020; Zou et al., 2006). The mechanical properties of a material are defined by how it responds to different conditions,

environments and external forces. As with traditional materials, the mechanical properties of nanomaterials include strength, brittleness, hardness, toughness, fatigue strength, plasticity, elasticity, ductility and yield stress (Joudeh & Linke, 2022a; Wu et al., 2020).

Nanomaterials behave quite differently from the identical materials at bigger dimensions for two key reasons: Quantum effects and surface effects (Roduner, 2006). The precise composition, shape and size of nanomaterials are what determine their physical and chemical properties.

Nanomaterials take on a surprise new persona with unique characteristics and abilities when their quantum structure and size are altered. Several types of nanomaterials, like nanostructures, nanoclusters, and nanosheets, can be distinguished entrenched on their extent of range. As particles interact with one another, their physical properties will evolve.

2. Classification of Nanomaterials based on dimension

Nanomaterials can be categorized in various ways due to their different sizes, shapes, dimensions, compositions, porosities, phases, and homogeneities. A material is classified as a nanomaterial if at least one of its dimensions is measured in nanometers. In other words, the dimensionality of a material is the fundamental characteristic that distinguishes the types of nanostructured materials (NSM) from one another(Paras et al., 2023; Tiwari et al., 2012). Generally speaking, there are four types of nanomaterials: zero-dimensional, one-dimensional, two-dimensional, and three-dimensional. Zero-dimensional, one-dimensional, two-dimensional, and three-dimensional nanomaterials are the four general categories into which nanomaterials fall.

2.1 Zero-dimensional nanomaterials

Materials with dimensions in the nanometer range are categorized as zerodimensional (0D) nanomaterials. Examples of 0D nanomaterials include nanoparticles, quantum dots, carbon nanodots and fullerene. Significant advances have been made in the field of 0D nanomaterials during past decades, with researchers developing new methods of synthesis and applying these to a variety of applications(Paras et al., 2023) . The structural properties of 0D nanomaterials, such as their ultra-small size and higher surface-to-volume ratio, provide enriched active regions per unit mass (Z. Wang et al., 2020).

The structural properties of 0D nanostructures, characterized by a high surface-to-volume ratio, ultra-small dimensions, and high active edge areas per unit mass, stem from their inherent spherical or semi-spherical geometries. The behavior of 0D nanoparticles is intrinsically linked to their electrical, optical, mechanical, and biological properties, as well as other intrinsic characteristics, due to their

diameters being less than 100 nm. This causes them to exhibit edge and quantum confinement effects (Liu et al., 2020).

2.2 One-dimensional nanomaterials

1 D Nanomaterials: These nanomaterials have two dimensions (x and y) in the nanoscale range, but one dimension of the nanostructure is outside the nanometric range (>10 nm). Examples of 1D nanomaterials include nanofibres, nanotubes, nanohorns, nanorods, thin films and nanowires. These are needle-shaped materials. The materials can be either amorphous or crystalline. They can be single crystalline or polycrystalline, and they can be chemically pure or impure. They can be used as standalone materials or they can be embedded within another medium, such as a metallic, ceramic, or polymeric one. 1D nanoparticles can be metallic, ceramic or polymeric (Afolalu et al., 2019; Buzea et al., 2007; Cho et al., 2019; Joudeh & Linke, 2022b; *Khan: Nanoparticle-Based Polymer Composites - Google Scholar*, n.d.; S. Khan & Hossain, 2022a; Mekuye & Abera, 2023).

Depending on parameters such as cross-sectional geometry and aspect ratio (the ratio of the diameter to the length of the nanostructure), one-dimensional nanostructures are divided into three categories: nanorods, nanowires and nanotubes. The largest dimension is measured in relation to the smallest dimension. A one-dimensional nanostructure of nanorods is characterized by a small aspect ratio, while a nanowire is characterized by a large aspect ratio If the cross-sectional geometry is spherical and hollow, it is called a nanotube (Adul-Rasool et al., 2024; Knauth & Schoonman, 2005; Law et al., 2005).

One-dimensional (1D) nanomaterials consist of hollow tubes or solid fiber structures. All 1D nanomaterials have a very high length-to-diameter ratio, which is important for their tubular structure. Carbon-based 1D nanomaterials are the most widely studied, and carbon nanotubes (CNTs) are the best-known example. Structurally, CNTs can be described as hollow cylindrical tubes composed of nanometer-scale coiled graphene sheets held together by van der Waals interactions. Following CNTs, carbon nanowires and nanofibers synthesized in solid form instead of hollow tubes have achieved similar importance (Lisuzzo et al., 2018).

2.3 Two-dimensional nanomaterials

Two-dimensional (2D) nanostructures have two dimensions that are outside the nanometric size range (Fatima et al., 2022; Tiwari et al., 2012). Because of this particular structure, electrons may travel freely in two dimensions, giving rise to extraordinary chemical and physical capabilities. Due to their unique shapes, 2D nanomaterials possess large surface and anisotropic physical/chemical properties (Chimene et al., 2015; Tan et al., 2017).

Two-dimensional (2D) nanomaterials are the thinnest materials. Strong in-plane connections and weak van der Waals forces between layers define its honeycomb-like layered structure (H. Zhang, 2015).

Common examples of 2D nanomaterials with plate-like structures are nanofilms, nanolayers, and nanocoatings. Graphene, a single layer of carbon that is a component of graphite's honeycomb structure, is the most recent example of this type (Hadef, 2018; Asha et al.2020).

By controlling their size and morphology, the physical properties of 2D nanomaterials can be effectively tuned. The electronic properties of 2D nanomaterials exhibit quantum confinement effects. This phenomenon occurs when their size approaches the characteristic length scale of electrons, resulting in discrete energy levels and bandgap modulation. As size decreases, the bandgap increases, resulting in semiconductor behavior. The ability to tune electronic properties based on size creates the opportunity to customize the conductivity and band structure of 2D nanomaterials for specific applications (Goggin & Samaniuk, 2021; Y. T. Guo & Yi, 2023).

2.4 Three-dimensional nanomaterials

Materials that are nanoscale in all three dimensions are known as 3D nanomaterials. These include colloids, free nanoparticles with different shapes, and thin films produced under circumstances that provide atomic-scale porosity.

3D nanostructures can be easily created by connecting several building blocks extending in all directions or by combining a 2D sheet-like structure with other building blocks(Shehzad et al., 2016). 3D nanomaterials are materials with crystalline nanoporous structures that have a rigid cage, framework or cavities. Examples of 3D nanostructures include aerogels, foams, fibres, pillars and layered structures skeletons (Hasanpour & Hatami, 2020; Saleh, 2020).

0D nanomaterials, such as nanoparticles, have a high surface-to-volume ratio. Surface area and accessibility is limited. These particles tend to aggregate. The greater accessible surface area of one-dimensional (1D) nanomaterials, such as nanowires, is due to their elongated structure. However, they still tend to clump together. This limits exposure of their surface area. Graphene and other two-dimensional (2D) nanomaterials have a very large theoretical surface area. However, the 2D sheets tend to stack on top of each other, thereby reducing the accessible surface area. By contrast, 3D nanomaterials have interconnected macro/mesoporous structures that prevent aggregation and restacking. This gives them the highest accessible surface area of all nanomaterials (Verma et al., 2023).

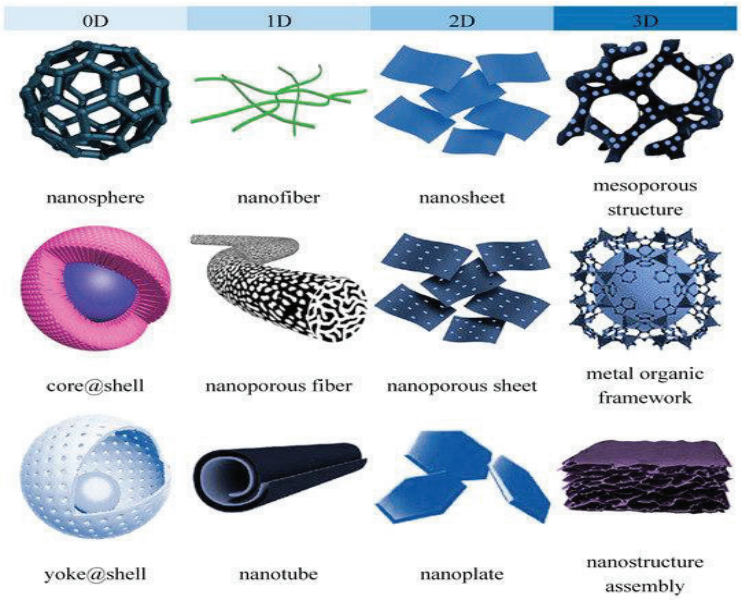


Figure1. Schemes for the application of nanomaterials of different dimensions (Goh et al., 2020)

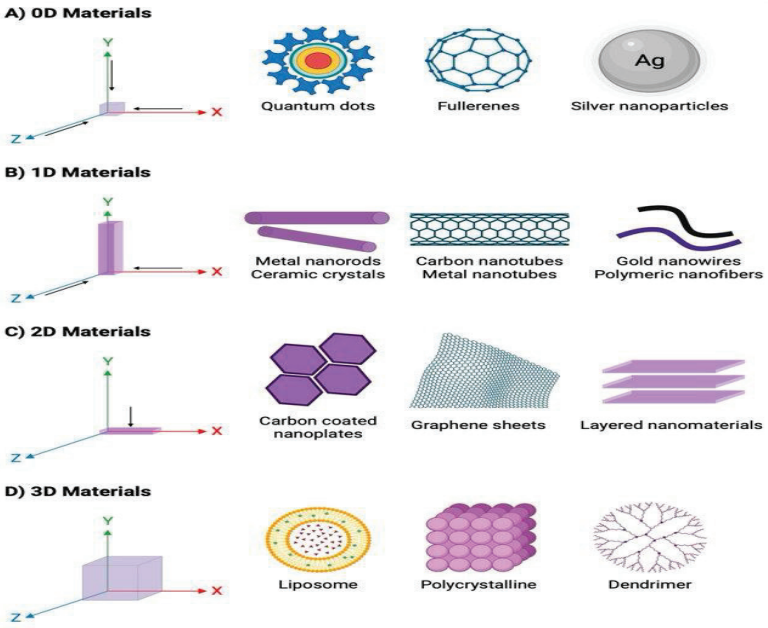


Figure 2. Classification of nanomaterials based their dimensionality. (A) 0D Materials: quantum dots, fullerenes and silver nanoparticles; (B) 1D Materials: metal nanorods, ceramic crystals, carbon nanotubes, metal nanotubes, and gold nanowires; (C) 2D Materials: carbon coated nanoplates, graphene she (Gutiérrez Coronado et al., 2025)

3. Classification of nanomaterials based on origin

Nanomaterials are divided into two groups, natural and artificial, based on their origin.

3.1 Natural nanomaterials

Viruses, protein molecules, minerals like clay, natural colloids like milk and blood (liquid colloids), fog (aerosol type), gelatin (gel type), mineralized natural materials like shells, corals, and bones, insect wings and opals, spider silk, lotus leaves, gecko feet, volcanic ash, and ocean spray are all examples of natural nanomaterials (Mekuye & Abera, 2023). These nanoparticles and nanostructured materials have no connection to human actions or processes; instead, they are created by mechanical or natural (bio)geochemical processes. Natural inorganic nanomaterials are formed by crystal growth. (Barhoum et al., 2022).

3.2 Artificial nanomaterials

Artificial nanoparticles and nanostructured materials are produced for specific applications based on their size and specific properties. Semiconductor nanoparticles and carbon nanotubes are prime examples of synthetic nanomaterials. They are produced deliberately through mechanical and manufacturing processes (Biswas et al., 2022; Cho et al., 2019; Mekuye & Abera, 2023).

4. Classification of nanomaterials based on the structural configuration/composition

4.1 Carbon-based nanomaterials:

The unique morphological and electronic properties of carbon-based nanomaterials make them of critical importance in multidisciplinary applications. Carbon-based nanomaterials frequently consist of carbon-rich structures. These structures appear in diverse geometries. These include hollow tubes, ellipsoids and spherical forms. Examples of these materials include carbon nanotubes, graphene, activated carbon, carbon black, fullerenes and carbon nanofibers. The potential applications of these materials are wide-ranging, including energy storage, water and wastewater treatment, sensor technologies, electrocatalysis, and biomedical fields. Their special qualities—such as their large surface area, robust chemical bonding structure, electrical conductivity, and functionalizable surface chemistry—are the reason for this. (Kumar & Kumbhat, 2019; Rao et al., 2021) .

4.2 Organic Nanomaterials

These organic nanoparticles (ONPs) are synthesized from organic molecules measuring 100 nm or less(Qi & Zhang, 2022). Proteins, carbohydrates, lipids, polymers and other organic compounds are used to make NPs, which are what this class comprises (Pan & Zhong, 2025). Common examples of organic nanoparticles or polymers are ferritin, liposomes, dendrimers, and micelles. These nanoparticles are non-toxic, biodegradable, and sensitive to heat and light as well as thermal and

electromagnetic radiation. Certain particles, such micelles and liposomes, have a hollow center, also referred to as nano capsules (Behari et al., 2008).

4.3 Inorganic Nanomaterials

Inorganic nanoparticles are those which do not contain carbon atoms. Metals and metal oxides are typically used to define what is meant by 'inorganic nanoparticles' (Y. Khan et al., 2022). The inorganic nanoparticles are not toxic and are therefore safe to use. They are biocompatible and hydrophilic. They are also more stable than organic nanoparticles. The synthesis of metal nanoparticles is influenced by their shape, facets and size. Well-known metal nanoparticles include those of aluminum, gold, iron, lead, silver, cobalt, zinc, cadmium and copper (Ijaz et al., 2020).The composition of metal oxide nanoparticles, also known as metal oxide nanomaterials, is of positive metallic ions and negative oxygen ions. Frequently synthesized examples include silicon dioxide (SiO_2), titanium dioxide (TiO_2), zinc oxide (ZnO) and aluminum oxide (Al_2O_3). These nanoparticles exhibit remarkable properties compared to their metallic counterparts (Y. Khan et al., 2022).

4.4 Composite Nanomaterials

A nanocomposite is a type of nanomaterial that consists of one or more phases in the nano-size range (zero-dimension, one dimension, and two dimensions) which are embedded in a ceramic, metal, or polymer material. Composites nanostructured materials are constituted of nanoparticles in combination with other nanoparticles, nanoparticles in combination with larger-scale materials, and nanomaterials in combination with bulk-type materials. The use of nanomaterials is already having a significant impact on the mechanical, thermal and flame-retardant properties of a wide range of products, from auto parts to packaging materials (Gu et al., 2022; S. Khan & Hossain, 2022b; P. Zhang et al., 2021).

Conclusion

The definitions, basic ideas, and categorization rules of nanomaterials have all been looked at in this chapter within a methodical framework. Because nanotechnology is developing so quickly and is used in so many different fields, terminology must be used correctly and consistently. Because of this, the chapter was written to remove terminological uncertainties and give readers a strong theoretical basis in the subject of nanomaterials. The creation of basic scientific knowledge and the appropriate direction of application-oriented material design processes are both aided by a thorough analysis of dimensions, morphological, and composition-based classifications. Selecting optimal

synthesis techniques, understanding characterisation results, and identifying material attributes appropriate for specified applications all depend on accurate classification. In this regard, the chapter provides a thorough method that enhances the conceptual foundation of nanomaterial science and allows researchers to assess various findings within a logical scientific framework.In terms of vocabulary and classification, this chapter provides a trustworthy place to start for students who are unfamiliar with nanomaterial research as well as a resource for seasoned researchers. It is anticipated that the conceptual framework offered here will offer a theoretical foundation that will direct the creation of application-oriented nanosystems, the modeling of multicomponent nanomatrices, and the design of next-generation nanostructures.

In summary, this chapter seeks to improve multidisciplinary communication among researchers and create conceptual coherence in the field of nanomaterials. It is expected that the all-encompassing strategy offered will add to the corpus of current literature and act as a reference for the scientific and technological development of nanotechnology.

Reference

Adul-Rasool, A. A., Athair, D. M., Zaidan, H. K., Rheima, A. M., Al-Sharify, Z. T., Mohammed, S. H., & kianfar, E. (2024). 0,1,2,3D nanostructures, types of bulk nanostructured materials, and drug nanocrystals: An overview. In *Cancer Treatment and Research Communications* (Vol. 40). Elsevier Ltd. https://doi.org/10.1016/j.ctarc.2024.100834

Afolalu, S. A., Soetan, S. B., Ongbali, S. O., Abioye, A. A., & Oni, A. S. (2019). Morphological characterization and physio-chemical properties of nanoparticle - Review. *IOP Conference Series: Materials Science and Engineering*, *640*(1). https://doi.org/10.1088/1757-899X/640/1/012065

Asha, A. B., & Narain, R. (2020). Nanomaterials properties. In *Polymer Science and Nanotechnology: Fundamentals and Applications* (pp. 343–359). Elsevier. https://doi.org/10.1016/B978-0-12-816806-6.00015-7

Atanassov, A., Trifonova, S., Saraivanova, J., Pramatarov, A., Atanassov, A., Trifonova, S., Saraivanova, J., & Pramatarov, A. (2017). Assessment of the administrative burdens for businesses in Bulgaria according to the national legislation related to the European Union internal market. *Management : Journal of Contemporary Management Issues*, *22*(Special Issue), 21–49. http://www.opac.gov-

Awan, M. S. U. D., Anwar, M. T., Khan, H. I., Asghar, M. R., Raza, M. R., Husnain, N., Hussain, M., & Rasheed, T. (2025). Nanoparticles' classification, synthesis, characterization and applications— A review. *Characterization and Application of Nanomaterials*, *8*(1). https://doi.org/10.24294/can8899

Baig, N., Kammakakam, I., Falath, W., & Kammakakam, I. (2021). Nanomaterials: A review of synthesis methods, properties, recent progress, and challenges. In *Materials Advances* (Vol. 2, Issue 6, pp. 1821–1871). Royal Society of Chemistry. https://doi.org/10.1039/d0ma00807a

Barhoum, A., García-Betancourt, M. L., Jeevanandam, J., Hussien, E. A., Mekkawy, S. A., Mostafa, M., Omran, M. M., Abdalla, M. S., & Bechelany, M. (2022). Review on Natural, Incidental, Bioinspired, and Engineered Nanomaterials: History, Definitions, Classifications, Synthesis, Properties, Market, Toxicities, Risks, and Regulations. *Nanomaterials*, *12*(2). https://doi.org/10.3390/nano12020177

Bathi, J. R., Moazeni, F., Upadhyayula, V. K. K., Chowdhury, I., Palchoudhury, S., Potts, G. E., & Gadhamshetty, V. (2021). Behavior of engineered nanoparticles in aquatic environmental samples: Current status and challenges. *Science of The Total Environment*, *793*, 148560. https://doi.org/10.1016/J.SCITOTENV.2021.148560

Behari, J., Tiwari, D. K., & Sen, P. (2008). Application of Nanoparticles in Waste Water Treatment. *World Applied Sciences Journal*, *3*(3), 417–433.

Bhushan, B. (Ed.). (2017). *Springer Handbook of Nanotechnology*. https://doi.org/10.1007/978-3-662-54357-3

Bilal Tahir, M., Sagir, M., & Asiri, A. M. (2021). *Nanomaterials: Synthesis, Characterization, Hazards and Safety*.

Biswas, M. C., Chowdhury, A., Hossain, M. M., & Hossain, M. K. (2022). Applications, drawbacks, and future scope of nanoparticle-based polymer composites. *Nanoparticle-Based Polymer Composites*, 243–275. https://doi.org/10.1016/B978-0-12-824272-8.00002-6

Boustani, I. (2020). Two-Dimensional Sheets. In *Springer Series in Materials Science* (Vol. 290, pp. 285–362). Springer. https://doi.org/10.1007/978-3-030-32726-2_6

Bowman, D. M., May, N. D., & Maynard, A. D. (2018). Nanomaterials in Cosmetics: Regulatory Aspects. *Analysis of Cosmetic Products: Second Edition*, 289–302. https://doi.org/10.1016/B978-0-444-63508-2.00012-6

Buzea, C., Pacheco, I. I., & Robbie, K. (2007). Nanomaterials and nanoparticles: Sources and toxicity. *Biointerphases*, *2*(4), MR17–MR71. https://doi.org/10.1116/1.2815690

Chen, X., & Ke, C. (2016). Structural and physical properties of boron nitride nanotubes and their applications in nanocomposites. *Boron Nitride Nanotubes in Nanomedicine*, 183–199. https://doi.org/10.1016/B978-0-323-38945-7.00012-2

Chimene, D., Alge, D. L., & Gaharwar, A. K. (2015). Two-Dimensional Nanomaterials for Biomedical Applications: Emerging Trends and Future Prospects. *Advanced Materials*, *27*(45), 7261–7284. https://doi.org/10.1002/ADMA.201502422;REQUESTEDJOURNAL:JOURNAL:15214095;WGROUP:STRING:PUBLICATION

Cho, G., Park, Y., Hong, Y. K., & Ha, D. H. (2019). Ion exchange: an advanced synthetic method for complex nanoparticles. In *Nano Convergence* (Vol. 6, Issue 1). Korea Nano Technology Research Society. https://doi.org/10.1186/s40580-019-0187-0

Dahm, M. M., Schubauer-Berigan, M. K., Evans, D. E., Birch, M. E., Bertke, S., Beard, J. D., Erdely, A., Fernback, J. E., Mercer, R. R., & Grinshpun, S. A. (2018). Exposure assessments for a cross-sectional epidemiologic study of US carbon nanotube and nanofiber workers. *International Journal of Hygiene and Environmental Health*, *221*(3), 429–440. https://doi.org/10.1016/J.IJHEH.2018.01.006

Dolez, P. I. (2015). Nanomaterials Definitions, Classifications, and Applications. *Nanoengineering: Global Approaches to Health and Safety Issues*, 3–40. https://doi.org/10.1016/B978-0-444-62747-6.00001-4

Fatima, J., Shah, A. N., Tahir, M. B., Mehmood, T., Shah, A. A., Tanveer, M., Nazir, R., Jan, B. L., & Alansi, S. (2022). Tunable 2D Nanomaterials; Their Key Roles and Mechanisms in Water Purification and Monitoring. In *Frontiers in Environmental Science* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fenvs.2022.766743

Goggin, D. M., & Samaniuk, J. R. (2021). 2D Colloids: Size- and Shape-Controlled 2D Materials at Fluid–Fluid Interfaces. *Langmuir*, *37*(48), 14157–14166. https://doi.org/10.1021/ACS.LANGMUIR.1C02418

Goh, P. S., Wong, K. C., & Ismail, A. F. (2020). Nanocomposite membranes for liquid and gas separations from the perspective of nanostructure dimensions. In *Membranes* (Vol. 10, Issue 10, pp. 1–29). MDPI AG. https://doi.org/10.3390/membranes10100297

Gu, Q., Lv, J., Mo, X., & Jiang, X. (2022). High aspect ratio metamaterials and their applications. *Sensors and Actuators A: Physical*, *335*, 113376. https://doi.org/10.1016/J.SNA.2022.113376

Guo, D., Xie, G., & Luo, J. (2014). Mechanical properties of nanoparticles: Basics and applications. In *Journal of Physics D: Applied Physics* (Vol. 47, Issue 1). https://doi.org/10.1088/0022-3727/47/1/013001

Guo, Y. T., & Yi, S. S. (2023). Recent Advances in the Preparation and Application of Two-Dimensional Nanomaterials. In *Materials* (Vol. 16, Issue 17). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ma16175798

Gutiérrez Coronado, O., Sandoval Salazar, C., Muñoz Carrillo, J. L., Gutiérrez Villalobos, O. A., Miranda Beltrán, M. de la L., Soriano Hernández, A. D., Beltrán Campos, V., & Villalobos Gutiérrez, P. T. (2025). Functionalized Nanomaterials in Cancer Treatment: A Review. In *International Journal of Molecular Sciences* (Vol. 26, Issue 6). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms26062633

Hadef, F. (2018). *An Introduction to Nanomaterials*. 1–58. https://doi.org/10.1007/978-3-319-76090-2_1

Hammani, S., Moulai-Mostefa, N., Samyn, P., Bechelany, M., Dufresne, A., & Barhoum, A. (2020). Morphology, Rheology and Crystallization in Relation to the Viscosity Ratio of Polystyrene/Polypropylene Polymer Blends. *Materials 2020, Vol. 13, Page 926*, *13*(4), 926. https://doi.org/10.3390/MA13040926

Hasanpour, M., & Hatami, M. (2020). Application of three dimensional porous aerogels as adsorbent for removal of heavy metal ions from water/wastewater: A review study. In *Advances in Colloid and Interface Science* (Vol. 284). Elsevier B.V. https://doi.org/10.1016/j.cis.2020.102247

Hischier, R., & Walser, T. (2012). Life cycle assessment of engineered nanomaterials: State of the art and strategies to overcome existing gaps. *Science of The Total Environment*, *425*, 271–282. https://doi.org/10.1016/J.SCITOTENV.2012.03.001

Ijaz, I., Gilani, E., Nazir, A., & Bukhari, A. (2020). Detail review on chemical, physical and green synthesis, classification, characterizations and applications of nanoparticles. *Green Chemistry Letters and Reviews*, *13*(3), 59–81. https://doi.org/10.1080/17518253.2020.1802517

Jeevanandam, J., Barhoum, A., Chan, Y. S., Dufresne, A., & Danquah, M. K. (2018). Review on nanoparticles and nanostructured materials: History, sources, toxicity and regulations. In *Beilstein Journal of Nanotechnology* (Vol. 9, Issue 1, pp. 1050–1074). Beilstein-Institut Zur Forderung der Chemischen Wissenschaften. https://doi.org/10.3762/bjnano.9.98

Joudeh, N., & Linke, D. (2022a). Nanoparticle classification, physicochemical properties, characterization, and applications: a comprehensive review for biologists. In *Journal of Nanobiotechnology* (Vol. 20, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12951-022-01477-8

Kallinger, C., Veffkind, V., Michalitsch, R., Bus., Y. V.-Nanotech. L. &, & 2008, undefined. (n.d.). Patenting nanotechnology: a European patent office perspective. *HeinOnlineC Kallinger, V Veffkind, R Michalitsch, Y VerbandtNanotech. L. & Bus., 2008•HeinOnline*. Retrieved December 15, 2025, from https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/nantechlb5&section=14

Kazakov, S., Ruiz-Alba, J. L., & Muñoz, M. M. (2020). The impact of information and communication technology and internal market orientation blending on organisational performance in small and medium enterprises. *European Journal of Management and Business Economics*, *30*(2), 129–151. https://doi.org/10.1108/EJMBE-04-2020-0068/FULL/PDF

Khan, S., & Hossain, M. K. (2022a). Classification and properties of nanoparticles. *Nanoparticle-Based Polymer Composites*, 15–54. https://doi.org/10.1016/B978-0-12-824272-8.00009-9

Khan, Y., Sadia, H., Ali Shah, S. Z., Khan, M. N., Shah, A. A., Ullah, N., Ullah, M. F., Bibi, H., Bafakeeh, O. T., Ben Khedher, N., Eldin, S. M., Fadhl, B. M., & Khan, M. I. (2022). Classification, Synthetic, and Characterization Approaches to Nanoparticles, and Their Applications in Various Fields of Nanotechnology: A Review. *Catalysts 2022, Vol. 12, Page 1386*, *12*(11), 1386. https://doi.org/10.3390/CATAL12111386

Knauth, P., & Schoonman, J. (2005). *Nanostructured materials: Selected synthesis methods, properties and applications*.

Kreyling, W. G., Semmler-Behnke, M., & Chaudhry, Q. (2010). A complementary definition of nanomaterial. *Nano Today*, *5*(3), 165–168. https://doi.org/10.1016/J.NANTOD.2010.03.004

Krishna Podagatlapalli, G. (2025a). The fundamentals of synthesis of the nanomaterials, properties, and emphasis on laser ablation in liquids: a brief review. *Discover Nano*, *20*(1). https://doi.org/10.1186/S11671-025-04235-5

Krishna Podagatlapalli, G. (2025b). The fundamentals of synthesis of the nanomaterials, properties, and emphasis on laser ablation in liquids: a brief review. In *Discover Nano* (Vol. 20, Issue 1). Springer. https://doi.org/10.1186/s11671-025-04235-5

Kumar, Narendra., & Kumbhat, Sunita. (2019). *Essentials in nanoscience and nanotechnology*.

Law, M., Greene, L. E., Johnson, J. C., Saykally, R., & Yang, P. (2005). Nanowire dye-sensitized solar cells. *Nature Materials 2005 4:6*, *4*(6), 455–459. https://doi.org/10.1038/nmat1387

Lisuzzo, L., Cavallaro, G., Lazzara, G., Milioto, S., Parisi, F., & Stetsyshyn, Y. (2018). Stability of halloysite, imogolite, and boron nitride nanotubes in solvent media. In *Applied Sciences (Switzerland)* (Vol. 8, Issue 7). MDPI AG. https://doi.org/10.3390/app8071068

Liu, Y., Huang, H., Cao, W., Mao, B., Liu, Y., & Kang, Z. (2020). Advances in carbon dots: from the perspective of traditional quantum dots. *Mater. Chem. Front*, *4*, 1586. https://doi.org/10.1039/d0qm00090f

Mekuye, B., & Abera, B. (2023). Nanomaterials: An overview of synthesis, classification, characterization, and applications. *Nano Select*, *4*(8), 486–501. https://doi.org/10.1002/nano.202300038

Moriarty, P. (2001). Nanostructured materials. *Reports on Progress in Physics*, *64*(3), 297. https://doi.org/10.1088/0034-4885/64/3/201

Mukhopadhyaya, J. (2017). *Synthesis and characterization of nanomaterials for uptake of radio nuclides and toxic metal ions*.

Murphy, C. J., & Jana, N. R. (n.d.). *Controlling the Aspect Ratio of Inorganic Nanorods and Nanowires*. https://doi.org/10.1002/1521-4095(20020104)14:1<80::AID-ADMA80>3.0.CO;2

Nasrollahzadeh, M., Sajadi, S. M., Sajjadi, M., & Issaabadi, Z. (2019). An Introduction to Nanotechnology. *Interface Science and Technology*, *28*, 1–27. https://doi.org/10.1016/B978-0-12-813586-0.00001-8

Pan, K., & Zhong, Q. (2025). Organic Nanoparticles in Foods: Fabrication, Characterization, and Utilization. *Annu. Rev. Food Sci. Technol*, *7*, 25. https://doi.org/10.1146/annurev-food-041715-033215

Paras, Yadav, K., Kumar, P., Teja, D. R., Chakraborty, S., Chakraborty, M., Mohapatra, S. S., Sahoo, A., Chou, M. M. C., Liang, C. Te, & Hang, D. R. (2023). A Review on Low-Dimensional Nanomaterials: Nanofabrication, Characterization and Applications. In *Nanomaterials* (Vol. 13, Issue 1). MDPI. https://doi.org/10.3390/nano13010160

Qi, H., & Zhang, C. (2022). Organic nanoparticles for electrogenerated chemiluminescence assay. In *Current Opinion in Electrochemistry* (Vol. 34). Elsevier B.V. https://doi.org/10.1016/j.coelec.2022.101023

Rao, N., Singh, R., & Bashambu, L. (2021). Carbon-based nanomaterials: Synthesis and prospective applications. *Materials Today: Proceedings*, *44*, 608–614. https://doi.org/10.1016/j.matpr.2020.10.593

Rauscher, H., Rasmussen, K., & Sokull-Klüttgen, B. (2017). Regulatory Aspects of Nanomaterials in the EU. *Chemie-Ingenieur-Technik*, *89*(3), 224–231. https://doi.org/10.1002/CITE.201600076;PAGEGROUP:STRING:PUBLICATION

Roduner, E. (2006). Size matters: Why nanomaterials are different. *Chemical Society Reviews*, *35*(7), 583–592. https://doi.org/10.1039/b502142c

Saleh, T. A. (2020). Nanomaterials: Classification, properties, and environmental toxicities. *Environmental Technology & Innovation*, *20*, 101067. https://doi.org/10.1016/J.ETI.2020.101067

Salem, S. S., & Fouda, A. (2011). Green Synthesis of Metallic Nanoparticles and Their Prospective Biotechnological Applications: an Overview. *Biological Trace Element Research*, *199*, 344–370. https://doi.org/10.1007/s12011-020-02138-3/Published

Schmid, Gunter. (20052004). *Nanoparticles : from theory to application*. Wiley-VCH Verlag.

Sharma, V., Kumar, A., & Dhawan, A. (2012). Nanomaterials: Exposure, Effects and Toxicity Assessment. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences 2012 82:1*, *82*(1), 3–11. https://doi.org/10.1007/S40011-012-0072-7

Shehzad, K., Xu, Y., Gao, C., & Duan, X. (2016). Three-dimensional macro-structures of two-dimensional nanomaterials. *Chemical Society Reviews*, *45*(20), 5541–5588. https://doi.org/10.1039/C6CS00218H

Shtykov, S. N. . (2018). *Nanoanalytics : nanoobjects and nanotechnologies in analytical chemistry*. 446.

Slezakova, K., Morais, S., Pereira, M. do C., Slezakova, K., Morais, S., & Pereira, M. do C. (2013). Atmospheric Nanoparticles and Their Impacts on Public Health. *Current Topics in Public Health*. https://doi.org/10.5772/54775

Soltani, A. M., & Pouypouy, H. (2019). Standardization and Regulations of Nanotechnology and Recent Government Policies Across the World on Nanomaterials. *Advances in Phytonanotechnology: From Synthesis to Application*, 419–446. https://doi.org/10.1016/B978-0-12-815322-2.00020-1

Tan, C., Cao, X., Wu, X. J., He, Q., Yang, J., Zhang, X., Chen, J., Zhao, W., Han, S., Nam, G. H., Sindoro, M., & Zhang, H. (2017). Recent Advances in Ultrathin Two-Dimensional Nanomaterials. *Chemical Reviews*, *117*(9), 6225–6331. https://doi.org/10.1021/ACS.CHEMREV.6B00558

Tiwari, J. N., Tiwari, R. N., & Kim, K. S. (2012). Zero-dimensional, one-dimensional, two-dimensional and three-dimensional nanostructured materials for advanced electrochemical energy devices. In *Progress in Materials Science* (Vol. 57, Issue 4, pp. 724–803). https://doi.org/10.1016/j.pmatsci.2011.08.003

Trotta, F., & Mele, A. (2019). *Nanomaterials: Classification and Properties*.

Verma, C., Berdimurodov, E., Verma, D. K., Berdimuradov, K., Alfantazi, A., & Hussain, C. M. (2023). 3D Nanomaterials: The future of industrial, biological, and environmental applications. *Inorganic Chemistry Communications*, *156*, 111163. https://doi.org/10.1016/J.INOCHE.2023.111163

Vi, S., & Tang, Y. (2014). *Synthesis of nanomaterials for biomedical applications*. http://eprints.nottingham.ac.uk/14101

Wang, J. L., & Meng, L. J. (2014). Influence of carbon nano-fiber on mechanical property of PALC. *Applied Mechanics and Materials*, *535*, 785–787. https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMM.535.785

Wang, X., Xu, C., Yi, M., & Zhang, H. (2011). Effects of Nano-ZrO2 on the microstructure and mechanical properties of Ti(C, N)-based cermet die materials. *Advanced Materials Research*, *154–155*, 1319–1323. https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMR.154-155.1319

Wang, Z., Hu, T., Liang, R., & Wei, M. (2020). Application of Zero-Dimensional Nanomaterials in Biosensing. In *Frontiers in Chemistry* (Vol. 8). Frontiers Media S.A. https://doi.org/10.3389/fchem.2020.00320

Warheit, D. B. (2018). Hazard and risk assessment strategies for nanoparticle exposures: How far have we come in the past 10 years? In *F1000Research* (Vol. 7). F1000 Research Ltd. https://doi.org/10.12688/f1000research.12691.1

Wu, Q., Miao, W. S., Zhang, Y. Du, Gao, H. J., & Hui, D. (2020). Mechanical properties of nanomaterials: A review. *Nanotechnology Reviews*, *9*(1), 259–273. https://doi.org/10.1515/ntrev-2020-0021

Zhang, H. (2015). Ultrathin Two-Dimensional Nanomaterials. *ACS Nano*, *9*(10), 9451–9469. https://doi.org/10.1021/acsnano.5b05040

Zhang, P., Tan, W., Zhang, X., Chen, J., Yuan, J., & Deng, J. (2021). Chemical Modification of Hydroxyl-Terminated Polybutadiene and Its Application in Composite Propellants. *Industrial & Engineering Chemistry Research*, *60*(10), 3819–3829. https://doi.org/10.1021/ACS.IECR.0C06172

Zidan, S., Silikas, N., Alhotan, A., Haider, J., & Yates, J. (2019). Investigating the Mechanical Properties of ZrO2-Impregnated PMMA Nanocomposite for Denture-Based Applications. *Materials 2019, Vol. 12, Page 1344*, *12*(8), 1344. https://doi.org/10.3390/MA12081344

Zou, B., Huang, C. Z., Wang, J., & Liu, B. Q. (2006). *Effect of Nano-Scale Ti N on the Mechanical Properties and Microstructure of Si3N4 Based Ceramic Tool Materials*. 154–158. https://documentsdelivered.com/source/046/866/046866302.php

Chapter 10

INTERVAL-VALUED DATA ANALYSIS: A SYSTEMATIC REVIEW OF METHODS, TRENDS, AND FUTURE DIRECTIONS

Sultan Turhan[1], Eralp Dogu[2]

1 Dr. AtaSancak Acıpayam Dairy Farm Denizli Türkiye, sultantrhn@hotmail.com, ahttps://orcid.org/0000-0002-9704-1700
2 Dr. Mugla Sitki Kocman University Department of Statistics Muğla Türkiye, eralp.dogu@mu.edu.tr, bhttps://orcid.org/0000-0002-8256-7304

1. Introduction

The exponential growth of information technology has fundamentally transformed data collection and storage practices across scientific and industrial domains. Traditional statistical methods, designed for single-valued observations, often prove inadequate when dealing with complex data structures that inherently contain variability and uncertainty. Interval-valued data represents one such structure where each observation is characterized by a range defined by lower and upper bounds rather than a single point value.

Interval-valued data naturally arises in numerous contexts: financial markets (daily trading ranges), environmental monitoring (temperature and humidity ranges), biomedical research (reference intervals for laboratory tests), and quality control (specification limits). The conventional approach of reducing such data to point estimates through averaging or other summary measures results in significant information loss and may lead to erroneous conclusions.

Symbolic Data Analysis (SDA), introduced by Diday (1988), provides a comprehensive framework for analyzing complex data structures including interval-valued data. This approach recognizes that statistical units may be characterized by more complex descriptions than simple numerical or categorical values, thereby preserving the inherent variability and structure within the data.

This systematic review addresses the following research questions:

1. What are the primary methodological approaches for interval-valued data analysis?
2. How has research in this field evolved over the past two decades?
3. Who are the key contributors and what are the main research clusters?
4. What gaps exist in current methodologies and what future directions are needed?

Understanding the current state of interval-valued data analysis is crucial for several reasons. First, it provides researchers with a comprehensive overview of available methodologies. Second, it identifies research gaps and opportunities for methodological development. Third, it offers guidance for practitioners seeking appropriate analytical techniques for their specific applications.

2. Methodology

2.1 Search Strategy

The search was conducted using the Web of Science Core Collection database, chosen for its comprehensive coverage of high-quality peer-reviewed literature. The following parameters are chosen for the filtering:

1. Search Terms: "interval valued data" OR "interval-valued data"
2. Database: Web of Science Core Collection
3. Time Period: 2000-2022
4. Language: English

5. Document Types: Articles, conference proceedings.

Studies focusing on statistical methods for interval-valued data, theoretical developments in symbolic data analysis, applications of interval-valued data analysis and methodological comparisons and evaluations were included in the study. Studies not primarily focused on interval-valued data, book chapters and editorial materials, duplicate publications, non-English publications were excluded.

2.3 Data Extraction and Analysis

The search yielded 163 articles meeting the inclusion criteria. For each article, we extracted:

- Bibliographic information (authors, year, journal, citations)
- Methodological approach
- Application domain
- Key contributions

Bibliometric analysis was performed using R software (version 4.2.0) and the bibliometrix package (Aria & Cuccurullo, 2017) to identify publication trends, author collaborations, and research clusters. Before presenting the results, we establish the theoretical foundation for interval-valued data analysis. The academic studies included in the study were collected at https://github.com/sultan-turhan/interval/blob/main/interval-valued-data.xls

2.4 Theoretical Framework: Interval-Valued Data

2.4.1 Definition and Notation

Symbolic data analysis (SDA) is an approach to statistical data analysis that focuses on the processing of complex and structured data. It is based on the idea that in addition to analyzing and drawing results at the individual level, it is generally necessary at the group level (Diday, 1988; Billard & Diday, 2000, Billard & Diday, 2003, Billard & Diday, 2006). Symbolic data can be described as categorical or weighting numbers. The first step in SDA is to create a symbolic data table, where rows are higher-level units and columns can take symbolic values. The second step is to examine this new type of data and extract new information using computational, statistical, and data mining extensions.

In general, symbolic data can contain intrinsic variability and be structured, unlike classical data, where each data point consists of a single (categorical or quantitative) value. It is the presence of this variation between data that necessitates the need for new techniques for analysis and will often differ from those for classical data. Note, however, that classical data represents special cases (for example, the classical point $x = a$ is equivalent to the symbolic range $\xi = [u, l]$). Then the symbolic data matrix $n * p$ for the object set $E = \{1, \dots n\}$ is represented by $X_j = (\xi_j)$ is the symbolic value observed for the variable $X_j, j = 1, \dots, p$ for the $u \in E$ object. The x_u of X is called the "symbolic upper bound " of the object u. The x_l of X is called the "symbolic lower bound " of the object u.

To illustrate an interval-valued symbolic data table, a dataset containing forensic data (Maile, 2017) could be a good example. Forensic science is an area with various applications (Turhan et al., 2022) that could be good test beds for interval data. The dataset includes postmortem interval (PMI) estimates from 9 death scenes in Lancaster County, Nebraska between March 2011 and March 2012. The postmortem interval is defined as the estimation of the time elapsed after death in cases where the exact moment of death is not known, and the detection of PMI is very important in terms of forensic medicine. This data set is given in Table 1.

Table 1. Sample Dataset: Symbolic Interval Forensic Data (Maile, 2017)

Case	Mass Loss (%)	Temperature (^{0}C)	Relative Humidity (%)	Estimated PMI (Days)
1	5-10	27	24.20	10-19
2	10-30	23	72.30	8-23
3	1-20	25	81.80	1-12
4	1-5	29	56.80	1-4
5	1-10	26	58.90	1-8
6	0.1-1	25	43.30	0-1
7	5-30	23	23.90	11-67
8	1-5	21	27.30	2-11
9	1-5	21	28.40	2-10

The data set consists of mass loss, temperature, relative humidity and estimated PMI variables. In this data set, mass loss and estimated PMI variables are interval variables.

So, for the data in Table 1, the first row is,

$$x_1 = \{[5-10], \quad 27^o, \quad 24.20, \quad [10-19]\}$$

It represents a case in mass loss 5%-10%, temperature 27^o degrees, relative humidity 24.20% and estimated PMI 10-19 days.

2.4.2 Basic Operations

Key operations for interval-valued data include for a certain j:

Center: $x_i^c = \frac{(x_i^u + x_i^l)}{2}$

Range: $x_i^r = x_i^u - x_i^l$, $i = 1,2,..,n$

3. Results

3.2 Methodological Landscape

The systematic review identified five primary methodological categories:

3.2.1 Regression Analysis (45 studies, 27.6%)

Regression analysis is the largest group of studies in this study. The notations will be borrowed from linear regressions since there are multiple types of regression (Billard & Diday, 2000). The general form of the classical linear regression model can be expressed in the matrix as:

$$Y = X\beta + \varepsilon$$

where the response vector $Y: n \times 1$ input variable $X: n \times p$. $\beta: p \times 1$ is the vector of unknown parameters, $\varepsilon: n \times 1$ is the random error vector. The general form of the interval valued linear regression model can be defined as follows (Türkşen & Metin, 2022):

$$[Y] = X[\beta] + [\varepsilon]$$

where [Y], [β] and [ε] are denoted as:

$$[Y] = \left[[Y_1^l, Y_1^u][Y_2^l, Y_2^u] \dots [Y_n^l, Y_n^u]\right]'$$

$$[\beta] = \left[[\beta_1^l, \beta_1^u][\beta_2^l, \beta_2^u] \dots [\beta_n^l, \beta_n^u]\right]'$$

$$[\varepsilon] = \left[[\varepsilon_1^l, \varepsilon_1^u][\varepsilon_2^l, \varepsilon_2^u] \dots [\varepsilon_n^l, \varepsilon_n^u]\right]'$$

Some methods used the center method, minmax method, center and range methods when creating the models.

Let $X_1, X_2, \dots, X_p$ is the exact input variables and $[Y_i] = [Y_i^l, Y_i^u]$, i=1,2,..,n is the interval valued response. Here, let Y_i^c be the center points of the interval-valued data obtained by $\mathrm{i} = 1,2,..,\mathrm{n}$

$$Y_i^c = \frac{(Y_i^l + Y_i^u)}{2}, \mathrm{i} = 1,2,..,\mathrm{n}$$

Then, the linear regression model for the center can be given in polynomial form as follows:

$$Y_i^c = \beta_0^c + \sum_{k=1}^{p} \beta_k^c X_{ik} + \sum_{k=1}^{p} \beta_{kk}^c X_{ik}^2 + \sum_{k=1}^{p} \sum_{m=1}^{p} \beta_{km}^c X_{ik} X_{im} + \varepsilon_i^c \quad \mathrm{i} = 1,2,..,\mathrm{n}$$

And, the model on the center point can be written as a classical model in matrix form as follows:

$$Y^c = X\beta^c + e^c$$

$Y^c = [Y_1^c\ Y_2^c \dots Y_n^c]'$, $\beta^c = [\beta_1^c\ \beta_2^c \dots \beta_p^c]'$ ve $\varepsilon^c = [\varepsilon_1^c\ \varepsilon_2^c \dots \varepsilon_n^c]'$. The parameter vector of the ordinary least squares (OLS) estimator $\hat{\beta}^c$ can be easily calculated using the following formula.

$$\hat{\beta}^c = (X'X)^{-1}X'Y^c$$

Billard and Diday (2002) proposed the MinMax method by taking into account the upper and lower values of interval value data. $X_1, X_2, \dots, X_p$ are the input variables and $[Y_i] = [Y_i^l, Y_i^u]$ i = 1,2,..,n range valued response.

$$Y^l = X\beta^l + e^l$$

$$Y^u = X\beta^u + e^u$$

$$\beta^l = [\beta_1^l\ \beta_2^l \dots \beta_p^l]'$$
$$\varepsilon^l = [\varepsilon_1^l\ \varepsilon_2^l \dots \varepsilon_n^l]'$$
$$\beta^u = [\beta_1^u\ \beta_2^u \dots \beta_p^u]'$$
$$\varepsilon^u = [\varepsilon_1^u\ \varepsilon_2^u \dots \varepsilon_n^u]'$$

$Y^u = [Y_1^u\ Y_2^u \dots Y_n^u]', \beta^l, \beta^u$

$$\hat{\beta}^l = (X'X)^{-1}X'Y^l$$
$$\hat{\beta}^u = (X'X)^{-1}X'Y^u$$

$$Y_i^l = \beta_0^l + \sum_{k=1}^{p} \beta_k^l X_{ik} + \sum_{k=1}^{p} \beta_{kk}^l X_{ik}^2 + \sum_{k=1}^{p}\sum_{m=1}^{p} \beta_{km}^l X_{ik}X_{im} + \varepsilon_i^l \quad \text{i} = 1,2,..,\text{n}$$

$$Y_i^u = \beta_0^u + \sum_{k=1}^{p} \beta_k^u X_{ik} + \sum_{k=1}^{p} \beta_{kk}^u X_{ik}^2 + \sum_{k=1}^{p}\sum_{m=1}^{p} \beta_{km}^u X_{ik}X_{im} + \varepsilon_i^u \quad \text{i} = 1,2,..,\text{n}$$

2.2.1.3. Center and Range Method (Türkşen & Metin, 2022)

The center and range method developed by Neto & De Carvalho (2008, 2018) consists of two independent linear regression models. This approach uses two key components: the interval centers and the interval ranges. The initial model is built using the center values and adheres to the same logic used in the center method. $X_1, X_2, \dots, X_p$ are the inputs, $[Y_i] = [Y_i^l, Y_i^u]$ i = 1,2,..,n is the response and i=1,2,..,n are the mid-points of the range-valued responses. The parameters for the center-based linear regression are estimated first. Then, the model corresponding to the interval-valued response is formulated as a polynomial:

$$Y_i^r = \beta_0^r + \sum_{k=1}^{p} \beta_k^r X_{ik} + \sum_{k=1}^{p} \beta_{kk}^r X_{ik}^2 + \sum_{k=1}^{p}\sum_{m=1}^{p} \beta_{km}^r X_{ik}X_{im} + \varepsilon_i^r \quad \text{i} = 1,2,..,\text{n}$$

Here, Y_i^r is the range of the interval-valued response and is obtained as follows:

$$[Y_i^r] = \left[Y_i^l, Y_i^u\right] \text{ i} = 1,2,..,\text{n}$$

can be written in matrix form as follows:

$$Y^r = X\beta^r + e^r$$

In which $Y^r = [Y_1^r\ Y_2^r \dots Y_n^r]', \varepsilon^r = [\varepsilon_1^r\ \varepsilon_2^r \dots \varepsilon_n^r]',\ \beta^r = [\beta_1^r\ \beta_2^r \dots \beta_p^r]',\ \text{i} = 1,2,..,\text{n}$ The ordinary least squared estimator is obtained as β_r. $\hat{\beta}^r = (X'X)^{-1}X'Y^r$

3.2.2 Fuzzy Logic Applications (26 studies, 16.0%)

Fuzzy logic is a mathematical theory that is based on thinking like a human and giving membership functions to certain outcomes so that the decision problem is not a 0-1 problem anymore where interval-value data analysis can be widely used (D'Urso et al., 2015). A fuzzy set is defined as $F: U \rightarrow [0,1]$. Contrary to definite sets, where an object may or may not be wholly belonging to a set, fuzzy sets allow the object $u \in U$ to partially belong to the F set at order $F(u)$. A fuzzy set X is a subset of a fuzzy set $\text{Y}, \text{X} \subseteq \text{Y}$ if $\text{X(u)} \leq \text{Y(u)}$, for all $\text{u} \in \text{U}$. Size of fuzzy set X is $|X| = \sum_{u \in U} X(u)$. Subsequently, we will work with closed intervals $[x_1, x_2] \subseteq [0,1]$. I denotes the set of all closed intervals above $[0,1]$ i.e., $I = \{[x_1, x_2] | [x_1, x_2] \subseteq [0,1]\}$. For $X = \{[x_1, x_2]\}$ the lower and upper bounds will be denoted by $\lfloor X \rfloor$ and $\lceil X \rceil$, i.e., $\lfloor X \rfloor = x_1$ and $\lceil X \rceil = x_2$ (Burda, 2020). Here, the interval valued data can be considered as a form of fuzzy numbers.

3.2.3 Clustering Methods (20 studies, 12.3%)

Clustering is a multivariate methodology that aims to organize similar entities into homogeneous classes based on the values observed in a set of variables (El-Sonbaty & Muhamed, 2002, Kejžar et al., 2021). Clustering for interval valued data, the aim is to classify the objects in Ω into $C_1, \dots, C_m$ clusters that are as homogeneous as possible internally and as different as possible externally. $D = (d_{ul})$ be the $n * n$ matrix of distance measures between objects $k, v \epsilon\ \Omega$. First, let us take quantitative-type variables and let us assume all Y_j are interval valued. Suppose we have the intervals $\xi_k = [X_{kl}, X_{ku}]$ and $\xi_v = [X_{vl}, X_{vu}],\ k, v \epsilon\ \Omega$. We seek a distance function $\delta_j\ (u, l)$ between objects k and v. There are a number of possible such functions used in the clustering process. We can define a symbolic Hausdorff distance for Y_j as

$$\delta_j\ (k, l) = max\ \{|X_{kl} - X_{vl}|, |X_{ku} - X_{vu}|\}$$

If we take a distance function $d(k, v)$

$$d(k,v) = \left(\sum_{j=1}^{p} [\delta_j (k,v)]^2 \right)^{1/2}$$

and use the specific $\delta_j(.)$ of we have

$$d(k,v) = \left(\sum_{j=1}^{p} [max \{|X_{ku} - X_{vu}|, |X_{kl} - X_{vl}|\}]^2 \right)^{1/2}$$

Notice that $d(k,v)$ in reduces to the Euclidean distance on $\Re^p$ when Y_j are classical variables.

3.2.4 Principal Component Analysis (18 studies, 11.0%)

Principal component analysis (PCA) is a data exploration method that aims to reduce the dimensions of a multivariate dataset by reconstructing the covariance matrix. In recent years, with the increase in knowledge reaching high dimensions and the difficulty of knowledge discovery, the effectiveness of PCA has been shown more and more. Therefore, some researchers have devoted themselves to presenting new PCA methods for complex situations such as symbolic data (Diday 1988; Bock & Diday, 2012; Diday & Noirhomme-Fraiture 2008). There are various versions of PCA for different symbolic data types in the literature. However, there are only two main differences for different PCA methods; one is how to construct a covariance matrix to determine the principal component space and the other is how to define a linear combination to create the projection of symbolic observations in the principal component space (Wang, 2016). The symbolic data table in this context contains $n \times p$ dimensions in its simplest form, which are explained by the p properties of the range type:

$$S^X = \begin{bmatrix} x_{11}^l, x_{11}^u & \cdots & x_{1p}^l, x_{1p}^u \\ \vdots & \ddots & \vdots \\ x_{n1}^l, x_{n1}^u & \cdots & x_{np}^l, x_{np}^u \end{bmatrix}$$

thus, each subject can be geometrically represented by a hyper-rectangle with 2^p corners in $\mathcal{R}^p$ space (Zuccolotto, 2007).

3.2.5 Time Series Analysis (10 studies, 6.1%)

Time series refers to the frequency of data points and is defined as measuring at regular time intervals, typically in successive time domains. When interval valued data is collected in chronological order, interval-valued data forms a time series. At each instant in time, $t = 1,2,\ldots,n$ is the number of

intervals observed in the time series, $X_{u_t}$and X_{l_t} as the upper and lower bounds of the interval $X_{l_t} \leq X_{u_t}$,

$$[X_{L_1}; X_{U_1}], [X_{l_2}; X_{u_2}], \dots, [X_{l_n}; X_{u_n}]$$

In general, in the methods presented here, two time series are considered: the X^c interval midpoint series and the X^r half interval interval series. For a time series, it can represent the midpoint and half-range interval series, respectively, as:

$$X_t^c = \frac{X_{u_t} + X_{l_t}}{2} \text{ and } X_t^r = \frac{X_{u_t} - X_{l_t}}{2} \quad t = 1,2, \dots, n$$

Thus, the values predicted by these models for the lower and upper bounds of the range $\hat{L}_{u_t}$) and $\hat{L}_{l_t}$, respectively, are given as:

$$\hat{L}_{U_t} = \hat{X}_t^c + \hat{X}_t^r \text{ ve } \hat{L}_{l_t} = \hat{X}_t^c - \hat{X}_t^r$$

$\hat{X}_t^c$ and $\hat{X}_t^c$ represent values predicted by linear adjustment for midpoint and half-pitch value series, respectively (Maia, 2008).

3.1 Publication Trends

The analysis of 163 articles reveals significant growth in interval-valued data research over the past two decades. Publication output shows an upward trend with notable increases after 2010, peaking at 20 publications in 2020. No publications were found for 2001, 2002, 2003, 2005, and 2009, indicating the field's relatively recent emergence. Figure 1 shows the number of papers per year.

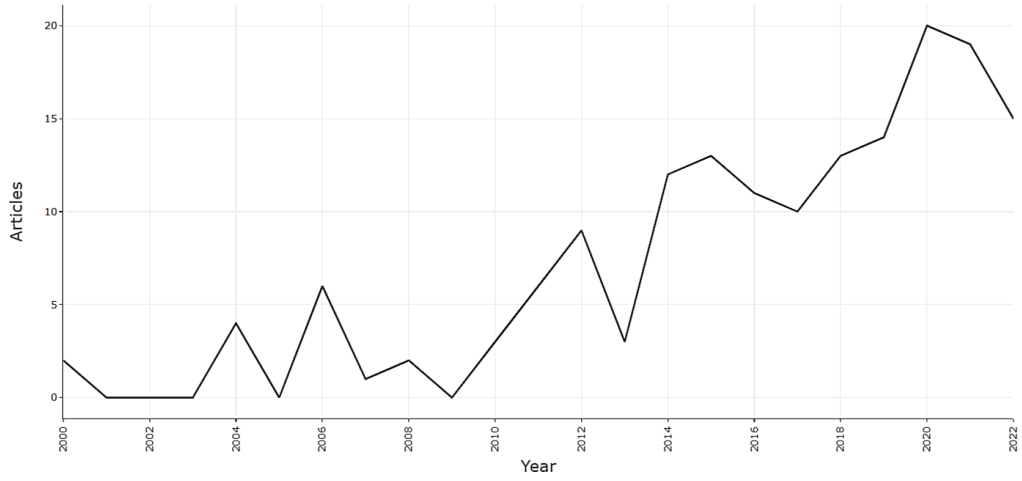


Figure 1. Annual Scientific Production for symbolic data analysis

The growth pattern suggests increasing recognition of interval-valued data analysis importance, likely driven by big data challenges and the need for uncertainty quantification in various applications. Leading Countries are listed as follows (Figure 2 shows the global distribution):

1. China: 153 publications (dominant contributor)
2. Brazil: 66 publications
3. USA: 62 publications
4. Other countries: <20 publications each

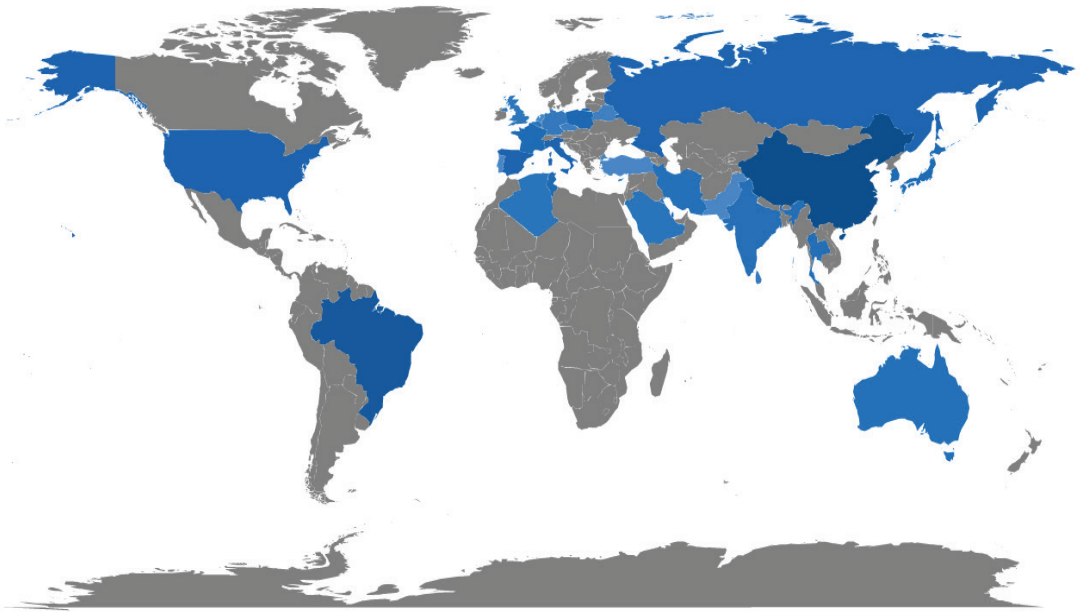


Figure 2. Countries' scientific production. Dark blue indicates the excess of studies, while light blue indicates the scarcity of studies. Gray shows the work in between.

Largest international collaboration are listed below (Figure 3 shows the international connections**):**

- Qatar-Tunisia: 6 joint publications
- China-USA: 4 joint publications
- Qatar-Kuwait: 4 joint publications

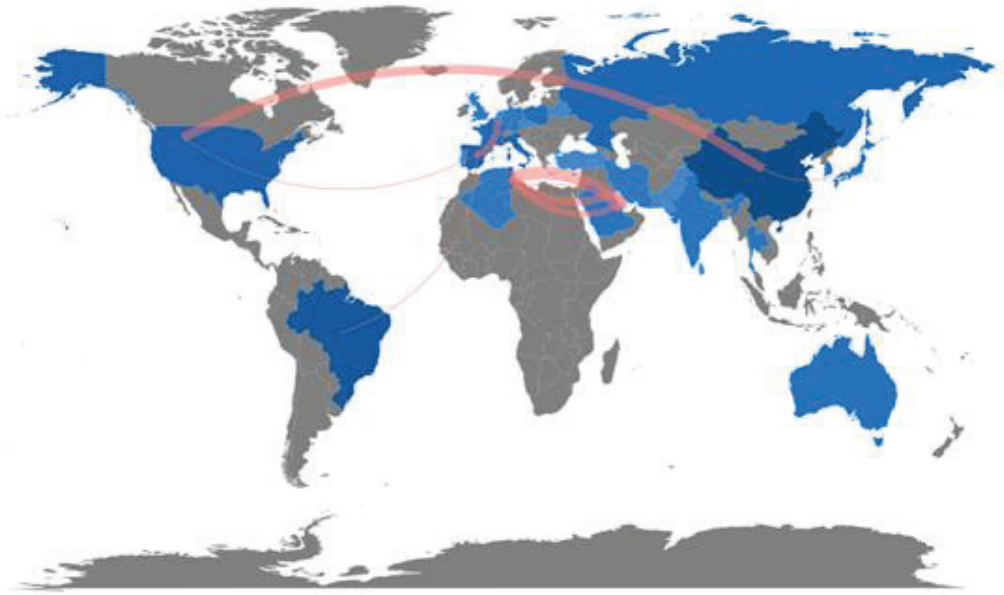


Figure 3. Collaboration network of the authors. Dark blue indicates the excess of studies, while light blue indicates the scarcity of studies. In addition, the thickness of the red line indicates the excess of relations between countries.

Figure 4 shows the production per top contributor over the years.

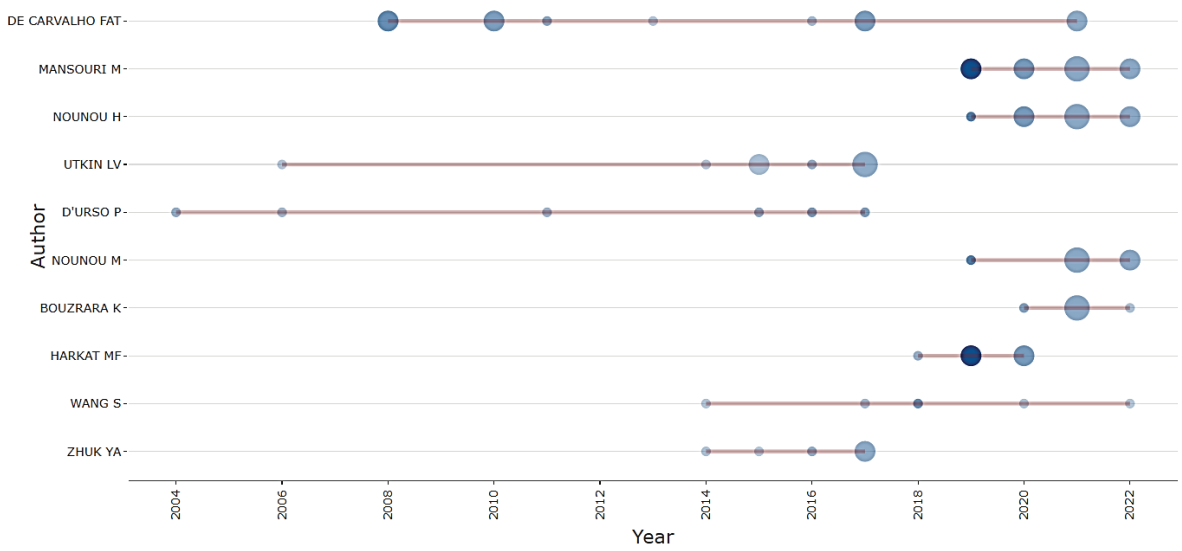


Figure 4. Authors' production over time. over time. The lines show the years between which the authors worked. If the bubbles are large and dark in color, the number of studies is high, and the small and light-colored ones indicate that the number of studies is low.

4. Conclusions

The systematic review reveals several notable strengths in the current state of interval-valued data analysis. First, the field demonstrates comprehensive coverage across multiple statistical domains,

with methods spanning regression, clustering, fuzzy logic, principal component analysis, and time series analysis. This indicates the versatility and broad applicability of interval-valued approaches. Second, researchers have established a strong theoretical foundation through rigorous mathematical framework development, particularly in symbolic data analysis theory. The field benefits from clear connections to real-world problems, with practical applications evident across diverse domains from finance to environmental monitoring. Finally, there is evidence of a growing research community, as demonstrated by the increasing number of researchers and publications over the past two decades.

Current applications of interval-valued data analysis are concentrated in several key application areas. Financial markets utilize these methods for analysing price ranges and volatility intervals, while environmental monitoring applications focus on measurement uncertainties and sensor data ranges. Quality control implementations emphasize specification limits and tolerance intervals, and biomedical research applications center on reference ranges and clinical decision intervals. However, several areas remain underexplored, presenting significant opportunities for expansion. Social sciences applications could benefit from interval-valued approaches for survey data and behavioural measurements. Network analysis with interval weights offers potential for modelling uncertain relationships, while spatial-temporal interval data analysis could advance geographic information systems and urban planning applications. Multi-modal interval distributions represent another frontier for handling complex uncertainty patterns.

The field requires substantial advancement in several methodological areas. Advanced machine learning methods represent a critical priority, including the development of ensemble methods specifically designed for interval-valued data, deep learning architectures adapted for interval inputs and outputs, reinforcement learning frameworks with interval rewards, and transfer learning techniques between different interval domains. Equally important is the establishment of a comprehensive statistical inference framework encompassing robust hypothesis testing procedures, bootstrap methods adapted for interval data, Bayesian approaches to interval uncertainty quantification, and model selection and validation techniques that account for interval structure. Big data challenges demand immediate attention through the development of distributed computing algorithms for interval data, streaming algorithms capable of processing interval-valued data in real-time, dimensionality reduction techniques for high-dimensional interval datasets, and privacy-preserving analysis methods that protect sensitive interval information.

Current software support for interval-valued data analysis remains limited, creating a significant barrier to widespread adoption. Priority areas for development include comprehensive R packages that provide integrated tools for interval analysis across all major methodological approaches, Python

libraries with seamless machine learning integration for modern data science workflows, specialized visualization tools that effectively communicate interval uncertainty and relationships, and cloud-based platforms designed for large-scale interval data analysis. These software developments are essential for translating theoretical advances into practical tools that researchers and practitioners can readily adopt across diverse application domains.

This systematic review has several important limitations that should be acknowledged. The reliance on a single database search (Web of Science only) may have resulted in the omission of relevant publications indexed in other databases such as Scopus, IEEE Xplore, or ACM Digital Library, potentially creating coverage gaps in the literature analysis. The English language restriction excludes potentially valuable contributions published in other languages, particularly from non-English speaking countries that may have significant research activity in this field. The use of limited search terms ("interval+valued+data") may have missed relevant studies using alternative terminology such as "range data," "bounded data," or "uncertain data," thereby underrepresenting the true scope of research in this area. Additionally, the focus on quantitative rather than qualitative analysis limits the depth of understanding regarding methodological nuances, research contexts, and theoretical developments that could provide richer insights into the field's evolution. Future reviews should address these limitations through comprehensive multi-database searches encompassing major academic databases, broader search strategies incorporating synonymous terms and related concepts, inclusion of grey literature such as conference proceedings and technical reports, and integration of qualitative analysis methods to provide deeper contextual understanding of the research landscape and methodological developments in interval-valued data analysis.

Data Availability Statement

The bibliographic data used in this study is publicly available through Web of Science. The complete dataset of analyzed articles is available at: https://github.com/sultan-turhan/interval/blob/main/interval-valued-data.xls

References

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975.

Billard, L., & Diday, E. (2000). Regression analysis for interval-valued data. In Data analysis, classification, and related methods (pp. 369-374). Berlin, Heidelberg: Springer Berlin Heidelberg.

Billard, L., & Diday, E. (2002). Symbolic regression analysis. In Classification, clustering, and data analysis: recent advances and applications (pp. 281-288). Berlin, Heidelberg: Springer Berlin Heidelberg.

Billard, L., & Diday, E. (2003). Symbolic data analysis: Definitions and examples. Dept. Statist., Univ. Georgia, Athens, GA, USA, Tech. Rep.

Billard, L., Diday, E. (2006). Descriptive statistics for interval-valued observations in the presence of rules. *Computational Statistics*, 21(2), 187-210.

Burda, M., Pavliska, V., & Murinová, P. (2020). Support and confidence calculation for interval-valued fuzzy association rules. *International Journal of Computational Intelligence Systems*, 13(1), 1014-1026.

Diday, E. (1988). The symbolic approach in clustering and related methods of data analysis: the basic choices. In Classification and Related Methods of Data Analysis, Proceedings of the First Conference of the International Federation of Classification Societies (IFCS-87: Technical University of Aachen (pp. 673-684). North Holland.

Diday, E., & Noirhomme-Fraiture, M. (Eds.). (2008). *Symbolic data analysis and the SODAS software*. John Wiley & Sons.

D'Urso, P., Giordani, P., & Leski, J. M. (2015). Fuzzy clustering of interval-valued data. *Pattern Recognition*, 48(3), 956-968.

El-Sonbaty, Y., & Ismail, M. A. (2002). Fuzzy clustering for symbolic data. *IEEE Transactions on fuzzy system*s, 6(2), 195-204.

Maciel, L., & Ballini, R. (2019). Interval fuzzy S-curve membership functions for time series forecasting. *Information Sciences*, 501, 69-90.

Neto, E. A. L., & de Carvalho, F. A. (2008). Centre and range method for fitting a linear regression model to symbolic interval data. *Computational Statistics & Data Analysis*, 52(3), 1500-1515.

Neto, E. A. L., & de Carvalho, F. A. (2018). An exponential-type kernel robust regression model for interval-valued variables. *Information Sciences*, 454, 419-442.

Wang, H., Chen, M., Shi, X., & Li, N. (2016). Principal component analysis for normal-distribution-valued symbolic data. *IEEE Transactions on Cybernetics*, 46(2), 356-365.

Kejžar, N., Korenjak-Černe, S., & Batagelj, V. (2021). Clustering of modal-valued symbolic data. *Advances in Data Analysis and Classification*, *15*(2), 513-541.

Maile, A. E., Inoue, C. G., Barksdale, L. E., & Carter, D. O. (2017). Toward a universal equation to estimate postmortem interval. *Forensic science international*, *272*, 150-153.

Bock, H. H., & Diday, E. (Eds.). (2012). *Analysis of symbolic data: exploratory methods for extracting statistical information from complex data*. Springer Science & Business Media.

Zuccolotto, P. (2007). Principal components of sample estimates: an approach through symbolic data analysis. *Statistical Methods and Applications*, *16*(2), 173-192.

Turhan, S., Tunç, M., Doğu, E., & Balcı, Y. (2022). Machine learning in forensic science and forensic medicine: Research on the literature. *Adli Tıp Dergisi*, *36*(1), 1-7.

Türkşen, Ö., & Ulu Metin, G. (2022). An adapted linear modeling method for interval-valued responses: golden center and range method. *Communications in Statistics: Case Studies, Data Analysis and Applications*, *8*(3), 463-483.

Maia, A. L. S., de Carvalho, F. D. A., & Ludermir, T. B. (2008). Forecasting models for interval-valued time series. *Neurocomputing*, *71*(16-18), 3344-3352.

Chapter 11

THE EVOLUTİON OF LİTHOGRAPHY TECHNİQUES İN MİCROCHİP MANUFACTURİNG

MERVE EREN YAKIŞIKLIER[1]

1 Assistant Professor, Yozgat Bozok University, Department of Medical Services and Techniques,

Orcid: 0000-0002-6920-359X

The Fundamental Role of Lithography in Microchip Manufacturing

The development of microchip manufacturing technologies is an engineering history shaped by the evolution of lithography processes. Goals such as miniaturizing semiconductor devices, increasing performance, and improving energy efficiency have necessitated the continuous expansion of the resolution capacity of manufacturing lines. The exponential scaling predicted by Moore's Law has led to a redefinition of lithography techniques every decade, resulting in a multilayer technological transformation ranging from optical lithography to deep ultraviolet (DUV) lithography, and then to extreme ultraviolet (EUV) lithography. While lithography is fundamentally based on the process of transferring patterns onto a semiconductor surface using light, it is considered the most critical step in microchip manufacturing due to its technological complexity and precise dimensional control. Resolution is limited by factors such as the wavelength of the light source, the numerical aperture, and the physical properties of the photoresist material used. Therefore, each new generation of lithography technology has been made possible not only by innovations in the optical system but also by advances in materials science, plasma engineering, and metrology. Currently, the high-resolution advantages offered by EUV technology are utilized in 5 nm and 3 nm process nodes, and the industry is now discussing more advanced methods such as High-NA EUV. This section examines in detail the historical development of lithography in microchip manufacturing, technological milestones, and the advanced methods achieved today.

The Birth of Optical Lithography and First-Generation Microchip Manufacturing Techniques

The foundations of the lithography process in microchip manufacturing were laid in the late 1950s and early 1960s, when semiconductor technology first entered commercial production. The lithography method used during this period was largely adapted from photographic techniques and was based on the principle of coating a light-sensitive photoresist onto a silicon surface and transferring the pattern to the substrate with the help of light passing through a mask.

Mercury vapor lamps were used as light sources in the first optical lithography systems; particularly g-line (436 nm) lithography, with resolution levels in the order of 10 μm, constituted the first generation process for microchip manufacturing (Mack, 2007).

During these years, lithography was considered a relatively simple pattern transfer technique rather than a critical stage in integrated circuit manufacturing. By the 1970s, the level of integration in the microelectronics industry began to increase, and higher-resolution lithography methods were needed to reduce component sizes. Thus, i-line (365 nm) lithography became widespread and, combined with advances in lens optics, made it possible to reduce resolution to levels of 1–2 μm (Burns & Colburn, 1989).

One of the most prominent features of optical lithography during this period was the simultaneous development of the chemical structure of photoresist materials with optical systems and the optimization of the process largely through empirical methods. Parameters such as exposure dose, light homogeneity, mask quality, and photoresist thickness became factors directly determining instrument efficiency; this increased the strategic importance of lithography in the production chain (Thompson et al., 1994).

The early 1980s were a period when lithography underwent a completely new technical transformation due to the downscaling pressure brought about by Moore's Law. During this period, it was clearly demonstrated through the Rayleigh criterion that the resolution limits are directly related to the wavelength of light; Therefore, it was accepted that further reduction of wavelength was necessary for the advancement of optical lithography (Ito & Willson, 1983). Deep UV (DUV) lithography studies were initiated to reduce semiconductor geometries below the 1 μm limit; these studies demonstrated that optical lithography is not only a production technique but also a fundamental determinant of technological competition in the semiconductor field. For this reason, the early days of optical lithography are considered a historical preparatory process that necessitated the emergence of subsequent generations of high-resolution lithography techniques.

The Development of Deep UV (DUV) Lithography

With optical lithography approaching its resolution limits, Deep Ultraviolet (DUV) lithography, using shorter wavelengths since the early 1980s, has become the strongest candidate to meet the scaling-down goals of the semiconductor industry. The fundamental motivation for DUV lithography is that the pattern transfer limit is directly related to the wavelength, as predicted by the Rayleigh resolution criterion. Therefore, after reaching the industrial limits of 365 nm i-line lithography, a new era of lithography began with the use of 248 nm (KrF) and later 193 nm (ArF) excimer laser light sources (Wong, 2001). The ability of excimer lasers to produce monochromatic and high-intensity light created a huge leap in resolution; thus, pattern transfer with high accuracy at sub-micron scales became possible. One of the most critical stages in the development process of DUV lithography is the application of the chemical amplification principle, which requires a complete redesign of photoresist materials. Developed by Ito and Willson in the late 1980s, this approach relied on accelerating the post-exposure reaction through acid production, thus achieving high sensitivity despite the low energy of DUV light (Ito & Willson, 1982; Willson & Ito, 1988). Chemically amplified photoresists (CARs) paved the way for the commercial production of DUV lithography and enabled 0.35 μm–0.25 μm geometries to become industry standards. This development demonstrated that lithography must advance not only with optical systems but also with materials science. By the mid-1990s, increasingly aggressive resolution targets began to push the boundaries of DUV lithography. Therefore, immersion lithography was developed to increase the resolution of 193 nm ArF lithography; The effective digital aperture of the system is increased by placing a high refractive index liquid (usually pure water) between the lens and the resistor (Levenson, 2006). The immersion approach has enabled scaling up to 45 nm and then 32 nm process nodes, exceeding the theoretical limits of 193 nm lithography, and has become the main production method of the industry since the early 2000s. However, immersion lithography has also brought new engineering problems such as bubble formation, surface tension effects, and resistor-liquid interactions. The development of DUV lithography, while enabling increased resolution, has also necessitated complex process steps such as double/quadruple patterning. By approaching the physical limits of 193 nm optics, the process has been implemented with a combination of multiple exposures and etching steps for line widths that cannot be obtained in a single exposure. This situation led to a significant increase in production costs, but it also played a bridging technology role by enabling the transition to the 20–10 nm scale in industry (Mulkens et al., 2015).

In this context, DUV lithography is considered a critical turning point in the evolution of optical lithography. DUV represents a period in which scientific and engineering disciplines became more

integrated in semiconductor manufacturing, both with the emergence of excimer laser technology and the development of chemically amplified photoresists. At the same time, DUV lithography created the necessary technological infrastructure for the birth of EUV lithography and became one of the fundamental steps that enabled industry to descend to nanometer scales.

The Emergence of Extreme Ultraviolet (EUV) Lithography

As the physical limits of deep UV lithography were approached, the semiconductor industry faced the need to further reduce the wavelength to increase resolution. While supporting 193 nm immersion lithography with multiple patterning techniques enabled production in processes of 22 nm and below, the cost and process complexity increased significantly. Therefore, from the second half of the 1990s onwards, Extreme Ultraviolet (EUV) lithography with a wavelength of 13.5 nm emerged as the next generation production technology (Attwood & Sakdinawat, 2017). The main advantage of EUV is that resolutions of 20 nm and below can be achieved in a single exposure thanks to the significantly shortened wavelength. In this respect, EUV represents the most radical transformation in the optical lithography roadmap that has been ongoing in the industry for decades. The biggest challenge in the development of EUV lithography was the inability to produce suitable light sources operating at this wavelength. Because EUV light is strongly absorbed by the atmosphere, the entire system must operate under a high vacuum; also, since conventional lenses are not transparent at the EUV wavelength, reflective multilayer mirrors are used. These mirrors are generally composed of molybdenum/silicon (Mo/Si) multilayer structures and direct EUV photons using the Bragg reflection principle (Bajt et al., 2002). Since the reflectivity of each mirror used in the optical system is only around 60–70%, a large portion of the EUV photons emitted from the light source are lost before reaching the system; therefore, developing a high-brightness EUV light source has become a critical engineering requirement. Currently, the most common light source used in EUV lithography systems is laser-produced tin plasma (LPP) technology. In this process, microscopic tin (Sn) droplets are bombarded with high-energy CO_2 lasers to produce EUV radiation at a wavelength of 13.5 nm (Bakshi, 2018). However, the risk of damage to optical components from particles scattered by tin plasma has posed a significant obstacle to the stability of EUV systems for many years. Therefore, sophisticated filter systems have been developed to protect the mirrors, and the cleaning of optical surfaces has been achieved through methods such as atomic hydrogen resurfacing (Morita et al., 2010). Another major obstacle to the production of EUV lithography is the complexity of mask structures. Unlike classic DUV masks, EUV masks are reflective and have a multilayered structure. The fact that defects on the mask can be magnified and lead to larger defects on the wafer has significantly increased the importance of mask metrology. Therefore, both defect-free multilayer production and pellicle development in EUV masks have been the subject of research for many years (Buitrago et al., 2016).

By 2019, manufacturers such as ASML, Intel, and Samsung had commissioned the first commercial systems enabling the use of EUV pellicle; thus, a significant milestone was reached in the protection of masks. Despite all the technical challenges, EUV lithography was successfully incorporated into commercial production by Samsung and TSMC in 2019 with a 7 nm process node. The adoption of EUV significantly reduced the need for multiple patterning and lowered the overall cost of the production chain. Furthermore, EUV lithography has become the standard technology in critical layers in 5 nm and 3 nm processes, leading to the development of High-NA EUV systems that will enable the industry to transition to 2 nm and smaller geometries in the future (Lio et al., 2021). The emergence of EUV lithography represents not only an increase in resolution achieved through a

reduction in wavelength, but also a technological revolution resulting from the convergence of a multidisciplinary scientific endeavor—materials science, plasma physics, optical engineering, and metrology.

The Evolution of Multiple Patterning Techniques: Double, Triple, and Quadruple Patterning Approaches

As the physical resolution limits of 193 nm lithography approached, the industry realized that it could not rely solely on optical system improvements to achieve sub-40 nm geometries; this necessitated the development of multiple patterning techniques. Producing narrow line spacing that cannot be achieved with a single exposure through the sequential application of multiple exposure, etching, and/or deposition steps has become a standard method, especially at 45 nm, 32 nm, and 22 nm process nodes (Mulkens et al., 2015). Multiple patterning is considered a critical bridging technology of the pre-EUV era, as well as extending the lifespan of DUV lithography. The most fundamental approach within multiple patterning techniques is the double patterning process called Double Patterning Lithography (DPL). This method is divided into two main categories: LELE (Litho–Etch–Litho–Etch) and Spacer-based DPL. In the LELE approach, the wafer is subjected to two consecutive exposures and two etching processes; thus, much narrower line spacing can be obtained than a single exposure can provide. However, the main disadvantage of this method is that mask alignment errors (overlay errors) grow in the second exposure, negatively affecting critical size stability (Vaidyanathan et al., 2006). Therefore, although technologically feasible, LELE is considered a method that increases production costs and variation extremes. Another important method is Self-Aligned Double Patterning (SADP). In SADP, a mandrel line is created in the first step; then a spacer (sidewall cladding) is grown on the sidewalls of this line. After the mandrel is removed, only the spacer structures remain, and these structures form the basis of the final pattern. The self-aligning feature of SADP has found widespread use in 20 nm class production because it largely eliminates overlay errors (Gu et al., 2012). Furthermore, the SADP method is considered more reliable than LELE in terms of critical endpoint stability. Self-Aligned Quadruple Patterning (SAQP), a more advanced version of SADP, makes it possible to quadruple the resolution using two consecutive spacer cycles. SAQP has become a necessary extension of DUV lithography for manufacturers targeting 10–7 nm geometries. Major manufacturers such as TSMC and Intel adopted the SAQP method as a core technique for high-performance logic circuits during the 2015–2018 period, when EUV was still in its maturation phase (Choi et al., 2015). The biggest advantage of SAQP is that it offers much higher dimensional stability; however, due to manufacturing challenges such as process complexity, material interactions, and thin-film strain, SAQP has only been economically feasible in high-tech nodes. In the evolution of multiple patterning techniques, intermediate methods such as Self-Aligned Triple Patterning (SATP) and LELE–EC (LELE with etch compensation), known as a hybrid approach, have also been developed. Although these methods have been used to fill technological gaps, they have found limited long-term application due to the complexity of the process flow, the increasing number of masks, and rising production costs (Ban & Zhang, 2014). While multi-patterning methods are generally effective for overcoming the resolution limitations of DUV lithography, they have not been considered a sustainable long-term solution due to factors such as increased cost, cycle time, and production variation. In this context, the multi-patterning approach can be considered a strategic part of the transition process to EUV lithography. Although the need for multi-patterning has largely decreased with the commercial use of EUV, methods such as SAQP are still used in some layers due

to EUV's low photon flux, mask defects, and pellicle limitations. It is predicted that the use of hybrid multi-patterning methods in conjunction with EUV will continue, especially in 3 nm and 2 nm process nodes (Liu et al., 2020). Therefore, multi-patterning techniques hold an important place in both the past and future of semiconductor manufacturing.

Metrology and Error Correction Processes After Lithography

The lithography process works in an integrated manner with a series of sub-processes that require high precision for successful pattern transfer. Among these processes, metrology is a fundamental component of process control in semiconductor manufacturing, enabling the measurement of parameters such as line width, overlay accuracy, critical dimension (CD) stability, and interlayer alignment. With resolutions reaching nanometer scales, metrology after lithography has become essential not only for ensuring accuracy but also for modeling process variations and guaranteeing device performance (Bunday, 2018). Therefore, metrology is of equivalent importance to lithography in modern microchip manufacturing. One of the most common measurement techniques used after lithography is CD-SEM (Critical Dimension-Scanning Electron Microscopy), which allows for direct visualization of nanoscopic line widths. CD-SEM systems provide critical information about line edge roughness (LER) and surface morphology at sub-nm dimensions thanks to the high resolution capacity of the electron beam. However, disadvantages such as the potential for photoresist degradation by high-energy electron bombardment and charge accumulation on the surface during measurement necessitate careful parameter control in the in-process application of CD-SEM (Leunissen & Timans, 2002). Nevertheless, CD-SEM has been accepted as the industry standard in all modern process nodes, especially starting from 90 nm. As an alternative to CD-SEM, optical scatterometry has emerged as a non-destructive and faster measurement method. Scatterometry indirectly calculates parameters such as line width, line height, and profile shape by analyzing the spectral distribution of reflected or diffracted light (Pang et al., 2016). This method is widely used in 300 mm wafer production lines, particularly in high-volume production, as it offers a speed advantage. Furthermore, scatterometry has begun to play a critical role in EUV lithography processes for examining mask defects and topographic variations on multilayer structures. Another critical element of lithography is overlay metrology. Overlay is the alignment of patterns in different layers relative to each other, and even a small alignment error at the nanometer scale can lead to significant degradation in instrument performance. As resolution decreases, overlay tolerances become increasingly tighter; therefore, optical overlay measurement systems have been developed to achieve sub-nanometer accuracy, supported by advanced image processing algorithms (Hsu et al., 2013). Minimizing overlay errors is critical, especially in processes where multiple patterning techniques are applied, because small alignment errors occurring in the second or third exposure can cause serious shifts in the final pattern. In post-lithography processes, the error correction stage is usually performed using model-based methods such as OPC (Optical Proximity Correction) and RET (Resolution Enhancement Techniques). OPC (Optical Pattern Correction) allows for the pre-correction of patterns on the mask, taking into account optical distortions that may occur during production. This ensures that the final shape on the wafer is closer to the desired geometry. RET (Reverse Optical Transformation) techniques include methods such as phase-shift masks (PSM) and off-axis illumination, which improve the optical transfer function of the lithography system and increase resolution (Saleh & Teich, 2019). With EUV lithography, OPC models have become more complex due to the much more pronounced mask defects, making the integration of physics-based simulations (FEM, waveguide modeling) a necessity. Today, advanced metrology systems not only perform measurements but are also integrated into closed-loop control mechanisms that enable

lithography tools to self-optimize by predicting process variations with AI-powered algorithms. These features play a critical role in increasing efficiency and reducing cost per chip in high-volume production. In conclusion, post-lithographic metrology, regularly applied in both DUV and EUV lithography processes, is one of the fundamental pillars of production reliability and its importance is expected to increase further as geometries become smaller.

Next-Generation Lithography Approaches

The continuously increasing resolution requirements in microchip manufacturing have necessitated the development of new lithography methods beyond traditional optical lithography. In this context, innovative techniques such as electron beam lithography, nanoimprint lithography, directed self-assembly (DSA), and high-NA EUV have emerged as significant alternatives for the future in terms of both resolution and production efficiency. Each technique offers different balances of resolution, speed, cost, and manufacturability, finding various applications in the semiconductor industry.

Electron Beam Lithography (E-Beam Lithography)

Electron beam lithography is based on the principle of directly writing high-energy electrons onto a resist material and is known for exceeding the diffraction limits of optical lithography, achieving resolutions at the 1–5 nm level (Chen, 2015). Because E-beam lithography is a maskless method, it is preferred in research laboratories and low-volume production applications. However, since the printing speed depends on scanning the electron dots individually, it is not economical for high-volume industrial production. The most critical role of e-beam lithography in industry is its use in printing EUV and DUV masks; it is an indispensable technology for high-resolution mask production in these fields (Hibbs et al., 2018). The main limitations of e-beam include electron scattering in resistors, unwanted energy emission known as the proximity effect, and the fact that this affects pattern accuracy. Therefore, in recent years, intensive research has been carried out on proximity correction algorithms and high-precision resistor materials.

Nanoimprint Lithography (NIL)

Nanoimprint lithography is a method that enables pattern transfer by mechanically imprinting high-resolution molds onto polymer resin. Because NIL is not subject to optical limitations, it allows pattern production at sub-10 nm resolution (Chou, Krauss & Renstrom, 1996). The most important advantage of NIL is its low cost and the fact that it provides high resolution without the need for complex optical systems. However, there are several challenges in the industrial applications of NIL. For example, mold-fast wear, particle entrapment, mold release stress, and uniformity problems at large wafer scales limit production yield (Resnick et al., 2005). Nevertheless, NIL is increasingly being used in some specialized applications such as memory devices, biosensors, and photonic structures.

Directed Self-Assembly (DSA) Lithography

The Directed Self-Assembly (DSA) method, which takes advantage of the self-assembly behavior of polymer block copolymers, stands out as a novel approach that makes it possible to obtain sub-20 nm patterns at low cost. The basic principle of DSA is to orient polymer chains into thermodynamically stable structures and combine them with lithography to obtain smaller patterns (Tang et al., 2008).

DSA is based on two main methods:

1. Graphoepitaxy – orientation via wall structures

2. Chemoepitaxy – orientation via chemical surface patterns

The biggest advantage of DSA is that it reduces the need for multiple patterning and can dramatically lower costs. However, controlling line defects, the limited line length, and managing production variations make full-scale use of DSA technology in industry difficult (Delaney & Fredrickson, 2017). Nevertheless, organizations such as Intel, IBM, and imec consider DSA a key component in their hybrid lithography strategies for nodes beyond 7 nm and 5 nm.

High-NA EUV Lithograph

The success of EUV lithography has been taken to a new level with the development of optical systems with higher numerical aperture (NA) to further reduce resolution limits. While the NA value of standard EUV systems is 0.33, this value has been increased to 0.55 in High-NA EUV systems. The increase in NA directly improves resolution according to the Rayleigh criterion and makes it possible to produce <20 nm pitch ranges in a single exposure (Lio et al., 2021).

The key advantages of High-NA EUV are:

• Further reduction in the need for multiple patterning

• Smaller-scale error tolerances in EUV masks

• Potential to reduce critical layer costs

However, this technology presents significant engineering challenges due to larger mirrors, more complex optical designs, much higher mask accuracy, and limited depth field (Banine et al., 2019). ASML's planned delivery of the first High-NA EUV scanners (EXE:5000 series) to research facilities by 2024 is considered a significant milestone in the commercialization process of the technology.

Future Perspectives and Industry Roadmap

The future of lithography in microchip manufacturing is being shaped by technological advancements, industry roadmaps, and scientific meetings that redefine the scalability limits of the semiconductor industry. In this context, reports published by the International Electron Devices Meeting (IEDM) and the International Roadmap for Devices and Systems (IRDS) offer important insights into which lithography strategies will stand out for manufacturing nodes of 1 nm and below. Both sources emphasize that a multi-step, hybrid technology path is inevitable to further increase resolution, reduce lithography costs, and improve process stability.

In recent years, IEDM presentations have clearly demonstrated that High-NA EUV lithography will be the most critical manufacturing tool for process nodes of 2 nm and below. Increasing the NA value to 0.55 allows for a reduction in line spacing by approximately 40% compared to the Rayleigh resolution limit, making it possible to transfer much finer details in a single exposure. At the IEDM 2022 and 2023 sessions, it was stated that High-NA EUV systems could go down to pitches below 16 nm, especially in logic circuits; thus significantly reducing the need for multiple patterning. However, it was noted that challenges such as increasing optical complexity, the need for higher accuracy in masks, and the limited depth of field of the system still need to be resolved. The IRDS 2023–2032 roadmap offers a longer-term perspective on the future of lithography technology. According to the report, it is predicted that after 2030, lithography performance in 1 nm and below process nodes will

be shaped not only by optical systems but also by innovations in materials science. Two areas in particular are highlighted:

Next-generation resistor chemistry:

Development of high-sensitivity molecular resistors with low LER (Line Edge Roughness) is essential to reduce the stochastic effects created by the low energy of EUV photons. Metal oxide resistors and non-CAR (non-chemically amplified) structures are being considered as potential solutions for sub-1 nm pattern stability.

Advanced masking technologies:

Larger masks required for High-NA EUV necessitate the investigation of holographic phase correction techniques, multilayer reflection optimization, and maskless lithography approaches due to angular differences in the optical path.

IRDS reports emphasize that after 2035, resolution increase may not be sustainable with EUV lithography alone; instead, hybrid lithography methods will gain importance. The main components of these hybrid systems are:

High-NA EUV + DSA (Directed Self-Assembly)

Integration of directed self-assembly techniques with EUV is envisioned to enable the low-cost production of sub-10 nm patterns.

Maskless Lithography (ML2 / Multi-beam E-beam)

If multi-beam printers make significant progress in speed, it is considered that maskless lithography may make a comeback in custom design layers.

Nanoimprint (NIL) Hybrid Processes

As we approach sub-1 nm physical limits, NIL processes, which are not dependent on optical limits, are expected to play a complementary role in certain areas.

With the approach of the physical limits of sub-1 nm processes, it does not seem possible for the technology to advance only with optical or electronic processes. Therefore, IEDM and IRDS state that a new lithography paradigm based on atomic manufacturing techniques may emerge in the long term. In this context, intensive research continues in the following areas:

• Atomic Layer Patterning (ALP)

• Quantum-scale self-organizing material systems

• Integration of 2D materials into the lithography process

• Laser-assisted atomic printing techniques

If these technologies mature, a completely new era in the history of lithography is expected to begin.

Generally speaking, the future of lithography points to a multidisciplinary transformation. While the industry's short-term strategy is based on High-NA EUV, the long-term roadmap is moving towards a new architecture based on a combination of hybrid lithography, self-organization, maskless printing, and atomically controlled manufacturing processes. Thus, it is predicted that the production of sub-1

nm geometries will be technically possible after 2030, but this will require a much more complex engineering effort in terms of cost, material stability, and physical limitations (Banine et al., 2019).

Conclusion and Evaluation

The evolution of lithography techniques in microchip manufacturing has been one of the most critical technological developments shaping the semiconductor industry over the last sixty years. The journey of optical lithography, starting from the 10 μm scale, has accelerated with chemically amplified DUV processes and multiple patterning techniques; and finally, it has reached a new dimension with the industrial maturity of EUV lithography operating at a wavelength of 13.5 nm. Each generation of lithography has not only increased resolution but has also necessitated multidisciplinary scientific progress in fields such as photoresist chemistry, optical engineering, metrology, and process integration. Therefore, the development of lithography technology plays a directly decisive role in the performance, power consumption, and integration density of semiconductor devices.

Although EUV lithography has become a standard manufacturing method in critical layers of advanced process nodes such as 5 nm and 3 nm in today's technology, the increase in resolution requirements brings new challenges. Low photon flux, mask defects, pellicle limitations, and optical losses remain fundamental elements determining the performance of EUV systems. In this context, High-NA EUV is positioned as a leading candidate to shape the future of the industry, with the potential to improve resolution and reduce the need for multiple patterning. On the other hand, alternative methods such as electron beam lithography, nanoimprint lithography, and oriented self-arranging have begun to be integrated into the production process not only in the research field but also in specific niche applications. These techniques herald a hybrid lithography future that enables multi-technology strategies not entirely dependent on EUV. Post-lithography metrology and error correction processes are also becoming increasingly important. Parameters such as overlay accuracy, critical dimensional stability, and edge roughness at sub-nanometer scales are key factors limiting lithography performance. Therefore, semiconductor manufacturing requires a holistic production approach based not only on the pattern transfer process but also on high-precision measurement and model-based error correction systems. In this context, metrology will play a decisive role in the lithography technologies of the future. In conclusion, the evolution of lithography technology is a story that constantly redefines the scalability limits of the microchip industry. Progressing at the pace predicted by Moore's Law has depended, and continues to depend, on the successful development of each new generation of lithography. The introduction of High-NA EUV, combined with advances in materials science and AI-powered process control mechanisms, will ensure that semiconductor manufacturing continues to advance towards geometries of 1 nm and below. Therefore, lithography will remain a fundamental driving force behind the computing technologies of the future, at the heart of innovations that push the boundaries of microchip architectures.

REFERENCES

Attwood, D., & Sakdinawat, A. (2017). *X-rays and extreme ultraviolet radiation: Principles and applications* (2nd ed.). Cambridge University Press.

Bajt, S., Alameda, J. B., Barbee, T. W., et al. (2002). Improved reflectance and stability of Mo/Si multilayers. *Optics Letters, 27*(19), 1616–1618.

Bakshi, V. (2018). *EUV lithography* (2nd ed.). SPIE Press.

Ban, H., & Zhang, G. (2014). Triple patterning lithography: A promising solution for 14 nm technology node and beyond. *Microelectronic Engineering, 119*, 12–18.

Banine, V., Koshelev, K., & Swinkels, G. (2019). Physical processes in EUV sources for high-volume manufacturing. *Journal of Applied Physics, 125*(23), 231101.

Buitrago, E., Wurm, S., & Benk, S. (2016). EUV mask challenges for high volume manufacturing. *Microelectronic Engineering, 155*, 85–92.

Bunday, B. (2018). Metrology for the semiconductor industry at the nanoscale: A review. *Journal of Micro/Nanolithography, MEMS, and MOEMS, 17*(4), 041010.

Burns, S. D., & Colburn, M. (1989). Optical lithography: Thirty years of innovation. *Solid State Technology, 32*(7), 75–82.

Chen, Y. (2015). Nanofabrication by electron beam lithography and its applications: A review. *Microelectronic Engineering, 135*, 57–72.

Choi, W., Lin, B., & Wise, A. (2015). Self-aligned quadruple patterning for advanced logic devices. *Journal of Micro/Nanolithography, MEMS, and MOEMS, 14*(3), 033505.

Chou, S. Y., Krauss, P. R., & Renstrom, P. J. (1996). Nanoimprint lithography. *Journal of Vacuum Science & Technology B, 14*(6), 4129–4133.

Delaney, K. T., & Fredrickson, G. H. (2017). Theory of block copolymer directed self-assembly. *Annual Review of Physical Chemistry, 68*(1), 321–346.

Gu, J., Smith, B., & Yan, P. (2012). Spacer-based double patterning: Process integration and challenges. *Proceedings of SPIE, 8326*, 83260E.

Hibbs, A., Lee, C., & Watanabe, K. (2018). High-resolution mask writing using e-beam lithography for advanced semiconductor nodes. *SPIE Photomask Technology, 10810*, 1081010.

Hsu, S., Holst, D., & Stowers, J. (2013). Overlay metrology and control for advanced semiconductor manufacturing. *Proceedings of SPIE, 8681*, 86810F.

Ito, H., & Willson, C. G. (1982). Chemical amplification in resist materials for microlithography. *Journal of Polymer Science: Polymer Chemistry Edition, 20*(4), 871–892.

Ito, H., & Willson, C. G. (1983). Chemical amplification in the design of dry developing resist materials. *Polymer Engineering & Science, 23*(18), 1012–1018.

Leunissen, L., & Timans, P. (2002). Critical dimension metrology for advanced semiconductor devices. *Microelectronic Engineering, 61–62*, 455–466.

Levenson, M. D. (2006). ArF immersion lithography: The epitome of optical lithography. *Journal of Micro/Nanolithography, MEMS, and MOEMS, 5*(2), 021001.

Lio, A., Malloy, M., Stoeldraijer, J., & Banine, V. (2021). High-NA EUV lithography: Enabling the next generation of semiconductor devices. *Journal of Micro/Nanolithography, MEMS, and MOEMS, 20*(4), 041013.

Liu, Y., Naulleau, P., & Goldberg, K. (2020). Patterning strategies for sub-3-nm semiconductor manufacturing. *Applied Physics Reviews, 7*(4), 041402.

Mack, C. A. (2007). *Fundamental principles of optical lithography: The science of microfabrication*. John Wiley & Sons.

Morita, T., Tanaka, T., & Miyake, Y. (2010). EUV optics contamination control using atomic hydrogen cleaning. *Surface Science, 604*(9–10), 879–885.

Mulkens, J., van de Mast, F., van de Kerkhof, M., & Streefkerk, B. (2015). Multiple patterning: From 193 nm lithography to EUV. *Proceedings of SPIE, 9422*, 94220A.

Pang, L., Xu, Y., & Zhang, Z. (2016). Scatterometry metrology for advanced lithography: Principles and applications. *Applied Optics, 55*(5), 1123–1132.

Resnick, D. J., Vladimirescu, M., Lin, Y., et al. (2005). Applications of nanoimprint lithography for semiconductor device fabrication. *Solid State Technology, 48*(7), 39–45.

Saleh, B. E. A., & Teich, M. C. (2019). *Fundamentals of photonics* (3rd ed.). Wiley.

Tang, C., Lennon, E. M., Fredrickson, G. H., et al. (2008). Directed self-assembly of block copolymers for nanolithography: Enabling technologies and challenges. *Macromolecules, 41*(18), 6926–6939.*

Thompson, L. F., Willson, C. G., & Bowden, M. J. (1994). *Introduction to microlithography* (2nd ed.). ACS Publications.

Vaidyanathan, R., Corliss, D., & Chen, R. (2006). Litho–etch–litho–etch double patterning for advanced logic. *Proceedings of SPIE, 6520*, 65200F.

Wong, R. H. (2001). *Excimer laser lithography*. SPIE Press.

Chapter 12

NAPHTHOQUINONE DERIVATIVES: SYNTHESIS, STRUCTURAL FEATURES, AND MULTIFACETED APPLICATIONS

Şenol YAVUZ[1]

1 Asist. Prof. Dr. Şenol YAVUZ, Hitit University, Osmancık Ömer Derindere Vocational School, Department of Property Protection and Security, ORCİD: 0000-0001-6261-9296

1. INTRODUCTION

Naphthoquinones are aromatic quinone compounds widely distributed in nature and play a key role in fundamental redox processes in many biological systems (Elhabiri et al., 2014; Futuro et al., 2018). Chemically, these compounds possess carbonyl groups at the 1,4-positions of the naphthalene ring and are remarkable molecules in both biochemical and industrial fields due to their strong electron-accepting properties (Elhabiri et al., 2014). The naphthoquinone scaffold constitutes the core structure of secondary metabolites naturally synthesized in plants, bacteria, fungi, lichens, and certain marine organisms (Futuro et al., 2018).

The biological significance of these compounds arises from their ability to participate in redox cycling, thereby mediating cellular oxidation–reduction reactions (Elhabiri et al., 2014). In particular, ubiquinones (coenzyme Q) and vitamin K derivatives, which function in the mitochondrial electron transport chain, are among the best-known biological examples of the naphthoquinone family (Futuro et al., 2018). In addition, natural naphthoquinones such as lapachol, plumbagin, juglone, and lawsone have attracted considerable attention in pharmaceutical research due to their antimicrobial, antifungal, anticancer, and antiparasitic activities (Futuro et al., 2018).

The biological activities of naphthoquinones are primarily associated with their redox properties. These compounds can accept one or two electrons in the cell to form semiquinone radicals, and their interaction with oxygen generates reactive oxygen species (ROS) (Elhabiri et al., 2014). Such ROS can induce oxidative damage to biomolecules, including DNA, lipids, and proteins, thereby causing cytotoxicity (Futuro et al., 2018). Therefore, in addition to their therapeutic potential, the toxicological aspects of naphthoquinones are also of significant importance. In particular, the interaction of compounds such as menadione (vitamin K_3) and plumbagin with biological targets through redox cycling has been extensively investigated from a drug design perspective (Elhabiri et al., 2014).

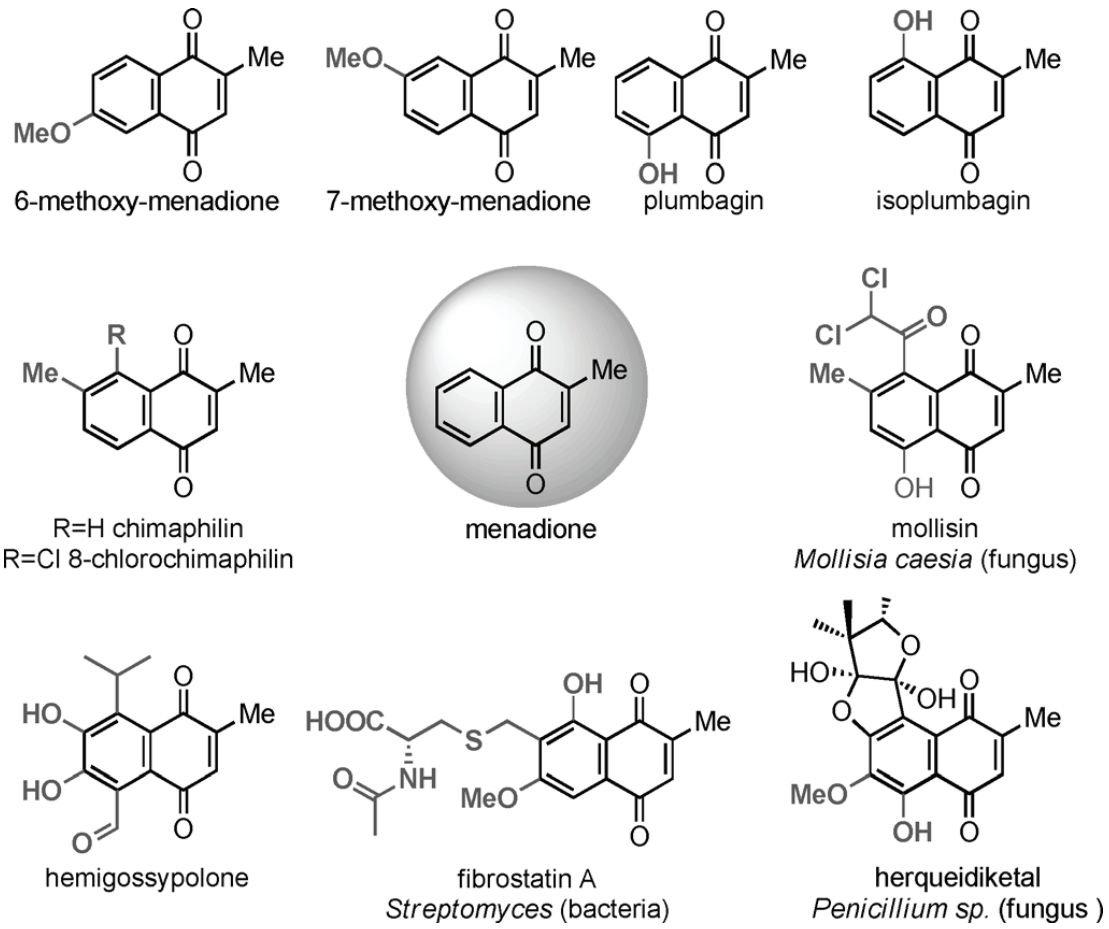


Figure 1. Natural menadione derivatives, including menadione, polysubstituted on the aromatic ring (Elhabiri et al., 2014).

Recent studies have demonstrated that introducing various functional groups onto the 1,4-naphthoquinone scaffold results in significant changes in the chemical and biological properties of the resulting derivatives (Sayıl & Ibis, 2010; Bayrak, 2019; Ibis & Deniz, 2010; Patan, Göksel, & Ayla, 2021). The incorporation of electron-donating or electron-withdrawing groups directly influences the redox potential, lipophilicity, and biocompatibility of these compounds (Elhabiri et al., 2014). For example, substitutions with groups such as $–NH_2$, –SH, $–OCH_3$, and $–CF_3$ markedly alter both the spectroscopic characteristics and pharmacological activities of the molecules (Bayrak, 2017; Ibis, Sahinler Ayla, & Gulsah Deniz, 2010; Ibis & Sahinler Ayla, 2011). Consequently, synthetic efforts over the past decade have focused on the development of N-, S-, O-, and N, S-substituted derivatives by attaching various nucleophiles (amines, thiols, phenols, etc.) to the 1,4-naphthoquinone ring (Aysecik Kacmaz et al., 2019; Tuyun et al., 2015, Patan, Göksel, & Ayla, 2021).

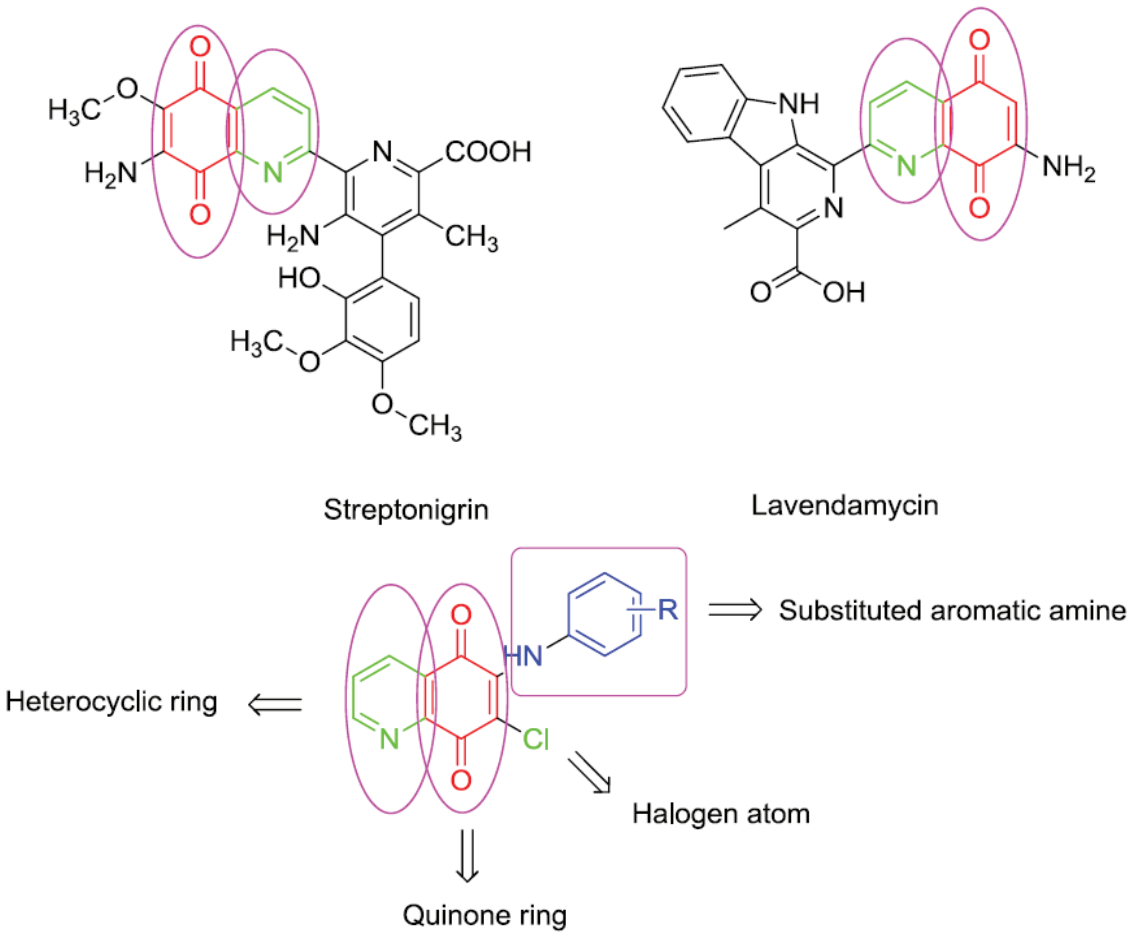


Figure 2. Azanaphthoquinone derivatives exhibiting antimicrobial activity (Bayrak, 2019).

The chemical versatility of naphthoquinones extends beyond biological systems. The electrochemical properties of these compounds have rendered them valuable for applications such as energy storage, sensor technologies, and organic electronics (Elhabiri et al., 2014). Owing to their electron-transfer capabilities, naphthoquinone derivatives have been employed in a variety of redox-active materials and photovoltaic systems (Ibis & Deniz, 2010). In addition, certain derivatives have been explored for use in textile dyes and photochromic materials due to their high stability and color-imparting properties (Sayıl, Deniz, & Çınarlı, 2016; Ibis & Deniz, 2010).

2. SYNTHETIC METHODS OF NAPHTHOQUINONES

Naphthoquinones are chemically conjugated systems with strong electron-withdrawing character and can therefore react with a wide variety of nucleophiles to form numerous derivatives. Synthetic approaches can generally be classified into three main categories: isolation from natural products, oxidation of precursor aromatic compounds, and nucleophilic substitution-based derivatization strategies (Patan, Göksel, & Ayla, 2021; Futuro et al., 2018). Today, N-, S-, O-, and N, S-substituted naphthoquinones obtained by combining these methods have become essential intermediates for applications in medicinal chemistry, materials science,

and biotechnology (Aysecik Kacmaz et al., 2019; Sayıl & Ibis, 2010; Bayrak, 2019; Ibis & Sahinler Ayla, 2011; Yavuz, 2023; Ibis, Sahinler Ayla, & Yavuz, 2019).

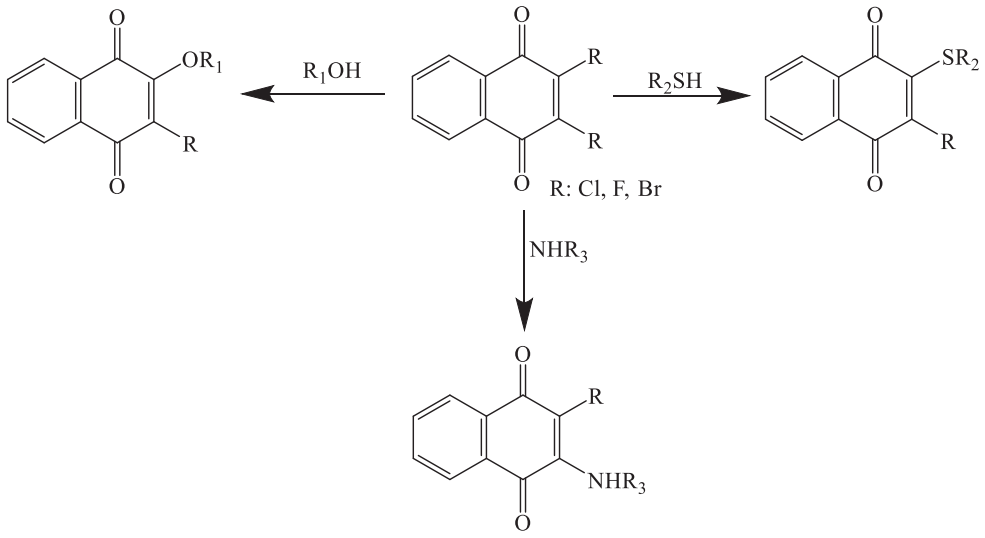


Figure 3. Schematic representation of the synthesis of N-, S-, O-, and N, S-substituted naphthoquinones (Yavuz, 2023; Ibis, Sahinler Ayla, & Yavuz, 2019).

2.1. Derivatives Obtained from Halogeno-Naphthoquinones

One of the most frequently employed strategies for the preparation of 1,4-naphthoquinone derivatives involves the reaction of halogenated derivatives, such as 2,3-dichloro-1,4-naphthoquinone (DCNQ), with suitable nucleophiles (Patan, Göksel, & Ayla, 2021). The interaction of DCNQ with nucleophiles such as piperidine, arylamines, or thiols affords N-substituted or N, S-substituted products in high yields. In the study by Patan and co-workers, the reaction of 2,3-dichloro-1,4-naphthoquinone with piperidinemethanol resulted in an N-substituted 1,4-naphthoquinone derivative, which was subsequently treated with various aromatic and aliphatic thiols to synthesize N, S-substituted derivatives (Patan, Göksel, & Ayla, 2021). FT-IR, ^{1}H NMR, ^{13}C NMR, and mass spectrometry analyses of the products showed complete agreement with the proposed structures.

Similarly, in a series of studies conducted by İbiş and co-workers, precursors such as 2-bromo-3-hydroxy-1,4-naphthoquinone and 2,3-dichloro-1,4-naphthoquinone were reacted with amine, thioamide, and phenol derivatives to synthesize novel S-, O-, and N-substituted naphthoquinones (Sayil & Ibis, 2010; Bayrak, 2017; Sayil & Ibis, 2010). In these reactions, nucleophilic substitution mechanisms carried out under basic conditions (Na_2CO_3, K_2CO_3, pyridine), typically in ethanol or acetone as the solvent, were found to be effective. Aromatic thiols, amines, and phenols act as nucleophilic centers, forming substitution at the halogenated carbon.

An essential advantage of these methods is the controllability of reaction selectivity. Parameters such as nucleophile strength, solvent polarity, and temperature play a decisive role in

determining product distribution. For instance, substitutions carried out with aromatic thiols generally afford high yields, whereas reactions involving aliphatic amines proceed more slowly (Bayrak, 2019; Patan, Göksel, & Ayla, 2021).

2.2. Nucleophilic Addition and Condensation Reactions

The naphthoquinone core exhibits electrophilic character due to its conjugated carbonyl system and is therefore well suited for the synthesis of new derivatives via Michael addition reactions or condensation mechanisms (Kaçmaz et al., 2019; Tuyun et al., 2015; Bayrak, 2017). In these approaches, amines, hydrazines, thiols, or carbonyl compounds are commonly employed as nucleophiles.

In a study by Nilüfer and co-workers, trifluoromethyl-containing arylamines were introduced onto the 1,4-naphthoquinone ring to obtain trifluoromethyl–arylamino–naphthoquinone derivatives (Yıldırım et al., 2017). Both spectroscopic and computational analyses (UV–Vis, NMR, DFT) were used to investigate the electron density distribution over the conjugated system, revealing a narrowing of the HOMO–LUMO gap due to the strong electron-withdrawing effect of the $–CF_3$ group.

In another approach, S, S-disubstituted or S, O-mixed-substituted 1,4-naphthoquinones were synthesized using thiosemicarbazone and sulfenyl derivatives (Tuyun et al., 2015; Becerra et al., 2021). Such compounds are typically formed via condensation or dual nucleophilic addition mechanisms. In methods developed by İbiş and co-workers, the sequential reaction of two different nucleophiles (for example, an amine followed by a thiol) enabled the construction of multifunctional structures (Bayrak, 2017;Ibis & Sahinler Ayla, 2011).

2.3. Oxidative Methods and Derivatization from Natural Products

The oxidation of hydroxynaphthalene derivatives obtained from natural sources represents an important alternative route for the synthesis of naphthoquinones (Futuro et al., 2018). For example, natural compounds such as lawsone (2-hydroxy-1,4-naphthoquinone) and juglone (5-hydroxy-1,4-naphthoquinone) can be rearranged and derivatized with new functional groups under appropriate oxidation conditions (Futuro et al., 2018). These approaches are particularly preferred for the structural modification of biologically active natural quinones.

In a comprehensive review by Futuro and co-workers (Futuro et al., 2018), the modification of naturally sourced naphthoquinones regarding their antimicrobial and antifungal activities was discussed. The study reported that condensation of lawsone derivatives with various amines yielded products with enhanced biological activity. A significant advantage of these methods is that they can be carried out under more environmentally friendly ("green") conditions, using solvents such as water or ethanol, and without forming toxic by-products.

2.4. Catalytic and Green Chemistry Approaches

In recent years, the use of transition-metal catalysts such as $CeCl_3$, CuI, and Pd/C in the synthesis of naphthoquinone derivatives has increased reaction rates and selectivity (Bayrak, 2019). In particular, $CeCl_3$-catalyzed nucleophilic substitution reactions offer high yields and short reaction times. Moreover, microwave-assisted syntheses and solvent-free reaction conditions have emerged as prominent strategies for environmental sustainability. These catalytic methods also facilitate the formation of structurally controlled derivatives. For example, in microwave-assisted reactions of aminonaphthoquinones, product selectivity is significantly increased while side-product formation is reduced (Tuyun et al., 2015). Consequently, both energy consumption and reaction times are shortened to minutes.

2.5. Overview of Reaction Mechanisms

The principal mechanisms involved in the synthesis of naphthoquinones are nucleophilic aromatic substitution (SNAr) and Michael-type addition reactions. The electron-withdrawing carbonyl groups render the carbons on the aromatic ring electrophilic; therefore, nucleophilic attack readily occurs at the position bearing the halogen atom (Patan, Göksel, & Ayla, 2021). The σ-complex formed as an intermediate subsequently regains aromaticity by eliminating a proton or a halogen. This mechanism is particularly efficient for 2,3-dichloro-1,4-naphthoquinone derivatives, affording high reaction yields.

Density Functional Theory (DFT)-based calculations have revealed the influence of substituent type on the HOMO–LUMO energy gap and electron density distribution (Ibis & Deniz, 2010; Elhabiri et al., 2014). In compounds bearing electron-withdrawing groups, a decrease in LUMO energy has been observed, leading to enhanced reactivity of electrophilic centers. These findings are consistent with experimental data and serve as a guide for the rational design of substitution strategies in naphthoquinone synthesis.

3. STRUCTURAL CHARACTERIZATION METHODS

Spectroscopic and analytical techniques used to confirm the chemical structures of products obtained during the synthesis of naphthoquinone derivatives and to determine the positions of functional groups are critical for understanding the physical, chemical, and biological properties of these compounds. In this section, the roles of FT-IR, UV–Vis, NMR, mass spectrometry, X-ray diffraction (XRD), and electrochemical analyses in the characterization of naphthoquinones are discussed in detail (Sayil, Deniz, & Cınarlı, 2016; Tuyun et al., 2015; Ibis & Deniz, 2010; Elhabiri et al., 2014; Patan, Göksel, & Ayla, 2021).

3.1. FT-IR Spectroscopy

FT-IR analysis is widely used, particularly for the identification of carbonyl (C=O) and hydroxyl (O–H) functional groups. In naphthoquinones, the characteristic carbonyl stretching

vibrations of the 1,4-dione system are typically observed in the range of 1660–1685 cm^{-1} (Patan, Göksel & Ayla, 2021). For example, in compounds obtained by derivatization of 2,3-dichloro-1,4-naphthoquinone with piperidine or piperonylamine, a slight shift of the carbonyl stretching frequencies toward lower wavenumbers (1670–1655 cm^{-1}) has been observed following N-substitution (Patan, Göksel & Ayla, 2021). This shift is attributed to an increase in electron density, which weakens the carbonyl bond.

Similarly, in O- or S-substituted derivatives, bands corresponding to C–O–C and C–S stretching vibrations are observed in the ranges of 1200–1300 cm^{-1} and 680–740 cm^{-1}, respectively (Bayrak, 2017; Becerra et al., 2021). Broad absorption bands associated with hydroxyl-containing derivatives generally appear in the region of 3300–3400 cm^{-1} (Tuyun et al., 2015). These characteristic bands directly confirm the presence of functional groups and, together with other spectroscopic data, play a complementary role in structural elucidation.

3.2. UV–Vis Spectroscopy

Naphthoquinones exhibit characteristic absorption bands due to their conjugated π-electron systems. Typically, π→π* transitions are observed in the range of 220–280 nm, while n→π* transitions appear between 310–410 nm (Patan, Göksel & Ayla, 2021). The wavelengths and intensities of these transitions vary depending on the electronic effects of substituents on the molecule.

In the study by Patan and co-workers, solvent-dependent shifts were observed in the UV–Vis spectra of N- and N, S-substituted derivatives; while the absorption maximum was around 350 nm in ethanol, a bathochromic shift to approximately 370 nm was observed in chloroform (Patan, Göksel & Ayla, 2021). This behavior indicates that solvent polarity and hydrogen-bonding interactions influence the energy of the π→π* transitions.

In spectroscopic studies conducted by İbiş and colleagues (Ibis & Deniz, 2010), it was reported that the color properties of S- and O-substituted naphthoquinone dyes exhibited a red shift upon the presence of electron-donating groups. Electron donor substituents increase the electron density on the aromatic ring, thereby lowering the transition energy and shifting the absorption maximum to longer wavelengths. This feature provides a significant advantage in the design of naphthoquinone derivatives for dye and photochromic material applications

3.3. NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is an indispensable technique for structural confirmation of compounds through both proton (^{1}H) and carbon (^{13}C) signals. In naphthoquinone derivatives, aromatic protons in the ^{1}H NMR spectrum generally resonate in the range of 7.2–8.2 ppm, while protons located near the carbonyl groups typically appear around 6.5–7.0 ppm (Sayil & Ibis, 2010; Bayrak, 2017; Patan, Göksel & Ayla, 2021).

In N-substituted compounds, signals corresponding to aliphatic groups attached to the nitrogen atom are observed in the 2.5–4.0 ppm range. For example, in the compound obtained from the reaction of 2,3-dichloro-1,4-naphthoquinone with 3-piperidinemethanol, the proton signal of the $–CH_2OH$ group was detected at 3.74 ppm [28]. Similarly, in S-substituted derivatives, $–SCH_2$ protons generally resonate between 3.3 and 3.6 ppm (Bayrak, 2017; Ibis & Deniz, 2010).

In the ^{13}C NMR spectra, carbonyl carbons are typically observed in the range of 180–185 ppm, aromatic carbons between 120–150 ppm, and aliphatic carbons between 20–60 ppm (Tuyun et al., 2015; Bayrak, 2017). ^{13}C NMR analyses are particularly decisive in determining whether the 1,4-quinone structure is preserved, owing to the presence of characteristic dual carbonyl signals.

3.4. Mass Spectrometry (MS)

Analysis of naphthoquinone derivatives by mass spectrometry (ESI-MS, EI-MS) is essential for determining molecular ion peaks and verifying structures. In the study by Patan and colleagues, the synthesized N, S-substituted compounds were observed in positive ion mode as sodium or potassium adducts, $[M+Na]^+$ or $[M+K]^+$ (Patan, Göksel & Ayla, 2021). This indicates that the ionization tendency of the compounds depends on the nature of the substituents.

In the reports by Tuyun et al. (Tuyun et al., 2015), amino-naphthoquinone derivatives were ionized in negative-ion mode as $[M–H]^–$, and their molecular weights were entirely consistent with the synthetic formulas. Additionally, in some cases, characteristic fragmentation patterns (e.g., loss of $–CH_2$ or –SH groups) were observed, which were used for structural confirmation.

3.5. X-Ray Diffraction (XRD) and Molecular Structure

The crystal structures of some naphthoquinone derivatives have been determined using single-crystal X-ray diffraction (SC-XRD) (Becerra et al., 2021; Ibis & Sahinler Ayla, 2011). The obtained crystallographic data provided insights into intramolecular hydrogen bonding and π–π stacking interactions. For instance, in an S, O-disubstituted derivative, intramolecular interactions between the carbonyl oxygens and sulfur atoms were observed and were found to enhance crystal stability (Becerra et al., 2021).

Crystal structure analyses also provide quantitative information on the planarity of the aromatic rings and the extent of conjugation in the system. These data are beneficial for comparing the electronic properties and optical behaviors of the compounds with theoretical calculations (Ibis & Deniz, 2010; Elhabiri et al., 2014).

3.6. Electrochemical Characterization

The electrochemical properties of naphthoquinone derivatives are crucial for examining their redox behavior. Cyclic voltammetry (CV) analyses have shown that 1,4-naphthoquinone

derivatives typically undergo two successive reduction steps (Elhabiri et al., 2014). The redox potentials of menadione (2-methyl-1,4-naphthoquinone) and its derivatives vary depending on whether the substituents are electron-withdrawing or electron-donating.

In the electrochemical studies by Elhabiri and colleagues (Elhabiri et al., 2014), it was reported that electron-withdrawing groups (e.g., $–NO_2$, $–CF_3$) increase the redox potential, whereas electron-donating groups (e.g., $–OCH_3$, $–NH_2$) decrease it. This indicates that substituents modulate redox behavior by altering the HOMO–LUMO energy levels. These experimental findings were also validated using theoretical QSPR (quantitative structure–property relationship) models (Elhabiri et al., 2014).

Electrochemical analyses are also valuable for modeling electron transfer processes in biological systems. In particular, it has been demonstrated that compounds such as menadione and plumbagin generate reactive oxygen species (ROS) via redox cycling during interactions with NADPH-dependent enzymes (Elhabiri et al., 2014; Futuro et al., 2018). This property underlies both their therapeutic effects and potential toxicity.

3.7. Computational Chemistry and Theoretical Models

In recent years, DFT (Density Functional Theory) calculations have been used as a complement to experimental data. Parameters such as the HOMO–LUMO energy gap, intramolecular charge distribution, dipole moment, and electrostatic potential maps are crucial for predicting the reactivity and photophysical properties of naphthoquinone derivatives (Ibis & Deniz, 2010; Elhabiri et al., 2014).

In the study by İbis and colleagues (Ibis & Deniz, 2010), the electron distribution of S- and O-substituted naphthoquinone dyes was examined, and the theoretical results were found to be consistent with UV–Vis data. This agreement demonstrates that theoretical calculations are a reliable tool for interpreting spectroscopic behavior.

In conclusion, when structural characterization techniques are used in combination, the chemical structure, functional groups, electronic configuration, and redox behavior of naphthoquinone derivatives can be comprehensively defined. The integration of experimental and theoretical methods provides a strong foundation for the rational design of these compounds and for expanding their application areas.

4. ELECTRONIC AND PHOTOPHYSICAL PROPERTIES OF NAPHTHOQUINONES

Naphthoquinones, due to their π-conjugated systems and strong electron-accepting character, are notable compounds in both chemical and electronic applications. These properties are directly related to their molecular orbital structures (HOMO–LUMO levels), the effects of substituents, solvent polarity, electron transfer capabilities, and photophysical behavior (Ibis & Deniz, 2010; Elhabiri et al., 2014; Patan, Göksel, & Ayla, 2021).

The naphthoquinone core inherently possesses a low-energy LUMO level due to the electron-withdrawing effects of its two carbonyl groups. This feature renders the molecule a strong electrophile and enhances its electron-accepting capacity in redox reactions (Elhabiri et al., 2014). Consequently, 1,4-naphthoquinone systems serve as suitable scaffolds for redox-active organic semiconductors, electrochromic materials, and photovoltaic components (Ibis & Deniz, 2010).

4.1. Substituent Effects on Electronic Structure

The nature and position of the substituents largely determine the electronic properties of naphthoquinone derivatives. Electron-donating groups (–NH_2, –OH, –OCH_3) increase the molecule's HOMO energy, thereby lowering the transition energy, whereas electron-withdrawing groups (–NO_2, –CF_3, –CN, –Cl) decrease the LUMO energy, resulting in an increased redox potential (Elhabiri et al., 2014). Studies by İbiş and colleagues have shown that in S- and O-substituted derivatives, the electron density is asymmetrically distributed over the aromatic ring. This asymmetry leads to significant changes in both optical absorption and electrochemical behavior (Ibis & Deniz, 2010). These effects manifest in UV–Vis spectra as bathochromic (red-shift) or hypsochromic (blue-shift) behavior.

In the trifluoromethyl-aryl-amino-naphthoquinone derivatives studied by Yıldırım and colleagues (Yıldırım et al., 2017), a significant narrowing of the HOMO–LUMO gap was observed due to the strong electron-withdrawing nature of the –CF_3 group. This narrowing caused the absorption maxima in the UV–Vis spectra to shift to 370–420 nm. DFT calculations supported this observation, indicating that the electron-withdrawing groups increase the molecule's dipole moment and enhance its polarization.

Electronic interactions depend not only on the nature of the substituent but also on its position. Substituents at the 2-position typically interact with the carbonyl groups through resonance, affecting $\pi \rightarrow \pi^*$ transitions, whereas groups at the 3- or 5-positions primarily exert an inductive effect, altering the redox potential (Ibis & Sahinler Ayla, 2011; Elhabiri et al., 2014). Therefore, substituting at different positions can be strategically used to tune the molecule's optical and electrochemical properties.

4.2. HOMO–LUMO Energy Levels and Theoretical Analyses

DFT (Density Functional Theory) calculations are widely used to understand the electronic structures of naphthoquinone derivatives (Ibis & Deniz, 2010; Elhabiri et al., 2014). Typically, functionals and basis sets such as B3LYP/6-311G(d,p) are employed in these computations. These analyses are used to determine the energies of the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO), and thus the energy gap ($\Delta E = E\_LUMO - E\_HOMO$).

In the study by Elhabiri et al. (Elhabiri et al., 2014), a strong correlation was observed between the redox potentials of menadione and 2-methyl-1,4-naphthoquinone derivatives and their theoretical energy gaps. It was reported that the addition of electron-withdrawing groups lowers the LUMO energy level, thereby increasing the redox potential observed in voltammetric measurements.

In the calculations performed by İbiş et al. (Ibis & Deniz, 2010) on S, O-substituted derivatives, it was observed that –S–R groups enhance the localization of the molecule's frontier orbitals. The electron density is particularly concentrated around the sulfur atom, which directly influences the compound’s photophysical behavior, such as absorption intensity. Thus, when theoretical and experimental data are considered together, it becomes clear that the type of substituent is a key determinant of the electronic character.

4.3. Photophysical Behavior and Solvent Effects

The photophysical properties of naphthoquinone derivatives are primarily investigated using UV–Vis and fluorescence spectroscopy. These compounds mainly exhibit π→π* transitions, while in some cases n→π* transitions also contribute. Solvent polarity has a significant effect on the absorption maxima and molar absorptivity (Ibis & Deniz, 2010; Patan, Göksel, & Ayla, 2021).

In the study by Patan and colleagues (Patan, Göksel, & Ayla, 2021), the absorption behavior of N-substituted derivatives was compared in ethanol and chloroform, showing a red-shift tendency in polar solvents. This effect is explained by solvents with high dipole moments stabilizing the excited state of the molecule.

Additionally, some derivatives have been reported to exhibit photoinduced electron transfer (PET) properties (Ibis & Deniz, 2010). This behavior is particularly pronounced in S- and O-substituted structures, as the binding energies of sulfur and oxygen atoms are low, facilitating electron transfer in the excited state. Such compounds have been suggested as potential materials for optoelectronic devices, including photodiodes, sensors, and OLEDs (Ibis & Deniz, 2010; Elhabiri et al., 2014).

Photophysical properties are also crucial in dye and pigment applications. In studies by İbiş and colleagues (Sayıl, Deniz, & Cınarlı, 2016; Ibis & Deniz, 2010), it was demonstrated that the color performance of naphthoquinone-based dyes on polyester fabrics is directly related to their absorption maxima and binding strengths. The presence of electron-donating groups shifts the color toward warm (red–orange) tones, whereas electron-withdrawing groups shift it toward cooler (blue–green) tones.

4.4. Redox Cycle and Electron Transfer Behavior

Naphthoquinones are potent redox agents that can participate in electron transfer processes in biological systems. This property is also critical for applications in semiconductors and biosensors (Elhabiri et al., 2014; Futuro et al., 2018). Cyclic voltammetry analyses have shown that these compounds typically undergo two consecutive reduction steps (Elhabiri et al., 2014). In the first reduction step, a semiquinone radical is formed, while the second step yields the hydroquinone form.

In the study by Elhabiri and colleagues (Elhabiri et al., 2014), it was reported that this reduction–oxidation cycle is fully reversible and that substituents modulate the redox potentials. Therefore, derivatives modified with appropriate substituents can be utilized as bio-redox systems or energy storage components operating within the desired redox potential range.

The electron transfer behavior of naphthoquinones is also closely related to their biological activities. Reactive oxygen species (ROS) are generated as naphthoquinones enter the redox cycle via NADPH-dependent enzymes (Futuro et al., 2018). This process forms the chemical basis for both therapeutic and toxic effects. In particular, this electron transfer capability plays a crucial role in their anticancer and antifungal activities.

4.5. Molecular Interactions and Conjugation Behavior

Crystal structure analyses have shown that π–π stacking, hydrogen bonding, and dipole–dipole interactions play significant roles in the intermolecular interactions of naphthoquinone derivatives (Becerra et al., 2021; Ibis & Sahinler Ayla, 2011). These interactions extend the conjugated system, improving electron mobility and strengthening optoelectronic properties.

In sulfur-containing derivatives, increased electron delocalization along the C=S–C bridges narrows the optical band gap and enhances conductivity properties (Bayrak, 2017; Becerra et al., 2021). Such interactions play a critical role, particularly in the design of organic semiconductors or redox polymers (Ibis & Deniz, 2010; Elhabiri et al., 2014).

In conclusion, the electronic and photophysical properties of naphthoquinones are susceptible to even minor changes in chemical structure and environmental factors. The choice of substituents, the solvent medium, and the crystal structure directly influence the redox behavior, optical response, and functional performance of these compounds. This versatile structure–property relationship makes naphthoquinones strategic molecules for both scientific research and advanced technological applications.

5. BIOLOGICAL ACTIVITIES AND PHARMACEUTICAL APPLICATIONS

Naphthoquinones are multifunctional secondary metabolites found in both plant and microbial sources. Their capacity for electron transfer, redox activity, and ability to generate reactive oxygen species (ROS) confer significant biological activity (Elhabiri et al., 2014; Futuro et al.,

2018). Recent studies have shown that both natural and synthetic naphthoquinone derivatives exhibit anticancer, antimicrobial, antifungal, antiparasitic, and antioxidant effects (Kacmaz et al., 2019; Yıldırım et al., 2017; Tuyun et al., 2015; Ibis et al., 2011; Futuro et al., 2018). These properties are attributed to versatile biochemical mechanisms arising from the interaction of the naphthoquinone scaffold with various biological targets.

5.1. Biochemical Mechanisms of Action of Naphthoquinones

The biological effects of naphthoquinones are fundamentally based on redox cycling. Intracellular NADPH-dependent enzymes reduce the compound to form a semiquinone radical. This radical reacts with molecular oxygen, generating reactive oxygen species (ROS) such as the superoxide anion (O_2^-) and hydrogen peroxide (H_2O_2) (Elhabiri et al., 2014; Futuro et al., 2018). These oxidative species cause damage to DNA, lipids, and proteins, leading to cellular stress and apoptosis.

In the comprehensive review by Futuro and co-workers (Futuro et al., 2018), this mechanism is emphasized as the chemical basis of both therapeutic and toxic effects. ROS generation leads to increased oxidative stress in cancer cells and disruption of the mitochondrial membrane potential. In contrast, because normal cells possess more effective antioxidant systems, selective cytotoxicity is possible.

Some naphthoquinones also exert their effects via quinone methide (QM) intermediates. These reactive species can form covalent bonds with DNA bases or protein thiol groups (Futuro et al., 2018). This mechanism of action is analogous to that of the anticancer drug mitomycin C, a Streptomyces-derived natural quinone. In addition, inhibition of topoisomerase I and II enzymes plays a critical role in the antitumor activity of β-lapachone and its derivatives (Futuro et al., 2018).

5.2. Anticancer Activity

Although the anticancer properties of naphthoquinones have long been recognized, the activity of newly synthesized derivatives has been increasingly investigated in recent years. Most 1,4-naphthoquinone derivatives exhibit inhibitory effects on DNA replication, mitochondrial function, and the cell cycle (Kacmaz et al., 2019; Ibis et al., 2011; Becerra et al., 2021).

In the study by Kaçmaz and co-workers (Kacmaz et al., 2019), N-, S-, and S, S-substituted 1,4-naphthoquinone compounds were reported to exhibit antiproliferative activity against the HeLa cell line. The presence of electron-withdrawing groups (e.g., –Cl, $–CF_3$) enhanced cytotoxicity, whereas electron-donating groups exerted a more moderate effect. This observation indicates that the redox potential is correlated with the rate of ROS generation.

In a study by Yıldırım and co-workers (Yıldırım et al., 2017) on trifluoromethyl arylamino naphthoquinone derivatives, these compounds were found to induce apoptosis via a β-

lapachone-like mechanism and to cause cell death by disrupting membrane potential. Density functional theory (DFT) calculations supported that the electron-withdrawing effect of the –CF_3 group enhances oxidative stress potential.

Additionally, some naphthoquinone derivatives act as topoisomerase II inhibitors, leading to double-strand DNA breaks (Futuro et al., 2018). This property has been particularly observed in β-lapachone and its derivatives, which have demonstrated activity even in chemoresistant cell lines. Therefore, the quinone scaffold is considered a promising core for the design of chemotherapeutic agents.

5.3. Antimicrobial and Antifungal Effects

Naphthoquinones exhibit broad-spectrum antimicrobial activity against bacteria and fungi. They disrupt cellular integrity either by forming covalent bonds with protein thiol groups in the bacterial membrane or by inducing oxidative stress through redox cycling (Tuyun et al., 2015; Ibiş et al., 2011; Futuro et al., 2018).

In the study by Tuyun and co-workers (Tuyun et al., 2015) on aminonaphthoquinone and benzo[b]phenazine-6,11-dione derivatives, pronounced activity was observed particularly against Gram-positive bacteria (Staphylococcus aureus and Bacillus subtilis). The same compounds were also reported to exhibit antifungal activity against *Candida albicans*, with minimum inhibitory concentrations (MICs) ranging from 16 to 64 μg/mL.

In the comprehensive review by Futuro and co-workers (Futuro et al., 2018), the antifungal activities of 68 naphthoquinones, both natural and synthetic, were compared, and 30 were found to be more potent than currently available antifungal drugs. These compounds exert their effects through multiple mechanisms, including disruption of cell wall synthesis, inhibition of ergosterol biosynthesis, and induction of mitochondrial dysfunction.

Similarly, some azonaphthoquinone derivatives synthesized via $CeCl_3$-catalyzed methods have shown high activity against *Escherichia coli* and *Pseudomonas aeruginosa* strains (Bayrak, 2019). In these compounds, it is thought that the increase in electron density on the conjugated system by aromatic amine groups facilitates the generation of reactive oxygen species at the cell membrane.

In comparative biological evaluations of S-, O-, and N-substituted naphthoquinones, sulfur-containing derivatives were found to exhibit notably higher antibacterial and antifungal activity (Ibiş et al., 2011; Becerra et al., 2021). This is attributed to sulfur's soft nucleophilic character, which enables more stable interactions with biological targets, such as enzyme thiol groups.

5.4. Antiparasitic and Antimalarial Activities

Menadione (vitamin K_3) and its derivatives exhibit potent antimalarial activity against *Plasmodium* species. Electrochemical studies by Elhabiri and co-workers (Elhabiri et al., 2014)

demonstrated that menadione derivatives inhibit NADPH-dependent glutathione reductase and thioredoxin reductase enzymes, thereby weakening the parasite's oxidative defense mechanisms.

In the same study, menadione was reported to act as a "subversive substrate" in the electron transfer chain, with the semiquinone intermediate reacting with O_2 to generate toxic superoxide and peroxide species for the parasite (Elhabiri et al., 2014). This mechanism has emerged as a particularly effective biochemical strategy against *Plasmodium falciparum*.

Plumbagin and its derivatives have been reported to exhibit similar effects against *Leishmania amazonensis* and *Trypanosoma rhodesiense* (Futuro et al., 2018). These compounds disrupt mitochondrial functions, leading to the collapse of energy metabolism. In particular, derivatives containing aromatic sulfur groups were observed to cross cell membranes more readily due to increased lipophilicity and, consequently, to be more effective in parasite cells (Tuyun et al., 2015; Elhabiri et al., 2014).

5.5. Antioxidant and Enzymatic Effects

Although naphthoquinones are generally known as inducers of oxidative stress, some derivatives can exhibit antioxidant behavior. Compounds substituted with electron-donating groups (e.g., –OH or $–OCH_3$) can stabilize free radicals and terminate chain reactions (Bayrak, 2017; Becerra et al., 2021). This behavior mainly arises from the reversible nature of the quinone/hydroquinone redox couple.

It has also been determined that some derivatives act as enzyme modulators. The interaction of menadione with enzymes such as glutathione reductase, thioredoxin reductase, and NADPH–quinone oxidoreductase can influence cellular defense systems (Elhabiri et al., 2014). This property is therapeutically valuable for targeting and modulating the balance of oxidative stress.

5.6. Structure–Activity Relationships (SAR) and Pharmacophore Features

There is a clear correlation between the biological activities of naphthoquinones and their structural features.

- Electron-withdrawing substituents (–Cl, $–CF_3$, $–NO_2$) increase the redox potential and accelerate ROS generation, thereby enhancing anticancer and antimicrobial activity (Yıldırım et al., 2017; Ibis et al., 2011).
- Electron-donating substituents ($–NH_2$, –OH, $–OCH_3$) generally exhibit better antioxidant properties and reduce toxicity (Bayrak, 2017; Becerra et al., 2021).
- The position of substitution (particularly the 2- and 3-positions) determines reactivity and lipophilicity; substituents at the 2-position are generally more influential on redox behavior (Elhabiri et al., 2014; Patan, Göksel, & Ayla, 2021).

These relationships are instructive for target-oriented drug design. For example, derivatization of 2,3-dichloro-1,4-naphthoquinone derivatives with piperidine or piperonylamine has been shown to both enhance biological activity and improve solubility (Patan, Göksel, & Ayla, 2021).

SAR analyses are also crucial for toxicity control. Maintaining the balance between ROS production and antitumor activity is critical to maintaining a safe therapeutic index (Futuro et al., 2018).

5.7. Clinical and Pharmaceutical Perspectives

Naphthoquinone-based compounds are already included in specific pharmaceutical formulations. Vitamin K derivatives are used in blood coagulation; β-lapachone and menadione-like derivatives are being investigated for antitumor applications; and plumbagin is being evaluated as an antimicrobial and antiparasitic prototype (Elhabiri et al., 2014; Futuro et al., 2018).

The pharmaceutical potential of these compounds has become particularly significant for redox-controlled therapeutic strategies (e.g., inducing oxidative stress in cancer cells). However, issues of toxicity and selectivity remain challenges that still need to be addressed. Preclinical studies have reported that some derivatives may cause damage to liver and kidney tissues by accumulating ROS (Futuro et al., 2018). Therefore, future research is expected to be more target-oriented and based on pharmacophore optimization.

In conclusion, naphthoquinone derivatives are versatile molecules that can exhibit significant biological differences due to minor structural changes. Their electron-transfer properties make them powerful agents from both biochemical and pharmaceutical perspectives, while redox-controlled toxicity management is considered a key strategy for future clinical applications.

6. INDUSTRIAL AND TECHNOLOGICAL APPLICATIONS

Naphthoquinones possess a wide range of applications not only in biological fields but also in industrial, energy, environmental, and materials technologies due to their redox-active structures, optical properties, chemical stability, and electron-transfer capabilities (Sayıl, Deniz & Cınarlı, 2016; Tuyun et al., 2015; Becerra et al., 2021; Ibis & Deniz, 2010; Elhabiri et al., 2014). In this section, the industrial-scale applications of naphthoquinone derivatives are examined under six main headings: (i) dye and pigment technology, (ii) photochromic and electrochromic materials, (iii) sensor and biosensor systems, (iv) energy storage and conversion systems, (v) corrosion inhibitors, and (vi) environmental and biotechnological applications.

6.1. Naphthoquinones in Dye and Pigment Technology

Naphthoquinone derivatives have long been used as natural and synthetic colorants. Their conjugated π-system and carbonyl groups enable absorption of light in the visible region. In

addition, the color tones can be readily tuned through the electronic effects of substituents (Sayil, Deniz, & Cınarlı, 2016; Ibis & Deniz, 2010).

In the study by Sayıl and co-workers (Sayil, Deniz, & Cınarlı, 2016), S- and O-substituted 1,4-naphthoquinone derivatives exhibited high color fastness, thermal stability, and washing resistance, particularly on polyester fabrics. The color intensity varied depending on the type of substituent, ranging from yellow to red and further to brownish tones.

Electron-donating groups (–OCH_3, –NH_2) shift the absorption maximum toward the red region, whereas electron-withdrawing groups (–Cl, –NO_2, –CF_3) cause a shift toward the blue region (Yıldırım et al., 2017; Ibis & Deniz, 2010). This electronic tunability makes naphthoquinone dyes attractive for use in textiles, plastics, optical filters, and artistic pigments. Moreover, owing to their high light fastness and photothermal stability, they are also employed in long-lasting coating materials (Ibis & Deniz, 2010).

6.2. Photochromic and Electrochromic Materials

The redox-active nature of naphthoquinones makes them ideal candidates for electrochromic (color-changing) materials (Ibis & Deniz, 2010; Elhabiri et al., 2014). During redox cycling, the partially reduced (semiquinone) or fully reduced (hydroquinone) forms of 1,4-naphthoquinone exhibit different absorption characteristics. This property enables color changes to be induced under electrical stimulation.

In theoretical and experimental studies by İbiş and co-workers (Ibis & Deniz, 2010), the electrochromic behavior of S, O-disubstituted naphthoquinone derivatives was investigated. These compounds were shown to undergo reduction, exhibit color changes at low potentials, and return to their original color upon re-oxidation. This property is significant for organic electrochromic windows and display technologies.

In addition, some derivatives have been found to exhibit photochromic behavior. These compounds, which undergo temporary color changes under UV light, are being explored for applications in photothermal energy conversion and optical storage systems (Ibis & Deniz, 2010)-the versatile optoelectronic properties of naphthoquinones open new horizons in both energy-efficient material design and bright surface technologies.

6.3. Sensor and Biosensor Applications

Naphthoquinones are widely used in sensor and biosensor design due to their electrochemically active structures and suitability for surface modification (Elhabiri et al., 2014). Their carbonyl groups can be readily reduced and oxidized during redox transitions on electrode surfaces, enabling the development of highly sensitive sensor systems.

In studies by Elhabiri and co-workers (Elhabiri et al., 2014), the redox couples of menadione (vitamin K_3) and its derivatives were immobilized on electrode surfaces. They demonstrated

the ability to detect biomolecules, including glucose, ascorbate, and NADH. The compounds' propensity for two-electron transfer enhanced the reproducibility of the sensor signals.

Moreover, it has been reported that sulfur- and oxygen-containing derivatives can bind more strongly to electrode surfaces, forming stable film layers on gold electrodes (Ibis & Deniz, 2010; Elhabiri et al., 2014). This property ensures long-lasting and reproducible biosensors.

Recent studies have also shown that naphthoquinone-based nanocomposite sensors can be used for electrochemical detection of environmental toxins, including heavy metal ions and pesticides. The surface redox transitions of naphthoquinone generate current signals that change upon analyte binding, providing high sensitivity (Elhabiri et al., 2014).

6.4. Energy Storage and Conversion Systems

Naphthoquinone derivatives play an essential role in energy storage systems (batteries, supercapacitors, fuel cells) due to their electron-accepting properties (Ibis & Deniz, 2010; Elhabiri et al., 2014). In particular, carbonyl redox couples (C=O/C–OH) enable the use of these compounds as organic electrode materials.

In the study by Elhabiri and co-workers (Elhabiri et al., 2014), cyclic voltammetry analysis of menadione and its derivatives revealed two reversible redox couples, indicating their suitability for energy storage applications. Similarly, the stable semiconducting phase formed during the reduction of 2,3-dichloro-1,4-naphthoquinone was found to be promising for organic solid electrodes.

An advantage of naphthoquinone-based systems is that they are low-cost, environmentally friendly, and free of heavy metals. In organic batteries, these compounds facilitate energy conversion through coupled electron–proton redox reactions. Additionally, the high redox reversibility of their derivatives supports stability in multi-cycle systems (Ibis & Deniz, 2010; Elhabiri et al., 2014).

The use of naphthoquinones as electron-acceptor materials in photovoltaic systems (organic solar cells) is also being explored (Ibis & Deniz, 2010). Their conjugated π-systems and low LUMO levels contribute to efficient charge transfer in donor–acceptor structured solar cells. By modifying substituents, the bandgap can be tuned, allowing optimization of device efficiency.

6.5. Corrosion Inhibitors

The redox behavior and aromatic structures of naphthoquinones make them suitable for use as protective agents on metal surfaces (Tuyun et al., 2015; Becerra et al., 2021). In particular, nitrogen- and sulfur-containing derivatives adsorb onto metal surfaces, preventing oxidative degradation.

Sulfur-substituted 1,4-naphthoquinones have been reported to provide cathodic protection on steel surfaces in corrosive environments (Becerra et al., 2021). The π-electron-rich rings of these compounds interact with the metal surface to form coordination bonds, thereby forming a protective film. These films provide both a physical barrier and chemical passivation.

These properties of naphthoquinone derivatives are essential for the development of environmentally friendly organic corrosion inhibitors, offering potential alternatives to traditional heavy metal-based inhibitors.

6.6. Environmental and Biotechnological Applications

Naphthoquinones have also attracted interest in environmental biotechnology. Microbially derived quinones, in particular, are used as electron carriers in biosensors, biocatalytic systems, and microbial fuel cells (Elhabiri et al., 2014; Futuro et al., 2018). Compounds such as menadione facilitate electron transfer by serving as redox mediators in microorganisms.

Additionally, naphthoquinone derivatives can serve as photoactive catalysts or as reactive materials for degrading dyes (Ibis & Deniz, 2010). Under UV light, they generate reactive oxygen species that decompose organic pollutants, making them especially valuable for environmentally friendly photocatalytic treatment systems.

The biological degradability of newly synthesized derivatives offers an advantage in reducing environmental toxicity and provides sustainable solutions for both energy and environmental engineering applications.

In conclusion, naphthoquinone derivatives are not only biologically active molecules at the laboratory scale but also versatile compounds central to optical, electronic, and chemical technologies. Their redox-controlled behavior has led to increasing applications in dyes, sensors, energy systems, corrosion protection, and environmental technologies. These properties elevate naphthoquinones beyond traditional pharmaceutical chemistry, positioning them as a strategically important family of molecules in advanced materials science

7. FUTURE PERSPECTIVES AND CONCLUSION

Naphthoquinone derivatives, owing to their redox-active structures, broad spectrum of biological activities, and electronic properties, occupy a central position in multidisciplinary research fields ranging from organic chemistry to materials science (Sayil, Deniz, & Cınarlı, 2016; Tuyun et al., 2015; Ibis & Deniz, 2010; Elhabiri et al., 2014; Futuro et al., 2018). The diversity of their synthetic routes, the ability to rationally tune structure–property relationships through functional groups, and their chemical similarity to natural derivatives make them robust model systems for both fundamental research and applied technologies.

7.1. Future Research Trends

Recent studies indicate that naphthoquinone chemistry is increasingly intersecting with emerging research areas such as biotechnology, nanotechnology, and green chemistry. Within this framework, future research is expected to focus on the following themes:

a) Green Synthesis and Sustainable Processes

The use of environmentally friendly solvents (ethanol, water, ionic liquids) and catalyst-free reactions in the synthesis of naphthoquinone derivatives provides lower energy consumption and reduced waste generation compared to traditional methods (Tuyun et al., 2015; Patan, Göksel, & Ayla, 2021). In the future, microwave- or ultrasound-assisted processes, biocatalytic transformations, and continuous-flow systems are expected to enhance efficiency and pave the way for industrially viable "green synthesis" approaches.

b) AI-Assisted Design and Computational Modeling

DFT-based calculations have increasingly become powerful tools for predicting the electronic structures, redox potentials, and interactions of naphthoquinones with biological targets (Ibis & Deniz, 2010; Elhabiri et al., 2014). In the future, combining these theoretical approaches with machine learning algorithms will enable rapid screening of hundreds of derivatives through QSPR (Quantitative Structure–Property Relationship) and QSAR (Quantitative Structure–Activity Relationship) models. This will allow for the prediction of potential biologically active compounds before synthesis and accelerate targeted drug design.

c) Multifunctional Material Design

The photophysical and electrochemical properties of naphthoquinones enable them to serve as components of hybrid systems that simultaneously provide energy storage and optical response (Ibis & Deniz, 2010). For example, bright surfaces in electrochromic devices can be designed to both change color and store energy, representing a novel example of the "multifunctional organic materials" concept.

Furthermore, naphthoquinone derivatives combined with metal complexes or nanostructures have the potential to create new synergies in photovoltaic efficiency, corrosion resistance, and sensor sensitivity (Becerra et al., 2021; Elhabiri et al., 2014).

d) Targeted Pharmaceutical Strategies

The anticancer, antifungal, and antiparasitic activities of naphthoquinones are promising; however, issues of selectivity and toxicity remain key challenges (Kacmaz et al., 2019; Yıldırım et al., 2017; Futuro et al., 2018). In the future, the development of targeted naphthoquinone derivatives via bioconjugation techniques (e.g., peptide or polymer carrier systems) may help reduce systemic toxicity.

Additionally, prodrug strategies could enable redox-potential–controlled release mechanisms. This approach would activate ROS generation specifically in tumor tissues, thereby protecting healthy cells (Elhabiri et al., 2014; Futuro et al., 2018).

e) Biotechnological and Environmental Integration

The use of menadione and similar compounds as electron-transfer mediators in microbial fuel cells is increasing (Elhabiri et al., 2014; Futuro et al., 2018). Naphthoquinones can become critical molecules for sustainable energy solutions, both in biosensor design and microbial energy conversion. Moreover, due to their photocatalytic properties, they hold promise for the development of environmentally friendly photoredox systems capable of degrading organic pollutants (Ibis & Deniz, 2010).

7.2. Scientific and Technological Added Value

The development of naphthoquinone chemistry provides scientific added value in three key dimensions:

1. **At the molecular level**, elucidating the relationships among electron density, redox behavior, and substituent effects has provided a deep knowledge base in organic and physical chemistry (Elhabiri et al., 2014; Patan, Göksel, & Ayla, 2021).
2. **At the applied level**, naphthoquinone derivatives provide a unifying chemical platform across diverse fields such as materials, energy, sensors, and biotechnology, paving the way for interdisciplinary, innovative approaches (Ibis & Deniz, 2010; Elhabiri et al., 2014).
3. **At the societal and environmental level**s, the development of controllable, biodegradable, and metal-free organic compounds contributes to the vision of sustainable chemistry (Tuyun et al., 2015; Ibis & Deniz, 2010).

7.3. Conclusion

Naphthoquinone derivatives are chemically inspired by nature but can be redesigned in the laboratory. Their common feature is remarkable versatility, encompassing chemical reactivity, biological activity, and functional applications. Thanks to their redox cycles, substituent modifications, and conjugated systems, naphthoquinones can engage strongly with biological targets while simultaneously serving as functional materials in electronic devices (Ibis & Deniz, 2010; Elhabiri et al., 2014; Futuro et al., 2018).

Comprehensive studies spanning from synthesis to application have demonstrated that these compounds play a multifaceted role not only in organic chemistry but also in biophysics, pharmaceuticals, materials science, and environmental engineering.

The chemistry of the future will be built on multifunctional molecules capable of providing holistic solutions to energy and health challenges. Naphthoquinones stand out as among the most promising building blocks for this transformation.

Thanks to their natural origins, tunable structures, and rich redox chemistry, naphthoquinones will continue to occupy a central role in the biotechnological and sustainable chemistry paradigms of the 21st century.

KAYNAKÇA

Elhabiri, M., Sidorov, P., Cesar-Rodo, E., Marcou, G., Lanfranchi, D. A., Davioud-Charvet, E., Horvath, D., & Varnek, A. (2014). Electrochemical properties of substituted 2-methyl-1,4-naphthoquinones: Redox behavior predictions. *Chemistry – A European Journal.* https://doi.org/10.1002/chem.201403703.

Futuro, D. O., Ferreira, P. G., Nicoletti, C. D., Borbasantos, L. P., Da Silva, F. C., Rozental, S., & Ferreira, V. F. (2018). The antifungal activity of naphthoquinones: An integrative review. *Anais da Academia Brasileira de Ciências, 90*(1 Suppl. 2), 1187–1214. https://doi.org/10.1590/0001-3765201820170815

Sayil, C., & Ibis, C. (2010). Synthesis and spectral properties of 1,4-naphthoquinone sulfanyl derivatives. *Russian Journal of Organic Chemistry, 46*(2), 209–215. https://doi.org/10.1134/S1070428010020119

Bayrak, N. (2019). Novel azanaphtoquinone compounds with aromatic amino moiety: Synthesis, structural characterization, and antimicrobial features. *Journal of Molecular Structure, 1195,* 411–416. https://doi.org/10.1016/j.molstruc.2019.05.093

Ibis, C., & Deniz, N. G. (2010). Synthesis and spectroscopic properties of S-,O-substituted naphthoquinone dyes. *Phosphorus, Sulfur, and Silicon and the Related Elements, 185*(11), 2324–2332. https://doi.org/10.1080/10426501003598671

Patan, A., Göksel, F. S., & Ayla, S. S. (2021). Reactions of 2,3-dichloro-1,4-naphthoquinone with piperidine, amine and some thiol nucleophile. *Phosphorus, Sulfur, and Silicon and the Related Elements.* https://doi.org/10.1080/10426507.2021.1901700

Bayrak, N. (2017). Novel straight-chained sulfanyl members of arylamino-1,4-naphthoquinones: Synthesis and characterization. *JOTCSA, 4*(2), 597–606. https://doi.org/10.18596/jotcsa.301558

Ibis, C., Sahinler Ayla, S., & Gulsah Deniz, N. (2010). Synthesis, Structural, Spectroscopic and Electrochemical Properties of Novel S- and S,O-Substituted 1,4-Naphthoquinones. *Asian Journal of Chemistry*, *22*(8), 6195–6204. https://asianpubs.org/index.php/ajchem/article/view/11802

Ibis, C., & Sahinler Ayla, S. (2011). Synthesis and spectroscopic evaluation of novel N-, S-, and O-substituted 1,4-naphthoquinone derivatives. *Phosphorus, Sulfur, and Silicon and the Related Elements, 186*(12), 2350–2356. https://doi.org/10.1080/10426507.2011.602376

Kacmaz, A., Deniz, N., Aydinli, S., Sayil, C., Onay-Ucar, E., Mertoglu, E. & Arda, N. (2019). Synthesis and antiproliferative evaluation of some 1,4-naphthoquinone derivatives against human cervical cancer cells. *Open Chemistry*, *17*(1), 337-345. https://doi.org/10.1515/chem-2019-0030

Tuyun, A. F., Bayrak, N., Yıldırım, H., Onul, N., Mataraci Kara, E., & Ozbek Celik, B. (2015). Synthesis and in vitro biological evaluation of aminonaphthoquinones and benzo[b]phenazine-6,11-dione derivatives as potential antibacterial and antifungal compounds. *Journal of Chemistry, 2015,* Article 645902, 8 pages. https://doi.org/10.1155/2015/645902

Sayil, C., Deniz, N. G., & Cinarli, A. (2016). Synthesis of N-, S-, O-substituted quinone dyes and their dyeability on polyester fibers. *Progress in Organic Coatings, 98,* 39–42. https://doi.org/10.1016/j.porgcoat.2016.04.017

Ibis, C., & Deniz, N. G. (2010). Synthesis and spectroscopic properties of S-,O-substituted naphthoquinone dyes. *Phosphorus, Sulfur, and Silicon and the Related Elements, 185*(11), 2324–2332. https://doi.org/10.1080/10426501003598671

Yavuz, Ş. (2023). Synthesis, Characterization, and Investigation of Antimicrobial Activities of New Naphthoquinone Compounds from 2-(butylthio)-3-chloronaphthalene-1,4-dione. Sinop Üniversitesi Fen Bilimleri Dergisi, 8(2), 156-168. https://doi.org/10.33484/sinopfbd.1362604

Ibis, C., Şahinler Ayla, S., & Yavuz, Ş. (2019). Reactions of quinones with some aryl phenols and synthesis of new quinone derivatives. *Synthetic Communications, 49*(2), 202–211. https://doi.org/10.1080/00397911.2018.1546403

Yıldırım, H., Bayrak, N., Tuyun, A. F., Mataracı Kara, E., Ozbek Çelik, B., & Gupta, G. (2017). 2,3-Disubstituted-1,4-naphthoquinones containing an arylamine with trifluoromethyl group: Synthesis, biological evaluation, and computational study. *RSC Advances, 7,* 25753–25764. https://doi.org/10.1039/c7ra00868f

Becerra, N. A., Espinosa-Bustos, C., Vázquez, K., Rivera, G., Paulino, M., Cantero, J., Nogueda, B., Chacón-Vargas, F., Castillo-Velazquez, U., Elizondo Rodríguez, A. F., Toledo, S., Moreno-Rodríguez, A., Aranda, M., & Salas, C. O. (2021). Expanding the chemical space of aryloxy-naphthoquinones as potential anti-chagasic agents: Synthesis and trypanosomicidal activity [Preprint]. *Research Square.* https://doi.org/10.21203/rs.3.rs-756093/v1

Ibiş, C., Tuyun, A. F., Özsoy-Güneş, Z., Bahar, H., Stasevych, M. V., Musyanovych, R. Y., Komarovska-Porokhnyavets, O., & Novikov, V. (2011). Synthesis and biological evaluation of novel nitrogen- and sulfur-containing hetero-1,4-naphthoquinones as potent antifungal and antibacterial agents. *European Journal of Medicinal Chemistry, 46*, 5861–5867. https://doi.org/10.1016/j.ejmech.2011.09.048