

TEHNIČKI GLASNIK - TECHNICAL JOURNAL

Scientific-professional journal of University North

Volume 16

Varaždin, December 2022

Issue 4

Pages 445-580

Editorial Office:

Sveučilište Sjever / University North – Tehnički glasnik / Technical journal
 Sveučilišni centar Varaždin / University Center Varaždin
 Jurja Križanića 31b, 42000 Varaždin, Croatia
 Tel. ++385 42 493 338, Fax. ++385 42 493 336
 E-mail: tehnickiglasnik@unin.hr
<https://tehnickiglasnik.unin.hr>
<https://www.unin.hr/djelatnost/izdavastvo/tehnicki-glasnik/>
<https://hrcak.srce.hr/tehnickiglasnik>

Founder and Publisher:

Sveučilište Sjever / University North

Council of Journal:

Marin MILKOVIĆ, Chairman; Anica HUNJET, Member; Goran KOZINA, Member; Mario TOMIŠA, Member;
 Vlado TROPŠA, Member; Damir VUSIĆ, Member; Milan KLJAJIN, Member; Anatolii KOVROV, Member

Editorial Board:**Domestic Members:**

Chairman Damir VUSIĆ (1), Milan KLJAJIN (1), Marin MILKOVIĆ (1), Krešimir BUNTAK (1), Anica HUNJET (1), Živko KONDIĆ (1), Goran KOZINA (1), Ljudevit KRPAJ (1), Krunoslav HAJDEK (1), Marko STOJČIĆ (1), Božo SOLDO (1), Mario TOMIŠA (1), Vlado TROPŠA (1), Vinko VIŠNJIĆ (1), Sanja ŠOLIĆ (1), Dean VALDEC (1), Duško PAVLETIĆ (5), Branimir PAVKOVIĆ (5), Mile MATIJEVIĆ (3), Damir MODRIĆ (3), Nikola MRVAC (3), Klaudio PAP (3), Ivana ŽILJAK STANIMIROVIĆ (3), Krešimir GRILEC (6), Biserka RUNJE (6), Sara HAVRLIŠAN (2), Dražan KOZAK (2), Roberto LUJIĆ (2), Leon MAGLIĆ (2), Ivan SAMARDŽIĆ (2), Antun STOJČIĆ (2), Katica ŠIMUNOVIĆ (2), Goran ŠIMUNOVIĆ (2), Ladislav LAZIĆ (7), Ante ČIKIĆ (1)/(2), Darko DUKIĆ (9), Gordana DUKIĆ (10), Srđan MEDIĆ (11), Sanja KALAMBURA (12), Marko DUNĐER (13), Zlata DOLAČEK-ALDUK (4), Dina STÖBER (4)

International Members:

Boris TOVORNIK (14), Milan KUHTA (15), Nenad INJAC (16), Marin PETROVIĆ (18), Salim IBRAHIMEFENDIĆ (19), Zoran LOVREKOVIĆ (20), Igor BUDAK (21), Darko BAJIĆ (22), Tomáš HANÁK (23), Evgenij KLIMENKO (24), Oleg POPOV (24), Ivo ČOLAK (25), Katarina MONKOVA (26), Berenika HAUSNEROVA (8), Nenad GUBELJAK (27), Charlie FAIRFIELD (28), Stefanija KLARIĆ (28), Bertrand MARESCHAL (29), Sachin R. SAKHARE (30), Suresh LIMKAR (31), Mandeep KAUR (32), Aleksandar SEDMAK (33), Han-Chieh CHAO (34), Sergej HLOCH (26), Grzegorz M. KRÓLCZYK (35), Djordje VUKELIĆ (21), Stanislaw LEGUTKO (17), Valentin POPOV (36), Dragan MARINKOVIC (36), Hamid M. SEDIIGHI (37), Cristiano FRAGASSA (38), Dragan PAMUČAR (39)

Editor-in-Chief:

Milan KLJAJIN

Technical Editor:

Goran KOZINA

Graphics Editor:

Snježana IVANČIĆ VALENKO

IT support:

Tomislav HORVAT

Print:

Centar za digitalno nakladništvo, Sveučilište Sjever

All manuscripts published in journal have been reviewed.**Manuscripts are not returned.****The journal is free of charge and four issues per year are published**

(In March, June, September and December)

Circulation: 100 copies**Journal is indexed and abstracted in:**

Web of Science Core Collection (Emerging Sources Citation Index - ESCI), EBSCOhost Academic Search Complete, EBSCOhost – One Belt, One Road Reference Source Product, ERIH PLUS, CITEFACTOR – Academic Scientific Journals, DOAJ – Directory of Open Access Journals, Hrcak – Portal znanstvenih časopisa RH

Registration of journal:The journal "Tehnički glasnik" is listed in the HGK Register on the issuance and distribution of printed editions on the 18th October 2007 under number 825.**Preparation ended:**

September 12, 2022

Published:

September 23, 2022

Legend:

(1) University North, (2) University of Slavonki Brod, (3) Faculty of Graphic Arts Zagreb, (4) Faculty of Civil Engineering Osijek, (5) Faculty of Engineering Rijeka, (6) Faculty of Mechanical Engineering and Naval Architecture Zagreb, (7) Faculty of Metallurgy Sisak, (8) Tomas Bata University in Zlin, (9) Department of Physics of the University of Josip Juraj Strossmayer in Osijek, (10) Faculty of Humanities and Social Sciences Osijek, (11) Karlovac University of Applied Sciences, (12) University of Applied Sciences Velika Gorica, (13) Department of Polytechnics - Faculty of Humanities and Social Sciences Rijeka, (14) Faculty of Electrical Engineering and Computer Science - University of Maribor, (15) Faculty of Civil Engineering - University of Maribor, (16) University College of Teacher Education of Christian Churches Vienna/Krems, (17) Faculty of Mechanical Engineering - Poznan University of Technology (Poland), (18) Mechanical Engineering Faculty Sarajevo, (19) University of Travnik - Faculty of Technical Studies, (20) Higher Education Technical School of Professional Studies in Novi Sad, (21) University of Novi Sad - Faculty of Technical Sciences, (22) Faculty of Mechanical Engineering - University of Montenegro, (23) Brno University of Technology, (24) Odessa State Academy of Civil Engineering and Architecture, (25) Faculty of Civil Engineering - University of Mostar, (26) Faculty of Manufacturing Technologies with the seat in Prešov - Technical University in Košice, (27) Faculty of Mechanical Engineering - University of Maribor, (28) College of Engineering, IT & Environment - Charles Darwin University, (29) Université Libre de Bruxelles, (30) Vishwakarma Institute of Information Technology (Pune, India), (31) AISSMS Institute of Information Technology (Pune, India), (32) Permtech Research Solutions (India), (33) University of Belgrade, (34) National Dong Hwa University - Taiwan, (35) Faculty of Mechanical Engineering - Opole University of Technology (Poland), (36) TU Berlin - Germany, (37) Shahid Chamran University of Ahvaz - Iran, (38) University of Bologna - Italy, (39) University of Defence in Belgrade - Military Academy - Serbia

CONTENT	I
Ana Svalina, Mario Tomiša, Marko Čačić, Krunoslav Hajdek* Synthesis of Current Knowledge and Research on the Design Thinking Methodology	445
Maosheng Zheng*, Yi Wang, Haipeng Teng Hybrid of "Intersection" Algorithm for Multi-Objective Optimization with Response Surface Methodology and its Application	454
Drago Pupavac*, Marija Ivaniš A Study on the Interdependence of Quality and Satisfaction with Public Transport in Rijeka	458
Hossein Safari, Parisa Sami Khaftani* The Role of Green Playground on Social Interactions of Children (Case Study: Pasdaran Park in Bandar Anzali)	464
Mehmet Şen*, Muciz Özcan Reducing Short Circuit Current with a Novel Approach Model in Wind Power Plants	471
Ilker Kilic A Lévy Flight Based BAT Optimization Algorithm for Block-based Image Compression	477
Oğuzhan Katar*, Erkan Duman Automated Semantic Segmentation for Autonomous Railway Vehicles	484
Adem Dalcalı*, Mehmet Akbaba Performance Optimisation of Single-Phase Variable Reluctance Shaded-Pole Motor	491
F. Kebire Bardak*, M. Nuri Seyman, Feyzullah Temurtaş EEG Based Emotion Prediction with Neural Network Models	497
A. Vinay, Nipun Bhat, Paras S. Khurana*, Vishruth Lakshminarayanan, Vivek Nagesh, S. Natarajan, T. B. Sudarshan AFMB-Net: DeepFake Detection Network Using Heart Rate Analysis	503
Matija Mikac*, Robert Logožar, Miroslav Horvatić Performance Comparison of Open Source and Commercial Computing Tools in Educational and Other Use — Scilab vs. MATLAB	509
Saida Hamioud Free Vibration of a Cantilever Euler-Bernoulli Beam Carrying a Point Mass by Using SEM	519
Nenad Bedeković, Ladislav Havaš, Tomislav Horvat*, Dražen Crčić The Importance of Developing Preventive Techniques for SQL Injection Attacks	523
Ivana Alpeza*, Ivna Linke, Biserka Runje How to Improve the Monitoring of Subjective Analytical Methods and Assessors - The Case of Wine	530
Atanas Kochov*, Martin Spasovski, Filip Serafimovski Tensile Strength of Carbon Fiber	537
Tomislav Mrčela*, Tamara Zadravec Information Model Development – Attributes and Power of Information	542
Muhammed Abdullah Ozel*, Selim Sefa Baysal, Cemil Sungur Development of a Test System to Make Circularity Measurements of Rubber Bushing After Stress Relief	546
Bircan Çalışır*, Ayhan Akbal A New RF Satellite Link Analyzing and Antenna Effect on Satellite Communication	550
Ahmad Farmahini Farahani Prioritizing the Researchers' Important Competencies in Research Centres	557
Andrija Bernik*, Damir Vusić, Voravika Wattanasoontorn Computer Game Elements and its Impact on Higher Education	566
Stefan Schmidt*, Benjamin Stefan Godwin Schmidt A Working from Home (WFH) 2022 - Best Practices Implementation and Case Study of Company Agreement	572
INSTRUCTIONS FOR AUTHORS	V

Co-organized by:

Universiti Malaysia Pahang (UMP)

Universiti Malaysia Perlis (UniMAP)

FH JOANNEUM University of Applied Sciences (Austria)



The success of ICSCAs include the organization of 11 conferences taking place worldwide. The history of ICSCA started in 2012 in Singapore, next it was hosted in quite a number of countries and regions including France, the Great Britain, Hong Kong, Thailand, Brunei, and last 4 years it was organized in Malaysia.

ICSCA 2023 is where you will meet and network with leaders, business innovators, researchers, academics and policy makers from across the globe. It's an annual event over the past 11 years and it provides a forum for researchers, practitioners and educators to present and discuss the most recent innovations, trends, experiences and concerns in the field of Software and Computer Applications.

Call For Papers**Topics of interest for submission include, but are not****Track 1. Software Engineering**

- ◆ Requirements Engineering
- ◆ Software Architecture and Design
- ◆ Software Quality Assurance and Usability Issues
- ◆ Empirical Software Engineering
- ◆ Formal Methods

Track 2. Computer Applications

- ◆ Operating Systems
- ◆ Cyber-Physical Systems
- ◆ IoT and Cloud Computing
- ◆ Information Systems
- ◆ Blockchain Technology

Track 3. Computational Intelligence and AI techniques

- ◆ Algorithms and Protocols
- ◆ Big Data
- ◆ Meta-heuristics Algorithms
- ◆ High Performance Computing
- ◆ Natural Language Processing

Visit this page for more topics: <http://www.icsca.org/cfp.html>

Submission

1. Full Paper (Publication and Presentation)
2. Abstract (Presentation Only)

※ For full paper(.pdf), please upload to <https://www.zmeeting.org/-submission/icsca2023>

※ For abstract, please send it to icsca_general@163.com

More details about submission, please visit <http://www.icsca.org/author.html>

Submission Deadline: Oct. 5, 2022

Call For Special Sessions

ICSCA 2023 solicits proposals for special sessions within the technical scope of the conference. Special sessions supplement the regular program of the conference and provide a sample of the state-of-the-art research in both academia and industry in special, novel, challenging, and emerging topics.

Special-session proposals should be submitted by the prospective organizer(s) who will commit to promoting and handling the review process of their special session as Chairs or Co-Chairs of the event

Proposal Submission Deadline: December 15, 2022

Paper Requirement

English is the official language of the conference; the paper should be written and presented only in English.

Articles submitted to the conference should report original, previously unpublished research results, experimental or theoretical and must not be under consideration for publication elsewhere. We firmly believe that ethical conduct is the most essential virtual of any academic. Hence any act of plagiarism is a totally unacceptable academic misconduct and cannot be tolerated. If an author is found to commit an act of plagiarism, the following acts of sanction will be taken:

- 1). Reject the article submitted or delete the article from the final publications.
- 2). Report the authors violation to his/her supervisor(s) and affiliated institution(s)
- 3). Report the authors violation to the appropriate overseeing office of academic ethics and research funding agency.
- 4). Reserve the right to publish the authors name(s), the title of the article, the name(s) of the affiliated institution and the details of misconduct, etc. of the plagiarist

Conference Schedule

February 23, 2023

Sign-in

February 24, 2023

Morning: Opening Ceremony & Keynote Speeches & Group Photo

Afternoon: Parallel Sessions

February 25, 2023

Parallel Sessions & Tutorial & Closing Ceremony

Contact

Freda C. Lee (Conference Secretary)

Email: icsca_general@163.com | icsca_general@outlook.com

Tel: +861-32-777777-0

Monday to Friday - 9 am till 6 pm (GMT+8)

Synthesis of Current Knowledge and Research on the Design Thinking Methodology

Ana Svalina, Mario Tomiša, Marko Čačić, Krunoslav Hajdek*

Abstract: Design thinking is an innovative way of thinking and solving problems in which the focus is on the user. This paper provides an overview of current research on the topic to gain insight into current practices and create a basis for further development of the area. Synthesis of knowledge is presented through historical development, methodology, and processes used in the field. The methodology used is a desk-research method through which a comparison of the literature was conducted. According to the results of the research, we can conclude that there is still no harmonized definition of design thinking. Therefore, due to the rapid development of this field, it is necessary to synthesize the latest research. Also, it is important to provide expert designers with specialized books and tools since the problems are becoming more complex, and designers with competencies required to solve such problems are needed.

Keywords: design methodology; design process; design thinking; problem-solving; user-centered design

1 INTRODUCTION

In today's age, designers are interdisciplinary experts who work in different fields depending on the problem they are solving, and their areas of expertise are no longer exclusively related to design. To address these problems, which are nowadays increasingly complex, they use design thinking methodology. Design thinking (DT) methodology has recently become more popular and the term itself has gained popularization. It is a way of thinking that puts the user at the center and within which innovative solutions are formed as a result of an iterative process. Design and business communities adopt and use DT [1] to improve innovation and gain competitive advantage. Some of today's largest corporations are using this methodology, such as Google, Apple, and Nike [2]. With the increasing development of technology, information, and modern techniques, the DT model is used more and more often in various sectors and industries. To provide insight into the observations related to DT, a cross-section of previous results and current practices will be presented in this paper.

2 RESEARCH AIM AND METHODOLOGY

Through literature analysis on this topic, this paper provides an overview of current research, conclusions, and existing theoretical assumptions to synthesize and compare knowledge about the DT process. Although DT stands out as a process that is for everyone (for non-designers), the focus of this paper is on the DT method used by expert designers. Also, although much of the research and literature is related to DT in the business context, in this paper we will focus on the process of a design thinker independently of the industry in question. Since the DT process itself differs from individual designers, teams, institutions, and companies, it is important that these kinds of reviews are done. In this way, findings are not dispersed, and we can summarize them to reach new conclusions. Although DT is closely related to the purpose and area of problem-solving, it is essential to approach the problem from a design perspective so that we can draw certain conclusions regarding processes and current

reach. Thoring and Mueller [3] point out that if DT wants to achieve scientific field status, then the field must refer to the working mechanisms of the process to explain and improve the way designers think. In this paper we will present a synthesis of research within this area for the following subtopics: (a) historical development of the DT field, (b) approach and features of the DT methodology, and (c) existing DT processes. The aim of this paper is to summarize the knowledge related to DT, review and compare the process and review the current state of the field to gain insight into the possibilities of further development of the area.

In this paper we used a narrative literature review methodology [4] which is a secondary desk-research method based on a comparative analysis of the relevant literature. Books, scientific papers, and articles from the field of design thinking were compared with an emphasis on recent research and basic theoretical background (books relevant to the field). We searched the frequently used and popular databases Google Scholar and Web of Science using keywords design thinking, design education, problem-solving, visual thinking, design process, design methodology, and perception. We examined a total of 52 references spanning from year 1926 to 2021. Out of the mentioned 52 references, 24 (46%) were journal articles (1972-2020), 9 (17%) were conference proceedings articles (1987, 2011-2020), and 19 (37%) were books of prominent authors in the field of DT (1926-2021). We focused primarily on sources in English (47, 90%) and secondarily in Croatian (5, 10%). The examined journal articles are from state-of-the-art journals such as *Design Issues*, *Design Studies*, *Journal of Organization Design*, *Harvard Business Review*, *Journal of Innovation and Entrepreneurship*, *Thinking Skills and Creativity*, *International Journal of Art & Design Education*, etc. We analysed information available in each of the references on the beforementioned subtopics that we defined as the basis for our paper ((a), (b), (c)). When talking about individual tools used in DT, Liedtka [5] concluded that the richness of DT interactions does not lie in individual tools and steps observed in isolation, but in the form that the tools and the whole process form together. For this reason, in this paper we will not focus on individual tools whose detailed descriptions can be found in existing papers [6-8], but on the

overall processes defined in the literature and different ways of their implementation in practice and education.

3 HISTORY OF DESIGN THINKING

Design thinking methodology has long been practiced in different ways and is not something new but has been developed in different areas. Thus, its history dates long before its official beginnings, through the workshops of industrial designers, artists, and architects. These are workshops by art history pioneers such as William Morris, Frank Lloyd Wright, and Ray and Charles Eames whose efforts were to make the world more meaningful. Design thinking arose from a creative process in which different activities alternate and then lead to the final product. Also, it is interesting that design as a discipline did not develop linearly, but expansively [9]. The development of design and design thinking was a logical extension of the work of graphic, industrial, and interaction designers who until then were solving the problems of designing the environment and systems that shape the human experience [10]. Antoljak and Kosović [11] divide the history of DT into four phases of concept development: (1) 1960-1980. - design of intangible products (H. Simon, V. Papanek, H. Rittel), (2) 1980-1990. - research of creative individuals (N. Cross, D. Schon, P. Rowe), (3) 1990-2005. - focus shifts to man (R. Buchanan, D. Kelley, T. Kelley) and (4) 2002-present - use of processes

in the business environment (T. Brown, H. Plattner). DT beginnings are found in the model of the Bauhaus school, which was praised and criticized precisely because of its approach to design education, which was radical and innovative and focused on the unity of art and technology [12]. Therefore, as phase zero, before the 1960s, we can single out the Bauhaus movement (1919-1933), Wallas [13] who represented four phases in the process of creative thinking (preparation, incubation, illumination, verification), and Dewey [14] who influenced later accomplishments related to DT. DT development began in the 1960s [6, 11], [15] which are also called the *Decade of Design Science* [16].

New research has focused on the way designers think [6] and the movement's desire was for the design process to be based on objectivity and rationality [16]. Herbert Simon is one of the pioneers of this way of thinking because he defined the parameters and classification of design in 1969 [11, 16]. In 1971 Papanek [17] pointed out that design must be an inventive tool that responds to people's needs, and McKim [18] defined visual thinking and provided a methodology that describes the process. One of the categories within DT are wicked problems related to solving extremely difficult problems. They are first mentioned in literature by Rittel and Webber [19, 20] and are defined as cultural problems that are complex and embedded in larger problems. Rowe [21] later divided problems into well-defined and poorly defined, and classified wicked problems as poorly defined.

Table 1 Significant design thinking literature through historical phases, 1919-2013

	Year	Author	Publication
Phase 0	1919-1933	Bauhaus school	
	1926	Graham Wallas	The Art of Thought
	1929	John Dewey	The Quest for Certainty: A Study of the Relation of Knowledge and Action
Phase 1	1969	Herbert Simon	The Sciences of the Artificial
	1971	Victor Papanek	Design for the Real World
	1972	Robert McKim	Experiences in Visual Thinking
	1972	Horst Rittel	On the Planning Crisis: Systems Analysis of the 'First and Second Generations'
	1973	Horst Rittel and Melvin Webber	Dilemmas in a General Theory of Planning
Phase 2	1980	Bryan Lawson	How Designers Think
	1982	Nigel Cross	Designerly Ways of Knowing
	1983	Donald Schon	The Reflective Practitioner: How Professionals Think in Action
	1987	Rolf Faste	Perceiving Needs
	1987	Peter Rowe	Design Thinking
Phase 3	1992	Richard Buchanan	Wicked Problems in Design Thinking
	1992	Elizabeth Sanders	Converging Perspectives: Product Development Research for the 1990s
	1996	Yu-Tung Liu	Is Designing One Search or Two? A Model of Design Thinking Involving Symbolism and Connectionism
	2001	Tom Kelley and Jonathan Littman	The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm
Phase 4	2008	Tim Brown	Design Thinking (Harvard Business Review)
	2009	Tim Brown	Change by Design
	2009	Roger Martin	The Design of Business: Why Design Thinking is the Next Competitive Advantage
	2011	Christoph Meinel, Larry Leifer, Hasso Plattner	Design Thinking: Understand – Improve – Apply
	2013	Tom Kelley and David Kelley	Creative Confidence: Unleashing the Creative Potential Within Us All

A significant development in the second phase is the emergence of new journals related to theory and methodology of design, such as *Design Studies* (1979), *Design Issues* (1984), *Research in Engineering Design* (1989), *Design Management Journal* (1989, former series, continues as *Design Management Review*) and many others [16]. During this period, the designer's way of thinking was studied, and researchers came to valuable conclusions related

to the methodology itself. Cross [22] pointed out that methods such as modeling and shaping patterns are used in design, and Schon [23] emphasized that the concept of design has expanded in the last twenty years. Rowe used the term design thinking in the title of his book, and Faste [24] emphasized individual needs and the perception of needs. In the third phase, the DT process begins to be presented to people outside the design profession, and many universities

introduce design courses into their curriculum [11, 15]. The biggest turning point in the DT area came with the founding of the design consulting company IDEO in 1991 which led to the widespread acceptance of the methodology in the last ten years [11, 15]. IDEO founders are David and Tom Kelley who apply DT systematically in IDEO. Buchanan [25] at the time was returning to the idea of wicked problems, while Sanders [26] pointed out that all future products would have to meet customer needs. Liu [27] stated that design is one of the most sophisticated human behaviors. The last phase continues today in which the focus has shifted to creating new ways of thinking with collaborative design and teams [11]. Brown [9] showed a cross-section of the IDEO methodology and defined the DT process. Hasso Plattner founded the Hasso Plattner Institute at Stanford (d.school) in 2005, and d.school in Potsdam, Germany opened shortly afterward [28]. To summarize the most significant

achievements according to the historical phases, Tab. 1 shows the significant literature through the stages of development of this field.

4 DESIGN THINKING APPROACH

There are various interpretations of the concept and ways of using design thinking depending on the meaning attributed to it in theoretical and practical situations. Although there is no harmonized definition of design thinking, common patterns in all definitions are user-centered approach, iterative exploration, prototyping, and teamwork [29]. For this reason, in the following chapters, an overview of different definitions and interpretations of the DT term will be made, as well as the definition of the characteristics, patterns, and mechanics of the procedures themselves.

Table 2 Design thinking definitions used in literature

SOURCE	DEFINITION
Herbert Simon, 1969	"Everyone designs who devises courses of action aimed at changing existing situations into preferred ones." [35]
Victor Papanek, 1971	"Planning and designing any action aimed at a desired and predictable goal is a design process." [17]
Nigel Cross, 1982	"Designerly ways of knowing rest on the manipulation of non-verbal codes in the material culture; these codes translate 'messages' either way between concrete objects and abstract requirements...." [22]
Richard Buchanan, 1992	"The new liberal art of design thinking ... points toward the impossibility of relying on any one of the sciences for adequate solutions to what are the inherently wicked problems of design thinking." [25]
Tim Brown, 2008	"Design thinking is a discipline that uses the designer's sensibility and methods to match people's needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity." [32]
Hasso Plattner, 2011	"The method of Design Thinking melds an end-user focus with multidisciplinary collaboration and iterative improvement and is a powerful tool for achieving desirable, user-friendly, and economically viable design solutions and innovative products and services." [28]
Tom Kelley and David Kelley, 2013	"Design thinking is a way of finding human needs and creating new solutions using the tools and mindsets of design practitioners." [36]
Antoljak and Kosović, 2018	"Design thinking is a way of solving problems with the help of a structured process and a set of methods and techniques that originally originated from designers." [11]
Richard Buchanan, 2019	"Design is the transformation of surroundings into environments for human experience." [10]
Michael Lewrick, Patrick Link and Larry Leifer, 2020	"Design thinking makes it possible for us to realize solutions that meet the wishes of the customers, solve a genuine problem, and thus provide value for the customer." [38]
Jeanne Liedtka, Karen Hold and Jessica Eldridge, 2021	"DT's human-centered focus engages us by giving us experiences that matter, that impact us at a deep personal level." [39]

4.1 Definition

The development of the DT theory depends on the development of human thinking in the field of cognitive sciences, artificial intelligence, and neuroscience [27]. Therefore, DT is comprehensive, interdisciplinary, and difficult to define unambiguously and simply. That is why there is no final definition of design thinking. Thinking in DT is a way of thinking that is specific in solving challenges through multidisciplinary. The term DT is often used as a unique approach to problem-solving in an innovative way [30] and is a link between process, product, and end-user [31]. This puts the focus on the user who is involved in all phases of the process. We can say that DT is a user-oriented approach to design in which innovation is achieved by direct observation of what people want and need [32]. Through research, testing, prototyping, and achieving empathy with users, they become the focus of every step and part of every decision. In this way, DT focuses on the solution, not the problem. In today's world, DT is a response to the complexity of modern technology [31] and can be viewed as

a collaborative approach [33]. A comparison of different definitions that are prominent in literature is available in Tab. 2. Although it is evident from various examples from practice and scientific literature that DT is a complex field and a methodology that brings innovative results, it is important to emphasize that certain authors do not agree with this concept. Some authors consider DT to be a useless process, which they call a new myth and a failed experiment [11]. Also, some authors argue that DT research has not produced a definitive account of DT [34]. Most of the criticism is related to the fact that the DT methodology is fluid and ever-changing. Because of this some authors claim that DT does not follow a methodology and that it neglects constraints that designers need [44].

4.2 Characteristics and Thinking Patterns

Dorst [37] points out that different creative activities in popular literature are often classified as design thinking, and for this reason, many researchers in the design profession are opposed to the DT approach. For this reason, it is important

to look at DT from the perspective of defining the process because it has specific and intentional ways of thinking [37]. Cross [22] identifies five aspects of design modes of knowledge: designers deal with poorly defined problems, they are solution-oriented, think constructively, use codes that translate abstract requirements into concrete objects, and use codes to read and write object languages. The most important DT principle is human-centered design [2, 9, 15] used as an inventive tool that meets the real needs of people [17]. With understanding human behavior, what users want and need, a solution is reached that is appropriate for the user. In addition to empathy and focus on the user, DT is also characterized by the culture of prototyping. With constant and frequent prototyping, ideas materialize, and solutions are found faster through iteration. Users are also involved in this process because by testing the prototype with users, insights that are important for solving the problem are gained. Solution generation, synthesis, and evaluation are prominent DT features. It is important to know how to determine the level of compromise when choosing a solution since in design there is no single solution. Experimentation is imperative in research because ideas can be explored through conceptual development [26].

Innovation as a feature of design thinking is achieved with user focus, creativity, and teamwork. As a result of teamwork, different profiles of people of various professions and skills collaborate. Team interdisciplinarity is important because of interdisciplinary nature of the DT process itself, which solves problems from different spheres of life. Thus, Papanek points out that most discoveries are made on the border of different techniques or disciplines [17]. Through DT we observe the human, business, and technological aspects of a problem [31]. We can look at DT as a process of a series of steps taken to solve problems and as a way of thinking about problems [1]. Most cognitive design theories point out that the dominant design model is cyclical and iterative [40]. This means that the DT process is not straightforward, but the steps can be repeated and can be returned to multiple times. DT is a nonlinear process in which knowledge can be obtained during different phases that can influence certain decisions [9, 41]. DT deals with solving complex problems so it is essential to define the logical thinking that occurs in such a procedure. Problem-solving can be described as a search within a maze of possibilities in which the maze represents the environment [35]. Successful problem-solving involves selectively searching the maze and reducing the maze to manageable proportions [35]. Given the fact that DT differs from traditional ways of solving problems, it is not surprising that it also differs in the way of logical thinking. Central practice for design disciplines is the framework which is the general implication that we will create a certain value by applying a working principle [37]. We can define key patterns of conclusion in design with models of formal logic [37]. Thus, logical ways of reasoning used to solve a problem are deduction, induction, and abduction [37]. The most used are the deductive and inductive way of logical thinking, and the less used is the abductive way, which is what is used in the DT process [42]. Logical reasoning in productive thinking is an abduction

process that can be carried out in two ways [37]. Abduction-1 is associated with conventional problem-solving and is a closed problem-solving process, while abduction-2 is associated with conceptual design and is an open problem-solving process, used in the DT process [37].

4.3 Procedure Mechanics

Three key strategic DT approaches are taking a broad systemic approach to problem-solving, distinctively framing the problem, and designing according to the first principles [43]. Thoring and Mueller [3] point out that three aspects of evolutionary creativity (generating, selecting, and retaining ideas) can explain most DT principles. The process begins with the identification and formulation of problems and constraints. Solving the set problem requires that the target space and the solution space overlap in such a way that they form an optimal solution between the two spaces [44]. The two main categories of productive thinking are convergence and divergence [45]. Divergent thinking is the ability to create different solutions and refers to the generation of ideas and the creation of choices, while convergent is the choice of one final solution from those previously defined [31], [46]. Design thinking encourages divergent thinking, which is then followed by convergent thinking [31]. Analysis and synthesis are complements to the divergent and convergent way of thinking [9]. It is these aspects that are the foundation of design thinking which is a continuous movement between divergent and convergent processes on the one hand, and analytical and synthetic on the other [9]. Stempfle and Badke-Schaub [44] propose generation, exploration, comparison, and selection as four basic cognitive operations for working with problem spaces. The first two expand the problem space, and the other two narrow it. Thoring and Mueller [3] present a typology of knowledge in which all patterns that enable decisions are defined as knowledge. Thus, they emphasize the typology of design knowledge with the following levels: physical (A), neural (B), symbolic (C), and model (D). The physical level refers to three-dimensional forms, the neural level to tacit knowledge, the symbolic level to explicit knowledge, and the model level to models and theories [3]. These four levels build on top of each other, and each level includes aspects of the previous level [3]. Buchanan [25] emphasizes the peculiar nature of design with the claim that design problems are wicked precisely because they do not have their subject of interest (one area), but the area of design interest is what the designer decides. Thus, the subject of interest is universal in scope because DT can be applied to any area of human experience [25].

5 DESIGN THINKING PROCESSES

The fluidity and flexibility of DT are interesting aspects of the process, but sometimes they can be a factor of misunderstanding. Given the use in various industries by different institutions and organizations, the process itself is not always the same. Various organizations have different DT approaches and the way they approach challenges. Because the process is not straightforward, it can often seem

chaotic and out of order, and it takes time to get used to this way of thinking [32]. However, to be able to methodologically define this technique as well as to conduct further research, it is necessary to set certain process definitions. Simon [35] points out that the design process can be seen as a process of generating different solutions and that it does not have to be just one test cycle but a multitude of cycles. In this chapter, we will compare some of the most used DT processes in education and the business world and show their similarities and differences. All processes differ in the way they are interpreted and in the number of stages, but they also have a lot in common. What they all have in common is a convergent and divergent way of thinking, analysis and synthesis, and the iterative nature of the process related to returning to certain phases or repeating them. Also, in all processes, the idea is to learn through failure. To explain and understand certain phases, processes are defined linearly, but they are never linear. All processes can be repeated several times and we can return to certain phases within the process. What is also common to all processes is

prototyping and obtaining feedback from users through testing [1]. Trošić [2] presents different organizational models in higher education that provide an educational context for DT such as Darden School of Business, Royal College of Art, University of Technology Sydney, Delft University of Technology, School of Economics and Business (Ljubljana), Hasso Plattner Institute at Stanford (d.school), in Potsdam, Cape Town, Malaysia, and Paris. Brown [9] believes that the most important opportunity for long-term impact of DT lies in education and emphasizes the importance of shaping education in a way that encourages experimentation. Srhoj and Morić Milovanović [6] divide the design process into micro and macro processes. Thus, Darden School of Business, Rotman School of Management, d.school, and IDEO are classified as those that use micro processes, while the University of St. Gallen is classified as using a macro process [6]. In recent years, DT has been developed in various industrial workshops led by leaders who are often non-design professionals [47]. Some of the prominent DT processes are shown in Tab. 3.

Table 3 Overview of the most used design thinking processes

Process	Approach	Phases
Micro process	Double Diamond design model, British Design Council	discover, define, develop, deliver
	University of Virginia Darden School of Business, USA	What is?, What if?, What wows?, What works?
	Hasso Plattner Institute of Design (HPI) at Stanford, USA (d.school)	empathize, define, ideate, prototype, test
	Hasso Plattner Institute (HPI) in Potsdam, Germany	understand, observe, point of view, ideate, prototype, test
	IDEO (Tim Brown)	inspiration, ideation, implementation
	IDEO (Tom Kelley and Jonathan Littman)	understand, observe, visualize, evaluate, implement
Macro process	Rotman School of Management, Toronto, Canada	initiation, investigation, integration, implementation
	University of St. Gallen, Switzerland	diverging (design space exploration, critical function prototype, dark horse prototype, funky prototype), converging (functional prototype, X prototype, final prototype)

Antoljak and Kosović [11] state that most DT approaches have a common fundamental "double diamond" process. The basic phases of "double diamond" are discover, define, develop, and deliver [11]. Through this process, divergent (discover, develop) and convergent (define, deliver) processes alternate. In the first step the problem is identified, in the second a deep understanding of the problem is gained, in the third ideas are developed, and in the last step, a prototype and final solution are created. The process of the Darden School of Business is based on four steps: "What is?", "What if?", "What wows?", and "What works?" [11]. In the step "What is?" the problems that users have are defined, in the step "What if?" concepts are developed and opportunities are recognized, in step "What wows?" the most promising one is selected from the multitude of concepts, and in the last phase "What works?" prototypes are shaped and implementation is carried out [6, 11].

The Hasso Plattner Institute (HPI) at Stanford and in Potsdam are representative examples of a typical educational DT institution and their processes are also the most commonly used [12]. HPI at Stanford (d.school) divides the design process into five steps: empathize, define, ideate, prototype, and test [41]. With empathy, designers understand users and their actions [31]. In the definition phase, the collected information is processed, and the challenge is defined [11]. In the ideation phase, rough ideas are

developed, and many ideas are devised, while in the prototyping phase, a functional model that helps to verify the design is created [11, 31]. The last step is testing in real conditions that can be carried out at all stages of the process and the purpose is to get feedback based on the prototype [11]. HPI d.school in Potsdam uses a similar six-step model: understand, observe, point of view, ideate, prototype, and test [3]. In phase understand the problem is defined, while insights into the user needs are gathered through observation [48]. By point of view, a micro theory of user needs is created [48]. During ideation, questions that address the defined problems and ideas for possible solutions are generated [48]. Through prototyping, a prototype is prepared, and through testing, feedback is collected from users and participants [48]. Tim Brown [32] points out that the DT process can best be described metaphorically as a system of spaces rather than as a predefined series of steps. Nevertheless, to define this process, Brown divides the DT process into three phases: inspiration, ideation, and implementation, and emphasizes that these steps overlap and are not sequential [9, 32]. The inspiration phase represents the problem that motivates the search for solutions and in which insights are gathered [9], [32]. Ideation is the process of generating, developing, and testing ideas, and the last stage is implementation, which is the path to the market [9, 32]. Through all three spaces, we can return to refine ideas and determine new directions [32].

Brown [9] also presents a model of constraints across three elements: feasibility, sustainability, and desirability. These are overlapping elements that the design thinker needs to balance [9]. Feasibility refers to what is functionally possible in the near future, sustainability to what is realistic to become part of a sustainable business model and desirability to what makes sense to people [9]. Kelley and Littman [49] presented the IDEO company methodology in five basic steps: understand, observe, visualize, evaluate, and implement. They pointed out that the interpretation of the method depends on the nature of the task to be solved. In the first step, the market, client, and technology are understood, in the observation step it is defined how users think, through visualization new concepts are visualized, in evaluation prototypes are evaluated through iterations, and in the last step, implementation is carried out [49]. Rotman School of Management divides the process into four phases: initiation, investigation, integration, and implementation.

As a macro process, we can highlight the University of St. Gallen which developed a process divided into two phases (divergent and convergent) within which micro processes related to prototypes are carried out [6]. The divergent phase consists of the design space exploration, critical function prototype, dark horse prototype, and funky prototype, while the convergent phase consists of a functional prototype, X prototype, and the final prototype [6]. Within the first step, the design challenge is explored, the first prototypes are developed in the critical function prototype, a prototype with neglected constraints is formed in the dark horse prototype, the best ideas are used in the funky prototype, functional prototype is shaped according to final requirements, and the X prototype defines key functionalities that meet customer needs [6]. Within each of the phases, a micro iteration process is performed [6]. As a result of the intersection of these processes of different organizations and institutions, we can see that all processes largely coincide with the already mentioned stages of the creative thinking process (preparation, incubation, illumination, verification) presented by Wallas [13] in 1926, which shows that DT evolved from this process and has been further expanded and changed since.

6 DESIGNER COMPETENCIES

Although DT often stands out as something that anyone can use, it is important to make a distinction between beginners and experts [50]. In this paper, we deal with the competencies of expert designers within this field. Creativity and artistic expression are only part of the competencies of designers whose work is often equated with the above-mentioned. The primary task of the designer is solving problems and the designer must have a sensibility in discovering existing problems [17]. DT skills enable readiness to face problems, thinking outside the box, and creating innovative solutions [50]. IDEO [8] singles out seven essential competencies of design thinkers: empathy, optimism, iteration, creative trust, creation, acceptance of ambiguity, and learning from failure. On the other hand, Brown [32] singles out empathy, integrative thinking,

optimism, experimentation, and collaboration as essential characteristics of any design thinker. Razzouk and Shute [50] highlight focus on man and the environment, the ability to visualize, a predisposition to multifunctionality, a systematic vision, and an affinity for teamwork. Designers possess highly developed skills in graphic communication [45] and competencies of intuitiveness, pattern recognition, and expression in various media [9]. Some of the characteristics of designers are flexibility, quick thinking, and spontaneity because they often must quickly assess the conditions of a situation [26]. As a result of integrative thinking in which all factors are considered, designers look at all aspects of a particular problem and achieve innovative results [32].

While DT presents itself as a process for everyone, not just designers, that doesn't mean anyone can come up with innovative solutions with equal expertise. That is why it is important to practice interdisciplinarity at the level of the team itself. Through DT, designers emphasize their abilities, and it is interdisciplinarity that connects the various aspects that make up design thinkers. Therefore, designers should be T-shaped persons, which means that they are experts in one area (vertical bar on the letter T), but also that they have a broad knowledge of similar fields and strong communication skills (horizontal bar on the letter T) [46]. T-shaped persons can recombine their knowledge better than other professionals, which leads to the generation of many ideas and enables better transfer of knowledge [3]. Expert strategies most often involve a top-down approach [50]. Thus, DT uses integrated teams composed of members of different professions and fields [9]. In this way, everyone contributes to the process and the final solution. Teams play an important role in the organizational context because groups can solve problems that individuals cannot [44]. The ideal team consists of five to six people of different backgrounds [46]. In this way, teams consist of individuals from different professions and the team becomes multidisciplinary. For this reason, T-shaped persons are needed within the team because then the members overlap in certain areas, but diversity is still achieved.

7 SIGNIFICANT CONCLUSIONS OF EXISTING RESEARCH

In recent years, much research has been conducted [3-6, 12, 29, 44, 51, 52] on the topic of design thinking and the way designers think and come up with solutions. We singled out and reviewed the mentioned relevant empirical research results and compared them with the DT theoretical settings presented so far. For example, Liedtka [5] conducted research in 22 organizations to observe the use of DT elements in practice. The results showed that teams try to achieve a qualitative understanding of the user context, team heterogeneity is carried out, multiple solutions are created, and a structured process is used [5]. Also, research showed that teams first conduct user research and only then start preparing solutions [5], which coincides with the theories of processes of Brown, HPI schools, and others presented in previous chapters. It often happens that during research teams notice that the problem statement needs to be reformulated and they return to the beginning [5]. This is

consistent with the iterative nature of design thinking that is common to all theoretical models. Aflatoony et al. [51] conducted a study evaluating students' thinking skills and their experience during a DT class in two high schools. What the results showed in terms of students' thinking skills was that students showed different levels of understanding of the design process and that they all improved their DT skills by the end of the course [51]. From this, we can conclude that DT is a skill that can be learned and needs to be practiced. A problem with teamwork in this research was leadership, even division of tasks, and team size [51]. This is very significant because the team entity is very important for the DT process. Pap et al. [31] conducted workshops with twelve students of different fields who used the DT method during the workshop. They concluded that mixing disciplines works well, that placing students in a DT environment encourages their creativity regardless of prior knowledge, and that working in the right place can create new value [31]. These conclusions coincide with citations from literature that cite teamwork and interdisciplinarity as key characteristics of design thinking. Thoring et al. [12] compared Bauhaus and d.school education and observed many similarities between the two modes of education, although they are separated by almost a century [12]. Thus, the greatest similarities are in the way of thinking, culture, environment, and conditions for innovation [12]. Neither approach use evaluation, they encourage experimentation and creativity, strive for radical innovation, and want to improve the world [12]. Costa Valentim et al. [52] conducted an empirical study with postdoctoral students to examine their perception of DT. Results showed that students had difficulties in using DT and in team interaction [52]. We can conclude that there are still some difficulties that need to be addressed in the process. Goldschmidt and Rodgers [29] compared the DT process of different groups of designers. From the analysis of the results, they concluded that the respondents did not follow a linear process but iterated between activities [29]. This fact confirms all previous models that emphasize iteration as an important feature of design thinking.

8 CONCLUSIONS

According to the results of the research, we can conclude that there is still no harmonized DT definition or strictly defined processes and phases used in the methodology. In this paper we provided a detailed list of DT literature through historic phases in Tab. 1, overview of DT definitions in Tab. 2, and synthesis of DT processes in Tab. 3. As DT continues to develop expansively and expands into different areas, we assume that definitions and phases will be increasingly dispersed and broadened as more organizations and institutions begin to implement it in their business and education. Therefore, due to the rapid growth of the field, it is necessary to do syntheses of the latest research. The analysis shows that DT courses are taught at many universities around the world. For further research, we propose an overview of the implementation of DT courses at Croatian universities and schools, examining the level and manner of using this methodology in Croatian companies and

comparison with foreign teams in which there is a DT education at the academic level. The first DT Croatian book focuses on the business side of solving problems and approaching the topic in a way that non-designers can understand it. Nevertheless, it is important to provide expert designers with structured processes and current research which was the goal and purpose of this review. In the latest edition of his book *Change by Design*, Tim Brown makes suggestions for further directions and development of design thinking. He cites redesign of outdated social systems, revival of participatory democracy, urban design, the humanization of artificial intelligence, biotechnology, and the circular economy. This instructs us that problems are becoming more wicked and therefore we need designers who have the necessary competencies and skills to solve such problems. The approach in which this is possible is design thinking. Therefore, we need to give designers a good background through education because they are the ones who know very well what the limitations are and where the boundaries lie. Even better than that they also know how, when, and how to move, expand, and break those boundaries. To be able to do so on an ever-increasing and complex level, they need to be supported in all aspects so that they can further develop this area.

9 REFERENCES

- [1] Dunne, D. (2018). Implementing design thinking in organizations: an exploratory study. *Journal of Organization Design*, 7(16), 1-16. <https://doi.org/10.1186/s41469-018-0040-7>
- [2] Trošić, D. (2019). Design thinking approach in business education. *Obrazovanje za poduzetništvo - E4E*, 9(1), 111-131. Retrieved from <https://hrcak.srce.hr/221254>
- [3] Thoring, K. & Mueller, R. M. (2011). Creating Knowledge in Design Thinking: The Relationship of Process Steps and Knowledge Types. *Diversity and Unity: Proceedings of IASDR 2011, the 4th World Conference on Design Research*, 1-9. Retrieved from <https://bit.ly/3n3ZSle>
- [4] Xiao, Y. & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 39(1), 93-112. <https://doi.org/10.1177%2F0739456X17723971>
- [5] Liedtka, J. (2017). Evaluating the Impact of Design Thinking in Action. *Academy of Management Proceedings*, 1-6. <https://doi.org/10.5465/AMBPP.2017.177>
- [6] Srhoj, S. & Morić Milovanović, B. (2016). Dizajn razmišljanje kao suvremeni pristup rješavanju poslovnih problema. *Zbornik Ekonomskog fakulteta u Zagrebu*, 14(2), 63-91. Retrieved from <https://hrcak.srce.hr/170203> (in Croatian)
- [7] Lewrick, M., Link, P., & Leifer, L. (2018). *The Design Thinking Playbook: Mindful Digital Transformation of Teams, Products, Services, Businesses and Ecosystems*. Hoboken: Wiley.
- [8] IDEO.org. (2015). *The Field Guide to Human-Centered Design*. San Francisco: IDEO.org.
- [9] Brown, T. (2019). *Change by Design, Revised and Updated: How Design Thinking Transforms Organizations and Inspires Innovation*. New York: Harper Business.
- [10] Buchanan, R. (2019). Systems Thinking and Design Thinking: The Search for Principles in the World We Are Making. *She Ji: The Journal of Design, Economics, and Innovation*, 5(2), 85-104. <https://doi.org/10.1016/j.sheji.2019.04.001>

- [11] Antoljak, V. & Kosović, M. (2018). *Design thinking za nedizajnere – kako riješiti poslovne probleme i uspješno inovirati*. Zagreb: Školska knjiga. (in Croatian)
- [12] Thoring, K., Mueller, R. M., Giegler, S., & Badke-Schaub, P. (2020). From Bauhaus to Design Thinking and Beyond: A Comparison of Two Design Educational Schools. *Proceedings of the Design Society: DESIGN Conference, 1815-1824*. Retrieved from <https://bit.ly/3tdNH6w>
- [13] Wallas, G. (1926). *The Art of Thought*. New York: Harcourt, Brace & Company.
- [14] Dewey, J. (1929). *The Quest for Certainty: A Study of the Relation of Knowledge and Action*. New York: Minton, Balch & Company.
- [15] Sabolović-Krajina, D. (2020). Design thinking kao poslovni koncept i alat u upravljanju knjižnicama. *Vjesnik bibliotekara Hrvatske*, 63(1-2), 65-82. Retrieved from <https://hrck.srce.hr/252864> (in Croatian)
- [16] Cross, N. (2001). Designly Ways of Knowing: Design Discipline versus Design Science. *Design Issues*, 17(3), 49-55. <https://doi.org/10.1162/074793601750357196>
- [17] Papanek, V. (1973). *Dizajn za stvarni svijet*. Split: Nakladni zavod Marko Marulić. (in Croatian)
- [18] McKim, R. H. (1980). *Experiences in Visual Thinking*. Monterey: Brooks-Cole Publishing.
- [19] Rittel, H. (1972). On the Planning Crisis: Systems Analysis of the 'First and Second Generations'. *Bedriftskonomien*, 8, 390-396. Retrieved from <https://bit.ly/2Ytz9Ey>
- [20] Rittel, H. W. & Webber, M. M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, 4(2), 155-169. <https://doi.org/10.1007/BF01405730>
- [21] Rowe, P. G. (1987). *Design Thinking*. Cambridge: MIT Press.
- [22] Cross, N. (1982). Designly ways of knowing. *Design Studies*, 3(4), 221-227. [https://doi.org/10.1016/0142-694X\(82\)90040-0](https://doi.org/10.1016/0142-694X(82)90040-0)
- [23] Schon, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- [24] Faste, R. A. (1987). Perceiving Needs. *SAE Future Transportation Technology Conference and Exposition*, 1-8. <https://doi.org/10.4271/871534>
- [25] Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues*, 8(2), 5-21. <https://doi.org/10.2307/1511637>
- [26] Sanders, E. (1992). Converging Perspectives: Product Development Research for the 1990s. *Design Management Journal (Former Series)*, 3(4), 49-54. <https://doi.org/10.1111/j.1948-7169.1992.tb00604.x>
- [27] Liu, Y.-T. (1996). Is designing one search or two? A model of design thinking involving symbolism and connectionism. *Design Studies*, 17(4), 435-449. [https://doi.org/10.1016/S0142-694X\(96\)00018-X](https://doi.org/10.1016/S0142-694X(96)00018-X)
- [28] Plattner, H., Meinel, C., & Leifer, L. (2011). *Design Thinking: Understand – Improve – Apply*. Berlin: Springer.
- [29] Goldschmidt, G., & Rodgers, P. A. (2013). The design thinking approaches of three different groups of designers based on self-reports. *Design Studies*, 34(4), 454-471. <https://doi.org/10.1016/j.destud.2013.01.004>
- [30] Pusca, D. & Northwood, D. (2018). Design thinking and its application to problem solving. *Global Journal of Engineering Education*, 20(1), 48-53. Retrieved from <https://bit.ly/38KfHoJ>
- [31] Pap, M., Vidović, R., & Baletić, B. (2019). Design Thinking metoda u znanstvenom istraživanju, edukaciji i poslovnoj praksi. *Prostor*, 27(2 (58)), 334-347. (in Croatian) [https://doi.org/10.31522/p.27.2\(58\).12](https://doi.org/10.31522/p.27.2(58).12)
- [32] Brown, T. (2008). Design Thinking. *Harvard Business Review*, 86(6), 84-92. Retrieved from <https://bit.ly/3yQsaBY>
- [33] Linton, G. & Kinton, M. (2019). University entrepreneurship education: a design thinking approach to learning. *Journal of Innovation and Entrepreneurship*, 8(3), 1-11. <https://doi.org/10.1186/s13731-018-0098-z>
- [34] Ghassan, A. (2016). A corpus-led study into how 'design' is represented in design thinking research. *Proceedings of the DESIGN 2016 14th International Design Conference*, 11-22.
- [35] Simon, H. A. (1996). *The Sciences of the Artificial*. Cambridge: MIT Press.
- [36] Kelley, T. & Kelley, D. (2013). *Creative Confidence: Unleashing the Creative Potential Within Us All*. New York: Crown Business.
- [37] Dorst, K. (2011). The core of 'design thinking' and its application. *Design Studies*, 32(6), 521-532. <https://doi.org/10.1016/j.destud.2011.07.006>
- [38] Lewrick, M., Link, P., & Leifer, L. (2020). *The Design Thinking Toolbox*. Hoboken: Wiley.
- [39] Liedtka, J., Hold, K., & Eldridge, J. (2021). *Experiencing Design: The Innovator's Journey*. New York: Columbia Business School Publishing.
- [40] Oxman, R. (2017). Thinking difference: Theories and models of parametric design thinking. *Design Studies*, 52, 4-39. <https://doi.org/10.1016/j.destud.2017.06.001>
- [41] Henriksen, D., Richardson, C., & Mehta, R. (2017). Design Thinking: A Creative Approach to Educational Problems of Practice. *Thinking Skills and Creativity*, 26, 140-153. <https://doi.org/10.1016/j.tsc.2017.10.001>
- [42] Martin, R. L. (2009). *The Design of Business: Why Design Thinking is the Next Competitive Advantage*. Boston: Harvard Business Review Press.
- [43] Cross, N. (2011). *Design Thinking: Understanding How Designers Think and Work*. New York: Berg.
- [44] Stempfle, J. & Badke-Schaub, P. (2002). Thinking in design teams - an analysis of team communication. *Design Studies*, 23(5), 473-496. [https://doi.org/10.1016/S0142-694X\(02\)00004-2](https://doi.org/10.1016/S0142-694X(02)00004-2)
- [45] Lawson, B. (2005). *How Designers Think*. New York: Architectural Press.
- [46] Thoring, K. & Mueller, R. M. (2011). Understanding the Creative Mechanisms of Design Thinking: An Evolutionary Approach. *DESIRE '11: Proceedings of the Second Conference on Creativity and Innovation in Design*, 137-147. <https://doi.org/10.1145/2079216.2079236>
- [47] Mosely, G., Wright, N., & Wrigley, C. (2018). Facilitating design thinking: A comparison of design expertise. *Thinking Skills and Creativity*, 27, 177-189. <https://doi.org/10.1016/j.tsc.2018.02.004>
- [48] Thoring, K. & Mueller, R. M. (2011). Understanding design thinking: A process model based on method engineering. *International Conference on Engineering and Product Design Education*, 493-498. Retrieved from <https://bit.ly/2WTH4Ki>
- [49] Kelley, T. & Littman, J. (2001). *The Art of Innovation: Lessons in Creativity from IDEO, America's Leading Design Firm*. New York: Currency.
- [50] Razzouk, R. & Shute, V. (2012). What Is Design Thinking and Why Is It Important? *Review of Educational Research*, 82(3), 330-348. <https://doi.org/10.3102/0034654312457429>
- [51] Aflatoony, L., Wakkary, R., & Neustaedter, C. (2017). Becoming a Design Thinker: Assessing the Learning Process of Students in a Secondary Level Design Thinking Course. *International Journal of Art & Design Education*, 37(3), 438-453. <https://doi.org/10.1111/jade.12139>
- [52] Costa Valentim, N. M., Silva, W., & Conte, T. (2017). The Students' Perspectives on Applying Design Thinking for the Design of Mobile. *2017 IEEE/ACM 39th International*

*Conference on Software Engineering: Software Engineering
and Education Track (ICSE-SEET), 77-86.*
<https://doi.org/10.1109/ICSE-SEET.2017.10>

Authors' contacts:

Ana Svalina, PhD student
Faculty of Graphic Arts, University of Zagreb,
Getaldićeva 2, 10000 Zagreb, Croatia
asvalina@grf.hr

Mario Tomiša, PhD, Full Professor
University North,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
mario.tomisa@unin.hr

Marko Čačić, PhD, Teaching Assistant
University North,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
marko.cacic@unin.hr

Krunoslav Hajdek, PhD, Associate Professor
(Corresponding author)
University North,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
krunoslav.hajdek@unin.hr

Hybrid of "Intersection" Algorithm for Multi-Objective Optimization with Response Surface Methodology and its Application

Maosheng Zheng*, Yi Wang, Haipeng Teng

Abstract: Recently, a new "intersection" method for multi-objective optimization was developed in the points of view set theory and probability theory, which introduces a new idea of favorable probability to reflect the favorable degree of the utility of performance indicator in multi-objective optimization, and the product of all partial favorable probabilities of entire utilities of performance indicators makes the overall / total favorable probability of the candidate. Here, in this paper, the new "intersection" algorithm for multi-objective optimization is combined effectively with response surface methodology (RSM) by taking each response as one objective, which transfers the multi-response optimization problem into a single response one with the help of the overall / total favorable probability of each scheme. The overall / total favorable probability is the uniquely decisive index of the scheme in the optimization. Applications of the hybrid approach with two examples in material technology are given, proper predictions are obtained.

Keywords: favorable probability; "intersection" method; hybrid; multi-object optimization; response surface methodology

1 INTRODUCTION

Experimentation is an indispensable part in scientific research and technical development [1-3]. A well-designed experiment is very powerful route to get informative achievement. In general, a well-designed experiment is of representative in limited number of tests. The special technique called "experimental design" aims to effectively reach to the goal. Excellent experimental design could give more effective consequences with limited time and cost.

Response surface methodology (RSM) is a collection of statistical and mathematical technique, which is useful for optimizing processes. It has been widely used in the design, development, and formulation of new products, as well as improvement of designs of existing product [2].

The most extensive applications of RSM are in the industrial world, particularly in situations where multiple input variables potentially affect quality characteristics or performance of product or process. These quality characteristics or performance of product or process are usually called the *response* in the treatment. In most actual cases, the RSM faces more than one response, i.e., multiple response problems. The input variables are subjected to adjust to appropriate status for purposes of optimizing the *responses* to their ideal position.

While, as to optimizing the *responses* to their ideal position, it is uneasy problem since several responses might be with competing relationship sometimes. So, a compromised technique is usually employed in the treatment. Till now, some algorithms has been employed as compromised technique for multi-objective optimizations [1], such as the so called "comprehensive balance method", "comprehensive scoring method", "grey relational analysis" [4], signal to noise ratio analysis [5], or signal to noise ratio (SNR) together with the technique for order preference by similarity to ideal solution (TOPSIS) [6], Pareto front optimization [2], etc. However, these algorithms involve many inherent defects, such as, personal factors, setting the beneficial index and unbeneficial index at unequal positions. So, these algorithms cannot be considered as full quantitative, but empirical ones instead. Particularly, the "additive"

algorithm is frequently used in the compromised treatments of the previous multi-objective optimizations, which diverges from the spirit of "simultaneous optimization" of multiple responses in the viewpoint of probability theory [7].

Recently, a new "intersection" method for multi-objective optimization was developed in the points of view set theory and probability theory [7], which aims to solve the inherent problems of personal factors and diverging from the spirit of "simultaneous optimization" of multiple responses in previous multi-objective optimization. A new concept of favorable probability was proposed to reflect the favorable degree of the candidate in the optimization, each performance utility indicator of the candidate contributes to one partial favorable probability quantitatively, and the overall / total favorable probability of a candidate is the product of all partial favorable probabilities in the viewpoint of probability theory; the total favorable probability is the uniquely decisive index in the competitive selection process, which thus transfers the multi-objective optimization problem into a single objective one.

As a further development to the newly proposed "intersection" method for multi-objective optimization, here in this paper, the new "intersection" algorithm for multi-objective optimization is combined hybridly with response surface methodology (RSM) to extend the experimental design and get an impersonal prediction. Applications of the hybrid approach with two examples in material technology are given to show the procedure of the approach.

2 COMBINATION OF THE "INTERSECTION" METHOD FOR MULTI-OBJECTIVE OPTIMIZATION ALGORITHM WITH RESPONSE SURFACE METHODOLOGY

2.1 Transformation of Multi-response Optimization Problem into a Single Response One

As to simultaneous optimization of multiple responses, several responses are involved. Lets' take each response of the scheme as one objective of new "intersection" method for multi-objective optimization [7]. Then some utilities of response might have the characteristics of "the higher the

better", i.e., beneficial type, but other utilities of response might have the characteristics of "the lower the better", i.e., unbeneficial type. Thus, each utility of response contributes to one partial favorable probability in linear manner according to its type individually [7].

Furthermore, as the scheme is an integral body of both beneficial and unbeneficial indicators of utility, the overall / total favorable probability of a scheme can be obtained by the product of all partial favorable probabilities according to probability theory for simultaneous optimization of multi-response. Thus, the overall / total favorable probability transfers the multi-response optimization problem into a single response one, which is the uniquely decisive index of the scheme in the optimization.

2.2 Regression Analysis

Moreover, regression analysis is conducted for the overall / total favorable probabilities of all schemes of the designed experiment to get a regressed function of the overall / total favorable probability. Then, the maximum value of the overall / total favorable probability and values of corresponding specific input variables are gained by common algorithm of mathematics. The next step is to regress each response to get its regressed function, and then substitute the values of corresponding specific input variables into each regressed function of the response to get its compromised result.

By far, the procedure of hybrid of "Intersection" algorithm for multi-objective optimization with response surface methodology is developed.

3 EXAMPLES OF APPLICATIONS

3.1 Optimal Design of PP / EPDM / GnPs / GF Composites

Niyaraki, et al employed RSM to optimize the mechanical properties of impact strength and elastic modulus of polypropylene (PP) / ethylene propylene dine monomer (EPDM) / grapheme Nano sheets (GnPs) / glass fiber (GF) hybrid nano-composites with Box-Behnken method of RSM [8]. The impact strength and elastic modulus are the two responses of the optimization of the nano-composite.

In literature [8], there are three levels as the input variables for the three parameters, i.e., EPDM (5, 10 and 15 wt.%), GnPs (0, 1 and 2 wt.%) and glass fiber (10, 20 and 30 wt.%). It was discovered that GnPs, glass fiber and EPDM played important roles in the impact strength and elastic modulus of the nano-composites.

Here the hybrid of "Intersection" algorithm for multi-objective optimization with response surface methodology is utilized to deal with this problem.

Tab. 1 shows the experimental results of impact strength and elastic modulus, together with their partial favorable probabilities and total probabilities; both impact strength and elastic modulus are beneficial performance utility indicators in the hybrid approach [7].

From Tab. 1, it can be seen that the scheme No. 12 exhibits the maximum of the overall / total favorable probability P_i , so it could be primarily chosen as one of the

optimal scheme of the multi-response RSM at the first glance directly.

Furthermore, regression analysis is conducted to the total favorable probability so as to get more accurate prediction. The regressed function for the total favorable probability is,

$$\begin{aligned} f_{P_i} \times 10^3 &= 0.5751 + 259.6483X_1 + 13.0371X_3 - \\ &- 5592.7607X_1^2 + 46.5832X_3^2 - 476.762X_1X_2 + \\ &+ 7.0492X_2X_3 - 167.1209X_1X_3, \\ R^2 &= 0.9931. \end{aligned} \quad (1)$$

The function $f_{P_i} \times 10^3$ reaches to its maximum value of 6.2262 at specific values of the input variables $X_1 = 0.0082$ wt.%, $X_2 = 0.3$ wt.% and $X_3 = 0.15$ wt.%.

The predicted values for impact strength and elastic modulus could be obtained by substituting the above specific values of input variables X_1 , X_2 and X_3 into the regression functions of impact strength and elastic modulus, respectively.

The regressed function of impact strength is,

$$\begin{aligned} f_{I_m} &= 18.5833 + 6333.333X_1 + 601.6667X_2 + \\ &+ 178.3333X_3 - 154167X_1^2 - 691.667X_2^2 + \\ &+ 1233.333X_3^2 - 12500X_1X_2 - 100X_2X_3 - 2000X_1X_3, \\ R^2 &= 0.9958. \end{aligned} \quad (2)$$

The predicted value for impact strength is 195.19 J/m at $X_1 = 0.0082$ wt.%, $X_2 = 0.3$ wt.% and $X_3 = 0.15$ wt.% from Eq. (2).

The regression function of elastic modulus is,

$$\begin{aligned} f_{E_l} &= 579.875 + 14025X_1 - 155X_2 - 710X_3 - \\ &- 171250X_1^2 + 2087.5X_2^2 + 2350X_3^2 - 23750X_1X_2 + \\ &+ 650X_2X_3 - 23750X_1X_3, \\ R^2 &= 0.9961. \end{aligned} \quad (3)$$

The predicted value for elastic modulus is 712.73 MPa at $X_1 = 0.0082$ wt.%, $X_2 = 0.3$ wt.% and $X_3 = 0.15$ wt.% from Eq. (3).

The tested result is 195.17 J/m for impact strength, and 713.08 MPa for elastic modulus [8], which agrees with the predicted data very well and not far from the experimental results of the test 12 of Tab. 1.

3.2 Maximizing Yield and Minimizing Molecular Weight with Desired Viscosity

Myers et al. one presented a simultaneous optimal problem of maximizing yield and minimizing molecular weight with desired viscosity through two input variables reaction time x_1 and temperature x_2 [2]. The three responses variables, i.e., the yield y_1 (%), the viscosity y_2 (cSt) and the molecular weight y_3 (Mr.) of the product, and the input

variables x_1 and x_2 are shown in Tab. 2.

The desired value for viscosity y_2 (cSt) is 65 cSt [2], so the utility u_2 of response y_2 might be reflected by the deviation from 65 cSt, i.e., $u_2 = |y_2 - 65|$. The utility u_2 has the

characteristics of "the lower the better", which pertains to unbeneficial performance index. Therefore, the assessment of partial favourable probability for desired yield y_2 is conducted by its utility u_2 as an unbeneficial index.

Table 1 Test results of impact strength and elastic modulus, and their partial favorable probabilities and total probabilities

Scheme	X_1 GnPs (wt.%)	X_2 Glass fiber (wt.%)	X_3 EPDM (wt.%)	Elastic Modulus Values (MPa)	Impact Strength Values (J/m)	Favorable probability		
						P for Elastic Modulus	P for Impact Strength	$P_t \times 10^3$
1	0	0.1	0.1	540	103	0.0550	0.0458	2.5144
2	0.02	0.1	0.1	660	138	0.0672	0.0613	4.1174
3	0	0.3	0.1	695	163	0.0707	0.0724	5.1212
4	0.02	0.3	0.1	720	148	0.0733	0.0657	4.8172
5	0	0.2	0.05	612	122	0.0623	0.0542	3.3753
6	0.02	0.2	0.05	703	136	0.0715	0.0604	4.3221
7	0	0.2	0.15	598	162	0.0609	0.0720	4.3795
8	0.02	0.2	0.15	642	172	0.0653	0.0764	4.9919
9	0.01	0.1	0.05	648	116	0.0659	0.0515	3.3981
10	0.01	0.3	0.05	737	158	0.0750	0.0702	5.2641
11	0.01	0.1	0.15	610	156	0.0621	0.0693	4.3019
12	0.01	0.3	0.15	712	196	0.0725	0.0871	6.3087
13	0.01	0.2	0.1	650	158	0.0661	0.0702	4.6427
14	0.01	0.2	0.1	645	160	0.0656	0.0711	4.6653
15	0.01	0.2	0.1	655	163	0.0667	0.0724	4.8265

The assessments of partial favourable probability for maximizing yield y_1 and minimizing molecular weight y_3 are performed according to the usual procedures of the "Intersection" algorithm for multi-objective optimization [7].

Table 2 Designed experiment and results of maximizing yield and minimizing molecular weight with desired viscosity

No.	Reaction time, x_1 /min	Temperature, x_2 /°C	Yield, y_1 /%	Viscosity, y_2 /cSt	Molecular weight, y_3 /Mr.
1	80	76.67	76.5	62	2940
2	80	82.22	77	60	3470
3	90	76.67	78	66	3680
4	90	82.22	79.5	59	3890
5	85	79.44	79.9	72	3480
6	85	79.44	80.3	69	3200
7	85	79.44	80	68	3410
8	85	79.44	79.7	70	3290
9	85	79.44	79.8	71	3500
10	92.07	79.44	78.4	68	3360
11	77.93	79.44	75.6	71	3020
12	85	83.37	78.5	58	3630
13	85	75.52	77	57	3150

Table 3 Consequences of assessments for partial and total favourable probabilities of the desired experiment

No.	Response Variables			Favourable Probability				Rank
	y_1 /%	y_2 /cSt	y_3 /Mr	P_{y1}	P_{y2}	P_{y3}	$P_t \times 10^3$	
1	76.5	62	2940	0.0750	0.1132	0.0869	0.7376	2
2	77	60	3470	0.0755	0.0755	0.0751	0.4275	7
3	78	66	3680	0.0765	0.1509	0.0704	0.8120	1
4	79.5	59	3890	0.0779	0.0566	0.0657	0.2897	10
5	79.9	72	3480	0.0783	0.0377	0.0748	0.2211	11
6	80.3	69	3200	0.0787	0.0943	0.0811	0.6021	5
7	80	68	3410	0.0784	0.1132	0.0764	0.6781	3
8	79.7	70	3290	0.0781	0.0755	0.0791	0.4662	6
9	79.8	71	3500	0.0782	0.0566	0.0744	0.3293	9
10	78.4	68	3360	0.0768	0.1132	0.0775	0.6743	4
11	75.6	71	3020	0.0741	0.0566	0.0851	0.3570	8
12	78.5	58	3630	0.0769	0.0377	0.0715	0.2075	12
13	77	57	3150	0.0755	0.0189	0.0822	0.1171	13

The consequences of the assessments for partial and total favourable probabilities P_{y1} , P_{y2} , P_{y3} and P_t of this product experiment are shown in Tab. 3.

From Tab. 3, it can be seen that the test No. 3 is with the maximum total favourable probability, followed by No. 1, No. 7 and No. 10.

Furthermore, regression of the total favourable probability can be done to get more accurate optimization. The fitted result for the total favourable probability is

$$\begin{aligned}
 P_t \times 10^3 = & -203375.1310 - 2654.9450x_1 + 215.4526x_2 + \\
 & + 15.5210x_1^2 - 2.6907x_2^2 - 0.0038x_1x_2 + \\
 & + 75610.091 \cdot \ln(x_1) - 0.0403x_1^3 + 0.0112x_2^3, \\
 R^2 = & 0.7675.
 \end{aligned}
 \quad (4)$$

$P_t \times 10^3$ gets its maximum value $P_{t \max} \times 10^3 = 0.9334$ at $x_1 = 91.0622$ minutes, and $x_2 = 77.6053$ °C.

Meanwhile, the fitted result for the yield y_1 is

$$\begin{aligned}
 y_1 = & -326843.3710 - 4410.0970x_1 + 48.0746x_2 + \\
 & + 0.0180x_1x_2 + 25.9987x_1^2 - 0.4803x_2^2 + \\
 & + 12747.8522 \cdot \ln(x_1) + 0.0682x_1^3 + 0.0014x_2^3, \\
 R^2 = & 0.9926.
 \end{aligned}
 \quad (5)$$

The yield y_1 gets its proper value of $y_{1 \text{opt}} = 78.2722$ % at $x_1 = 91.0622$ minutes, and $x_2 = 77.6053$ °C.

Simultaneously, the fitted result for viscosity y_2 is

$$y_2 = 1454310.2050 + 20242.9592x_1 + 2436.0490x_2 - 0.0900x_1x_2 + 117.4347x_1^2 - 29.7790x_2^2 - 580304 \cdot \ln(x_1) + 0.3025x_1^3 + 0.1215x_2^3, \quad (6)$$

$$R^2 = 0.9723.$$

The viscosity y_2 gets its appropriate value $y_{2Opt} = 68.8928$ cSt at $x_1 = 91.0622$ minutes, and $x_2 = 77.6053$ °C.

Subsequently, the fitted result for molecular weight y_3 is

$$y_3 = -176217474.000 - 2438309.5500x_1 - 13078.9964x_2 - 5.7600x_1x_2 + 14549.0282x_1^2 + 170.7936x_2^2 + 68063495.1500 \cdot \ln(x_1) - 8.5337x_1^3 - 0.7128x_2^3, \quad (7)$$

$$R^2 = 0.9238.$$

The optimal molecular weight y_3 gets its proper value $y_{3Opt} = 3590.0681$ Mr. at $x_1 = 91.0622$ minutes, and $x_2 = 77.6053$ °C.

Obviously, the optimal status for this problem is close to test No. 3 of Tab. 2.

4 DISCUSSION

From above study, the hybrid of "Intersection" algorithm for multi-objective optimization with response surface methodology is developed. Through the procedure, the multi-response optimization problem is transformed into a single response one with the help of favorable probability; the overall / total favorable probability is the uniquely decisive index of the scheme in the optimization. Furthermore, regression analysis is conducted for the overall / total favorable probabilities of all schemes to complete the entire optimization.

5 CONCLUSION

The hybrid of "Intersection" algorithm for multi-objective optimization with response surface methodology is an effective approach in the designed experiment, which makes it possible to obtain more accurate prediction for the multi-response optimization. Considering the advantages and physical essence of the newly proposed "intersection" method for multi-objective optimization, this hybrid could supply a novel and simply way for multi-objective experimental designs.

Conflict Statement

There is no conflict of interest.

6 REFERENCES

- [1] Montgomery, D. C. (2017). *Design and Analysis of Experiments*, 9th Edition, John Wiley & Sons, Inc., 460-485, 489-554.
- [2] Myers, R. H. & Montgomery, D. C. (2016). *Response Surface*

Methodology, Process and Product Optimization Using Designed Experiments, 4th edition, John Wiley & Sons Inc., New Jersey.

- [3] Fang, K.-T., Liu, M.-Q., Qin, H., & Zhou, Y.-D. (2018). *Theory and Application of Uniform Experimental Designs*, Springer Jointly published with Science Press, Beijing, China. 102-200.
- [4] Obara, C., Mwema, F. M., Keraita, J. N., Shagwira, H., & Obiko, J. O. (2021). A Multi-response optimization of the multi-directional forging process for aluminum 7075 alloy using grey-based Taguchi method. *SN Applied Science*, 3(596), 1-20. <https://doi.org/10.1007/s42452-021-04527-2>
- [5] Teruo, M. (2011). *Taguchi Methods, Benefits, impacts, Mathematics, Statistics, and Applications*, ASME Press, New York, 7-204.
- [6] Cheng J., Tan J., & Yu J. (2011). Multi-objective robust optimization of injection molding process parameters based on TOPSIS. *Journal of Mechanical Engineering*, 47, 27-32. <https://doi.org/10.3901/JME.2011.06.027>
- [7] Zheng M., Wang Y., & Teng H. (2021). A New "Intersection" Method for Multi-Objective Optimization in Material Selection, *Tehnicki Glasnik*, 15(4), 1-7. <https://doi.org/10.31803/tg-20210901142449>
- [8] Niyaraki M. N., Ghasemi F. A., Ghasemi I., & Daneshpayeh S. (2021). Predicting of Impact Strength and Elastic Modulus of Polypropylene/EPDM/Graphene/Glass Fiber Nanocomposites by Response Surface Methodology. *Tehnicki Glasnik*, 15(2), 169-177. <https://doi.org/10.31803/tg-20190204023624>

Authors' contacts:

Maosheng Zheng

(Corresponding author)

School of Chemical Engineering, Northwest University,
No. 229, Taibai North Road, Xi'an, 710069, Shaanxi Province, China
E-mail: mszhengok@aliyun.com

Yi Wang

School of Chemical Engineering, Northwest University,
No. 229, Taibai North Road, Xi'an, 710069, Shaanxi Province, China

Haipeng Teng

School of Chemical Engineering, Northwest University,
No. 229, Taibai North Road, Xi'an, 710069, Shaanxi Province, China

A Study on the Interdependence of Quality and Satisfaction with Public Transport in Rijeka

Drago Pupavac*, Marija Ivaniš

Abstract: The paper aims to determine the interdependence between quality ratings and the level of satisfaction of Rijeka's inhabitants with the city's public urban transport system. The purpose of the study is to identify the effects that the individual elements of citizens' satisfaction with public urban transport services have on public urban transport quality ratings. To accomplish the study's objective and purpose, and to prove the main hypothesis that there is a statistically strong and positive relationship between quality ratings and satisfaction with public urban transport, numerous scientific methods have been applied, in particular the survey method, the method of analysis and synthesis, and the methods of descriptive and inferential statistics. The study's main finding points to the low quality of Rijeka's public urban transport system and to the low level of satisfaction of citizens with the system. A statistically strong and positive relationship was found between quality ratings and the satisfaction of citizens with public urban transport services. Citizens are the least satisfied with the frequency of public transport intensity and the availability of urban transport in evening.

Keywords: city; public transport; satisfaction; urban transportation

1 INTRODUCTION

Cities are the greatest invention of humankind. They make us richer, smarter, greener, healthier and happier [1]. Around the world, there are more than 30 cities that have more than 10 million inhabitants and this number is still growing [2]. The purpose of cities is to bring people together, but that is not possible without well-designed and sustainable public urban transport systems. Public urban passenger transport refers to the regular carriage of passengers within the immediate area of an urban agglomeration as well as between that area and suburban settlements, along fixed routes or lines on which most passengers are daily commuters [3]. For any modern city to function, it must have a public urban transport system. The UN [4] estimates that more than half of the world's population lives in cities, so it is not surprising that urban transport is said to be the bloodstream of modern cities [5]. Accordingly, the development of modern cities should focus on public transport. Namely, there can be no sustainable mobility without the use of public urban transport on a mass scale [6-8]. Achieving sustainable mobility calls for curbing car reliance and building cities for people rather than for cars.

In the town of Rijeka, public urban passenger transport is based on bus transport. Rail transport is the only other type of public transport available in urban traffic in Rijeka. It can be argued [9], however, that rail traffic in Rijeka is in its introductory or popularisation phase. Considerable investment in rail traffic infrastructure is required for the growth phase, which could substantially increase the share of rail traffic in Rijeka's urban traffic. How people perceive public urban transport is also of vital importance to city authorities. Studies of this type can provide local officials with information about what is important to urban transport service users and how they perceive the quality of current urban transport services. Dissatisfaction with public urban transport is of an emotional nature, and city authorities are seeking ways to resolve this problem in order to raise citizens' level of satisfaction. The trend in addressing the issue of public urban transport and citizen satisfaction is not

linear, however. Namely, when public urban transport reaches the minimum level required for a city to function normally, the increase in citizen satisfaction levels begins to slow down [10].

This paper explores the interrelationship between the satisfaction of citizens with public urban transport and the quality of public urban transport in Rijeka. In addition to the main hypothesis, two auxiliary hypotheses have also been formulated. To confirm the hypotheses, a number of scientific methods have been used, in particular the survey method and the methods of descriptive and inferential statistics.

2 THEORETICAL FRAMEWORK

2.1 Literature Review

Cities have always been hubs of numerous economic, cultural, social, political and other activities, and there has always been an interactive connection between the development of cities and the development of urban passenger traffic. Modern forms of transport have helped to continuously expand the maximum radius of cities. Public urban traffic has enabled the rapid circulation of people within cities while meeting the needs of a city's inhabitants, such as going to and from work, running administrative affairs, and shopping, as well as other needs related to education, health-care, culture, sports, etc. Some ten million people ride the Tokyo subway system daily, more than the number of people living in New York. As lifestyles change, so do the reasons for travelling [11]. Every segment of society – individuals, households, local communities, businesses – benefits from a well-developed urban transport system. Public urban traffic directly helps to alleviate congestion, reduce noise, ensure a clearer environment, facilitate local businesses, enhance safety in urban traffic and increase employment. Cities across the world are investing heavily in their public transport systems to make them competitive with passenger car transport [12]. A 5-dollar return on investment can be made on every dollar invested in public urban transport [13].

Public transport in European cities was traditionally provided by publicly owned enterprises [14]. The same applies to cities and towns in Croatia.

Fig. 1 illustrates the percentage of passengers using local urban transport in EU countries, based on a survey conducted in September 2014 and involving 28 050 respondents.

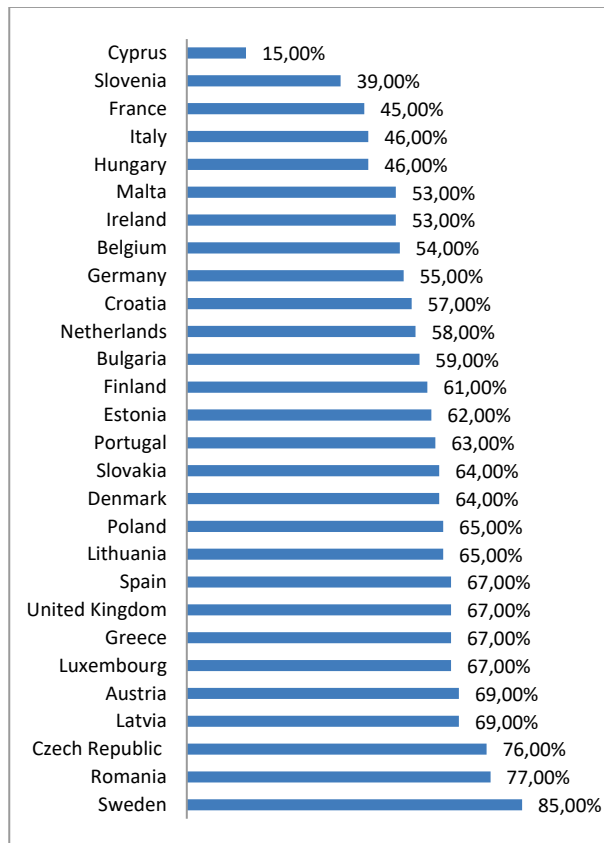


Figure 1 Percentage of commuters who used local urban transport (trams, buses, metro, commuter trains, etc.) in the EU in 2014, by country [15]

Based on Fig. 1, it can be seen that of the then EU-28 countries, Swedes mostly (85 %) used public urban passenger transport services, while the least public passenger transport services were used by Cypriots, only 15 %. In Croatian cities, 57 % of the population uses public urban passenger transport, which ranks Croatia 19th.

2.2 Research Concept

Public urban transport in Rijeka is carried out by KD Autotrolej, a bus transport service company, and HŽ Putnički prijevoz (HŽPP), the Croatian rail passenger transport company. On 1 September 2018, the two companies signed an agreement introducing integrated passenger transport in the Rijeka region. KD Autotrolej provides urban (local) transport on 18 bus lines that service the entire area of the city. The main bus stations for public urban transport in Rijeka are located on Jelačić Square and the Delta Terminal and are connected with all parts of the city, including neighbouring towns such as Opatija and Kastav. Public urban and county transport is divided into four zones (see Fig. 2).

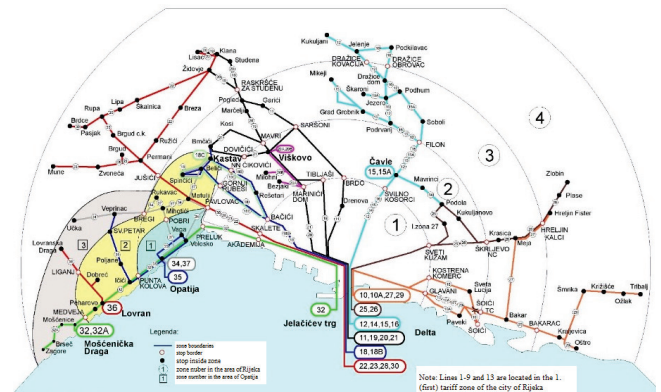


Figure 2 County lines network with tariff zones [16]

The HŽPP zones encompass the following stations and stops (see Fig. 3): Zone 1: Sušak Pećine, Krnjevo and Opatija-Matulji; Zone 2: Škrlevo, Meja, Rukavac, Jušići, Jurdani and Permani; Zone 3: Melnice, Plase and Šapjane; and Zone 4: Zlobin, Drivenik, Lič and Fužine.



Figure 3 HŽPP zones [15]

A joint ticket allows for transport by train within the four HŽPP zones (to a distance of 40 km from Rijeka) and within Zone 1 of the Rijeka bus transport system. The introduction of joint monthly tickets enables passengers to purchase tickets at a price 40 % lower than the regular price of separately bought train tickets and bus tickets.

The European Committee for Standardization issued the standard EN 13816 which defines a set of recommended criteria to measure the quality of public transport services and those are [17]: availability, accessibility, information, time, customer care, comfort, safety and ecological impact. Accordingly, in this scientific discourse, the construct "Satisfaction with public transport services in Rijeka" is measured using five items: reliability of public transport (R_{PT}), frequency of public transport intensity (F_{PT}), availability of stops (A_S), availability of public transport in evening hours (A_{ET}), and price of public transport (P_{PT}). All ratings used the following scale: (0) Don't know, (1) Not at all satisfied, (2) Not satisfied, (3) Partially satisfied, (4) Satisfied, and (5) Completely satisfied.

The reliability of the measuring instrument was analysed using Cronbach's alpha coefficient, which was 0.90. This value suggests excellent internal consistency and stability of the factors and high scale reliability (cf. Tab. 1).

Table 1 Reliability of the measuring instrument

Summary for scale: Mean=11,4739 Std.Dv.=5,41108 Valid N:403 Cronbach alpha: ,906214 Standardized alpha: ,907808 Average inter-item corr.: ,668259					
	Mean if - deleted	Var. if - deleted	Stdv. if - deleted	Item-Totl - Correl.	Alpha if - deleted
R PT	9,28	18,94	4,35	0,81	0,87
F PT	9,41	19,25	4,38	0,82	0,87
A S	8,61	19,24	4,38	0,72	0,89
A ET	9,39	19,24	4,38	0,73	0,89
P PT	9,17	18,96	4,35	0,73	0,89

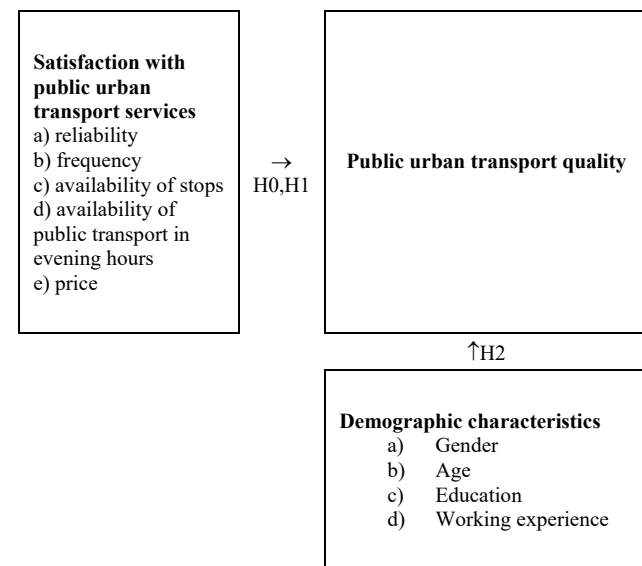


Figure 4 Conceptual research model of the interdependence of satisfaction with public urban transport services and the ratings of public urban transport quality

Although often studied in the field of economic sciences, the concept of satisfaction has rarely been the subject of research in transport [18]. Instead, research has mostly focused on the modal redistribution from passenger car to other means of urban transport [19] or on the selection of and preferences for various types of transport [20]. Felleson and Friman [18] highlight the important role of safety/security, system (supply and reliability), comfort, and staff behaviour for traveller satisfaction. According to Friman, Edvardsson and Garling [21], users who had negative experiences with urban transport primarily reported dissatisfaction with staff (drivers), the unreliability of transport, and the lack of timely information. Morfoulaki, Tyrinopoulos, and Ayfantopoulou [22] argued that the main factors of satisfaction with urban transport are waiting time, circumstances on board vehicles, frequency and on-time performance, ticketing policy, and terminal safety. According to Castillo and Benitez [23], the overall satisfaction of public urban transport users can be measured by collecting the opinions of users regarding their satisfaction with transport network coverage, transfers between lines, and waiting times, among other things.

The construct "Public urban transport quality" was measured with the question "On a scale of 0 to 5, how would you rate the quality of public urban transport services?"

In line with the above-mentioned, a conceptual research model was developed (Fig. 4).

Accordingly, the following hypotheses were constructed:

H0: There is a statistically strong and positive relationship between satisfaction with public urban transport services and the ratings of public urban transport quality.

H1: There is a statistically strong and positive relationship between all of the elements used to measure public urban transport user satisfaction and the ratings of public urban transport quality.

H2: There are no statistically significant differences in the ratings of public urban transport quality with regard to the demographic characteristics of the respondents.

Table 2 Demographic characteristics of respondents

	N	%
SEX		
Male	119	29,53
Female	284	70,47
EDUCATION		
Elementary school (ES)	-	-
Secondary school (SC)	99	24,56
Higher expertise (3 years Professional schools or Polytechnic according Bologna Declaration or 2 years according to old system of education) (HS)		
Faculty (3 years - university baccalaureus) (F_UB)	45	11,17
Faculty (4-6 years according to old system of education) (F_OS)	33	8,19
Faculty (5 years according Bologna Declaration masters degrees) (F_MD)	121	30,02
Master of science (MSc)	61	15,14
Master of science (MSc)	12	2,98
Doctorate (PhD)	32	7,94
AGE		
18-20	2	0,50
21-30	37	9,18
31-40	90	22,33
41-50	142	35,24
51-60	91	22,58
61 and more	41	10,17
WORKING EXPERIENCE		
Less than 3 years	18	4,66
From 3-5 years	12	2,97
From 5-10 years	46	11,41
From 10-20 years	131	32,51
More than 20 years	193	47,89
No working experience	3	0,74

3 SAMPLE

Research was conducted from May to August 2021. The targeted sample consists mostly of highly educated people, followed by people with secondary school qualifications, all of whom are resident in Rijeka, thus making the sample representative and reliable. A survey was publicly distributed via Facebook and Facebook Groups (*Volim grad koji teče* and *Nova riječka enciklopedija – Fluminensia*). In the three months that the survey was accessible, 403 valid questionnaires were collected. The questionnaire was

completed by inhabitants aged 18 to 61 and over. Only three of the 403 respondents had no work experience (Tab. 2).

4 RESEARCH RESULTS AND DISCUSSION

Survey participants were supposed to evaluate with grade from 0 to 5 the quality of public urban traffic in Rijeka. Based on the gathered data (cf. Tab. 3) a brief descriptive analysis of their evaluation was made.

The average grade of public urban transport is ($M = 2,23$; $SD = 1,23$). Low median value ($M_e = 2$) that divides the set into two equal parts shows low quality of public urban transport in Rijeka. This assertion is supported also by the fact that the first quartile is made up of survey participants who gave the grade to public urban transport less than one and the last quartile is made up of survey participants who evaluate public urban transport with a grade higher than 3.

The remaining 50 % of survey participants evaluate the quality of public urban transport with a grade ranging from 1 to 3.

Table 3 Descriptive statistics of evaluation of the quality of public urban transport

	<i>Q PT</i>
ARITHMETIC MEAN (M) case 1-403	2,23
MEDIAN case 1-403	2
STANDARD DEVIATION (SD) case 1-403	1,23
VALID N case 1-403	403
SUM case 1-403	900
MIN case 1-403	0
MAX case 1-403	5
25th% case 1-403	1
75th% case 1-403	3

The data in Tab. 4 indicate that the average satisfaction of respondents with public urban transport is also low ($M = 2,29$; $SD = 1,26$).

Table 4 Descriptive statistics of elements of satisfaction with public urban transport services

	<i>R PT</i>	<i>F PT</i>	<i>A S</i>	<i>A ET</i>	<i>P PT</i>
MEAN case 1-403	2,18	2,05	2,85	2,07	2,30
MEDIAN case 1-403	2- Not satisfied	2- Not satisfied	3 – Partially satisfied	2- Not satisfied	2- Not satisfied
SD case 1-403	1,23	1,17	1,30	1,29	1,33
VALID N case 1-403	403	403	403	403	403
SUM case 1-403	881	828	1152	836	927
MIN case 1-403	0 – Don't know	0 – Don't know	0 – Don't know	0 – Don't know	0 – Don't know
MAX case 1-403	5 – Completely satisfied	5 – Completely satisfied	5 – Completely satisfied	5 – Completely satisfied	5 – Completely satisfied
25th% case 1-403	1 – Not at all satisfied	1 – Not at all satisfied	2- Not satisfied	1 – Not at all satisfied	1 – Not at all satisfied
75th% case 1-403	3 – Partially satisfied	3 – Partially satisfied	4 – Satisfied	3 – Partially satisfied	3 – Partially satisfied

Respondents are the less unsatisfied with the availability of stops ($M = 2,85$; $SD = 1,30$) and the least satisfied with the frequency of public urban transport intensity ($M = 2,05$; $SD = 1,17$) as well as with the availability of public urban transport in evening hours ($M = 2,07$; $SD = 1,29$).

The constructed hypotheses are tested in the following section.

H1: *There is a statistically strong and positive relationship between all of the elements used to measure public urban transport user satisfaction and the ratings of public urban transport quality.*

Table 5 Correlation analysis of the interdependence between public urban transport quality and the elements used to measure the respondents' satisfaction with public urban transport services

Correlations Marked correlations are significant at $p < ,01$ $N = 403$ (Casewise deletion of missing data)									
	Means	Std.Dev	<i>Q PT</i>	<i>R PT</i>	<i>F PT</i>	<i>A S</i>	<i>A ET</i>	<i>P PT</i>	
<i>Q PT</i>	2,23	1,23	1,00	0,84	0,78	0,65	0,64	0,69	
<i>R PT</i>	2,18	1,23	0,84	1,00	0,81	0,67	0,66	0,65	
<i>F PT</i>	2,05	1,17	0,78	0,81	1,00	0,65	0,70	0,67	
<i>A S</i>	2,85	1,30	0,65	0,67	0,65	1,00	0,59	0,62	
<i>A ET</i>	2,07	1,29	0,64	0,66	0,70	0,59	1,00	0,61	
<i>P PT</i>	2,30	1,33	0,69	0,65	0,67	0,62	0,61	1,00	

In proving the main hypothesis of this paper, it was deemed appropriate to first test hypothesis H1, which is based on the assumption of interdependence between public urban transport quality and all of the elements used to measure the respondents' satisfaction with public urban transport services. The results of the test are shown in Tab. 5.

The data in Tab. 5 indicate that there is a statistically positive relationship between all of the elements used to measure the respondents' satisfaction with public urban transport services and the quality of public urban transport. A statistically strong correlation was found between the reliability of public transport and the quality of public transport (Pearson's correlation coefficient $r = 0,84$; $p < 0,01$) and a moderately strong correlation between public transport quality and the other elements used to measure the respondents' satisfaction with public urban transport services ($0,50 < r > 0,80$). Based on this finding, hypothesis H1 can be conditionally accepted.

H0: *There is a statistically strong and positive relationship between satisfaction with public urban transport services and the ratings of public urban transport quality.*

Multiple regression analysis was conducted to examine the correlation between satisfaction with public urban transport services and the quality of public urban transport (see Tab. 6).

The obtained finding indicates that there is a statistically strong and positive relationship between satisfaction with public urban transport services and the ratings of public urban transport quality ($r = 0,87$; $p < 0,01$), thus confirming the main hypothesis H0. Based on the finding, a conclusive multiple regression analysis model was also developed to further highlight the elements upon which the ratings of public urban transport quality in Rijeka are crucially dependent (see Tab. 7).

Table 6 Multiple regression analysis

Regression Summary for Dependent Variable: Q_PT ; Pearson's correlation coefficient $R = ,87110190$ Coefficient of determination $R^2 = ,75881851$ Adjusted $R^2 = ,75578096$ $F(5,397)=249,81$ $p < 0,0000$ SE 0,60879						
	b^*	Std.Err. - of b^*	b	Std.Err. - of b	$t(397)$	p -value
Intercept			0,12	0,08	1,63	0,10
R_PT	0,54	0,05	0,54	0,05	11,92	0,00
F_PT	0,17	0,05	0,17	0,05	3,57	0,00
A_S	0,06	0,04	0,05	0,03	1,57	0,12
A_ET	0,02	0,04	0,02	0,03	0,52	0,61
P_PT	0,18	0,04	0,16	0,03	4,91	0,00

Table 7 Conclusive multiple regression analysis model

Regression Summary for Dependent Variable: Q_PT ; $R = ,87006151$ $R^2 = ,75700704$ Adjusted $R^2 = ,75518002$ $F(3,399)=414,34$ $p < 0,0000$ SE 0,60954						
	b^*	Std. Err. - of b^*	b	Std. Err. - of b	$t(399)$	p -value
Intercept			0,18	0,07	2,72	0,01
R_PT	0,57	0,04	0,57	0,04	12,97	0,00
F_PT	0,19	0,04	0,19	0,05	4,19	0,00
P_PT	0,20	0,03	0,18	0,03	5,73	0,00

The model with the rated parameters can be written in the following form:

$$Q_PT = 0,18 + 0,56R_PT + 0,18F_PT + 0,19P_PT$$

$$(R = 0,87; F(3,399) = 249,81; p < 0,01) \quad (1)$$

Based on the obtained model (1), it is evident that the ratings for public urban transport in Rijeka first and foremost depend upon the reliability of public urban transport and then upon the frequency and price of public urban transport.

H2: *There are no statistically significant differences in the ratings of public urban transport quality with regard to the demographic characteristics of the respondents.*

The t-test and ANOVA test were conducted to prove the hypothesis that there are no statistically significant differences between the demographic characteristics of the respondents and the ratings of public urban transport quality.

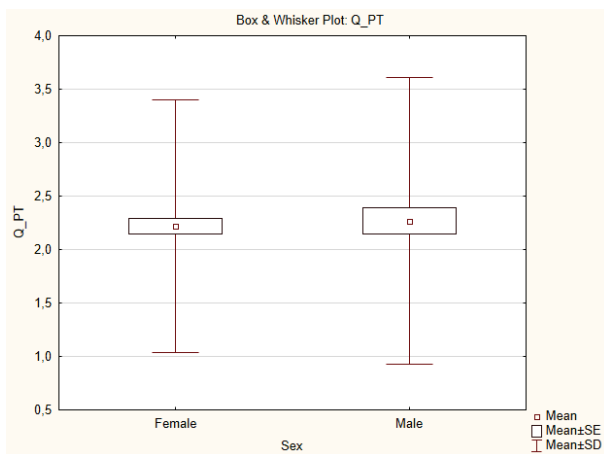

Figure 5 t-test results

Fig. 5 presents the results of the t-test carried out to establish whether there are any differences among the respondents with regard to gender.

Fig. 5 demonstrates that there are no statistically significant differences between men and women ($M_{\text{male}} = 2,26$; $SD = 1,34$; $M_{\text{female}} = 2,21$; $SD = 1,18$; t -value = $-0,37$; p -value = $0,707$) with regard to the rating of public urban transport quality. The conducted ANOVA tests also did not find any statistically significant differences in the ratings of public urban transport quality with regard to age (cf. Tab. 8), education (cf. Tab. 9) or work experience (cf. Tab. 10).

Table 8 ANOVA age groups and ratings of public urban transport quality

Age; LS Means Current effect: $F(5, 397) = 1,2481$, $p = ,28584$					
Age	Q_PT - Mean	Q_PT - Std. Err.	Q_PT - 95,00%	Q_PT - +95,00%	N
18-20	1,50	0,86	-0,20	3,20	2
21-30	2,40	0,20	2,00	2,80	37
31-40	2,15	0,12	1,90	2,41	90
41-50	2,23	0,10	2,02	2,43	142
51-60	2,09	0,12	1,84	2,35	91
61+	2,58	0,19	2,20	2,96	41

Table 9 ANOVA level of education and ratings of public urban transport quality

Ed; LS Means Current effect: $F(6, 396) = 1,3148$, $p = ,24927$					
Ed	Q_PT - Mean	Q_PT - Std. Err.	Q_PT - 95,00%	Q_PT - +95,00%	N
SS	2,36	0,12	2,12	2,60	99
HS	2,08	0,18	1,72	2,44	45
F OS	2,27	0,11	2,05	2,49	121
F UB	2,03	0,21	1,60	2,45	33
F MD	2,00	0,15	1,69	2,30	61
MSc	2,08	0,35	1,38	2,78	12
PhD	2,59	0,21	2,1	3,02	32

Table 10 ANOVA working experience and ratings of public urban transport quality

W_E ; LS Means Current effect: $F(5, 397) = ,20206$, $p = ,96152$					
W_E	Q_PT - Mean	Q_PT - Std. Err.	Q_PT - 95,00%	Q_PT - +95,00%	N
Less than 3 years	2,50	0,29	1,92	3,07	18
From 3-5 years	2,16	0,35	1,46	2,86	12
From 5-10 years	2,17	0,18	1,81	2,53	46
from 10-20 years	2,23	0,10	2,02	2,44	131
More than 20 years	2,22	0,08	2,04	2,39	193
No working experience	2,33	0,71	0,92	3,73	3

These findings indicate the acceptance of hypothesis H2.

5 CONCLUSION

Rapid and omnipresent public urban transport is a precondition to a city fulfilling its economic, cultural, social and political duties. Achieving sustainability mobility means breaking car dependency and fostering development focused on public transport. In Rijeka, public urban transport is based on bus transport. Although still in the popularisation phase, the inclusion of rail transport in Rijeka's public urban transport system aims to enrich the city's transport offering and provide more options to persuade people to use their cars as little as possible. To this end, integrated transport was introduced in Rijeka to directly improve the quality of public urban transport and enhance the satisfaction of citizens. Despite these efforts, however, the results of the study point

to the low quality of public urban transport services ($M = 2,23$; $SD = 1,23$) and to the low satisfaction of citizens with urban transport ($M = 2,29$; $SD = 1,26$). Citizens' satisfaction with public urban transport services was measured using five elements: 1) reliability of public transport, 2) frequency of public transport intensity, 3) availability of stops, 4) availability of public transport in evening hours, and 5) the price of public transport. A strong or moderately strong, positive relationship was found between all five elements and public urban transport quality ratings. Citizens are the most satisfied with the availability of stops ($M = 2,85$; $SD = 1,30$) and the price of public urban transport ($M = 2,30$; $SD = 1,33$), but the least satisfied with the frequency of public urban transport intensity ($M = 2,05$; $SD = 1,17$) and availability of public urban transport in evening hours ($M = 2,07$; $SD = 1,29$). A statistically strong and positive relationship ($r = 0,87$; $p < 0,01$) was found between the overall satisfaction of citizens with public urban transport services and the quality ratings of public urban transport services. Accordingly, a multiple regression model was developed, with public urban transport quality as the dependent variable and reliability of public transport, frequency of public transport intensity, and price of public transport as the independent variables. No statistically significant differences were found between the demographic characteristics of the respondents (gender, age, educational level and work experience) and the ratings of public urban transport service quality.

6 REFERENCES

- [1] Glaeser, E. (2011). *Triumph of the City*, Penguin Books, New York.
- [2] Jacenty Sznajder, M. (2017). *Metropolitan Commuter Belt Tourism*, Routledge, New York.
<https://doi.org/10.4324/9781315607221>
- [3] Zelenika, R. (2001). *Traffic Systems*, (In Croatian: Prometni sustavi), Faculty of Economics Rijeka, Rijeka.
- [4] <https://www.un.org/development/desa/en/news/population/world-urbanization-prospects.html>
- [5] Vuchic, V. R. (1999). *Transportation for Livable Cities*. New Brunswick, New Jersey: Centre for Urban Policy Research.
- [6] Pucher, J. & Dijkstra, L. (2003). Promoting Safe Walking and Cycling to Improve Public Health: Lessons from The Netherlands and Germany. *American Journal of Public Health*, 93(9), 1509-1516. <https://doi.org/10.2105/AJPH.93.9.1509>
- [7] Tolley, R. (2003). *Sustainable Transport*. 1st ed.; Elsevier: Amsterdam, The Netherlands.
- [8] Sarker, R., Mailer, M., & Sikder, S. (2019). Walking to a public transportstation: Empirical evidence on willingness and acceptance in Munich, Germany. *Smart and Sustainable Built Environment*, Available from: https://www.researchgate.net/publication/331439326_Walking_to_a_public_transport_station_Empirical_evidence_on_willingness_and_acceptance_in_Munich_Germany_OPEN_ACCESS [accessed Nov 29 2021]. <https://doi.org/10.1108/SASBE-07-2017-0031>
- [9] Kostelić, H. (2018). Connection of the eastern part of the primorje-gorski kotar county by a new public urban-suburban passenger transport system (In Croatian: Povezivanje istočnoga dijela primorskogoranske županije novim prijevoznim sustavom javnog gradskog i prigradskog putničkog prijevoza), *Željeznice* 21, 17(3), 27-43. Available at: <https://hrcak.srce.hr/255724> (Accessed: May 23, 2021). (in Croatian)
- [10] Knapfer, S., Pokotilo, V., & Woetzel, J. (2018). Elements of Success: Urban transportation systems of 24 global cities, McKinsey center for future mobility, McKinsey & Company, dostupno na: <https://www.mckinsey.com> (Accessed: October 18, 2021).
- [11] Ceder, A. (2021). Urban mobility and public transport: future perspectives and review. *International Journal of Urban Sciences*, 25(4), 455-479.
<https://doi.org/10.1080/12265934.2020.1799846>
- [12] Sharav, N.; Shiftan, Y. Optimal Urban Transit Investment Model and Its Application. *Sustainability* 2021, 13, 8706. <https://doi.org/10.3390/su13168706>
- [13] Taylor, L. (2021). Want to create 5 million green jobs? Invest in public transport in cities, <https://www.weforum.org/agenda/2021/04/here-s-why-cities-should-invest-in-public-transport/> (Accessed: September 15, 2021).
- [14] Wickham, E. & Latniak, E. (2017). European urban public transport: Towards a single European employment model, Work Organisation. *Labour & Globalisation*, 4(1), 160-174.
- [15] <https://www.statista.com/statistics/429239/local-urban-transport-use-in-the-eu-by-country/>
- [16] <https://www.hzpp.hr/od-1-rujna-integrirani-prijevoz-i-na-rijeckom-području> (Accessed: May 23, 2021). (in Croatian)
- [17] Comité Européen de Normalisation. EN 13816, Transportation – Logistics and services – Public passenger transport – Service quality definition, targeting and measurement. Brusel: Comité Européen de Normalisation; 2002.
- [18] Felleson, M. & Friman, M. (2008). Perceived Satisfaction with Public Transport Service in Nine European Cities. *Journal of the Transportation Research Forum*, 47(3), 93-103. <https://doi.org/10.5399/osu/jtrf.47.3.2126>
- [19] Brčić, D., Slavuj, M., & Šojat, D. (2015). Bus Subsystem Analysis in the Scope of Sustainable Urban Mobility (In Croatian: Analiza funkcije autobusnog podsustava u održivoj urbanoj mobilnosti), *Suvremeni promet - Modern Traffic*, Hrvatsko znanstveno društvo za promet, Zagreb, 35(1-2), 99-103. (in Croatian)
- [20] Hensher, D. & Prioni, P. (2002). A Service Quality Index for Area-Wide Contract Performance Assessment. *Journal of Transport Economics and Policy* 36, 93-113.
- [21] Friman, M., Edvardsson, B., & Garling, T. (1998). Perceived Service Quality Attributes in Public Transport: Inferences from Complaints and Negative Critical Incidents. *Journal of Public Transportation*, 2(1). <https://doi.org/10.5038/2375-0901.2.1.4>
- [22] Morfoulaki, M., Tyrinopoulos, Y., & Ayfantopoulou, G. (2007). Estimation of Satisfied Customers in Public Transport Systems: A New Methodological Approach. *Journal of the Transportation Research Forum*, 46(1), 63-72.
- [23] del Castillo, M. J. & Benitez, F. G. (2013). Determining a public transport satisfaction index from user surveys. *Transportmetrica A: Transport Science*, 9(8), 713-741. <https://doi.org/10.1080/18128602.2011.654139>

Authors' contacts:

Drago Pupavac, PhD, Professor
(Corresponding author)
Polytechnic of Rijeka,
Vukovarska 58, 51000 Rijeka, Croatia
E-mail: drago.pupavac@gmail.com

Marija Ivaniš, PhD, Associate Professor
University of Rijeka, Faculty of Tourism and Hospitality Management,
Primorska 42, naselje Ika, 51410 Opatija, Croatia
E-mail: mivanis@fthm.hr

The Role of Green Playground on Social Interactions of Children (Case Study: Pasdaran Park in Bandar Anzali)

Hossein Safari, Parisa Sami Khaftani*

Abstract: Playground is effective on behavior of children; however, the severity of this effect and its variety in different playgrounds is still unclear. Identifying the playing area, reviewing local observations, evaluating interactions and their extent, and analyzing the level of communication between children can lead to correct ranking of playgrounds. Features of playground and the created space along with vegetation and plants are effective factors in attracting children and encouraging them to increase interactions. Various features of playground can determine the extent to which a child participates in social interactions. These features, along with other environmental factors, allow the development of social behaviors for children. Therefore, playgrounds must have potential facilities and capabilities for influencing children to provide high quality. This study will analyze the Pasdaran Park playground using quantitative method and distributing questionnaires. Field study is required to obtain the results. To this end, a questionnaire containing 18 questions was developed using parameters and factors previously defined and scored on a 5-point Likert scale. Fifty park visitors were questioned. According to the results, the most important factors, which can increase satisfaction and meet visitor needs include providing new access routes and improving the existing ones, increasing children interactions in green spaces, using vegetation in the environment, improving the environment according to the needs of users. The results of this study can be used as guideline for designing and developing this space and similar spaces.

Keywords: Bandar Anzali; children; green playground; Pasdaran Park; social interactions

1 INTRODUCTION

Urbanization and industry development have deprived millions of children of their natural rights and have reduced their presence in the society, to the point where children are referred to as invisible beings. The external information of urban space, such as environmental configuration, may affect decision-making in way finding [1]. Urban organization systematizes urban spaces to integrate spaces, provide comfortable access, create a beautiful environment, provide necessary physical spaces while preserving urban identity and sustainable urban development [2]. People, including children, spend considerable time in urban spaces. Children are more affected by the environment. Therefore, proper design of urban spaces is very important to children. Society can be advanced when problems of children are considered and irreparable damages can be incurred when they are neglected [3]. Children have more complex psychological needs; thus, it is very effective to design urban spaces considering the developmental psychology, mental characteristics, health and safety of children in fostering creativity, enhancing the sense of cooperation and interactions of children.

Most playgrounds are just a separate part of the urban space not related to the surroundings and decorated with just a little vegetation [4]. It is not surprising that many scholars have extensively criticized the structures of these areas, children activity and interactions in these areas, because they are repetitious, predictable and tedious and are not attractive for children and do not promote their interactions. Considering the importance of accessibility in the busy modern life, each project requires to be located properly [5].

Most playgrounds are defined by several colorful structures and secure rubber surfaces (to prevent injury); they not only have limited equipment, but also malfunction as a social place [6]. The primary expectation of a playground

considering social aspects of children and their families is not met in almost any of these areas [7].

Children environment should be stimulating as much as possible so that their potential talents are realized. Talents of children will be realized better in more enriched and peaceful living and social environments [8]. If children have to be educated in a space where their researcher soul is destroyed and their growth and creativity are prevented, the efforts to improve the quality of learning and creativity growth in the future will be useless.

Emphasis on comfortable, safe, available, responsive, attractive and creative spaces underlies physical and mental growth of children. Obviously, dynamic interaction of children and discovery of environmental capabilities require their active participation in shaping and exploiting the environment. Children will understand the social sense of a space, environmental awareness, sense of belonging and effectiveness through participation in the environment and they will not consider themselves as indifferent user of the environment [9]. Children can be educated by architectural elements in physical and functional organization; in this regard, the elements fitted to physical and psychological conditions of children can be used to bring diversity and attractiveness for children.

This may remind the contemporary playground principles, defined by the McDonald's model and evoked by Eric Schollerser as commercial part of product development, which is in fact a cold spiritless place.

It is not surprising that many authors question the concept of playground and its functions, and believe that many childish games are not played in these grounds. In this context, advantages of existing playgrounds are also limited because playground sites are usually selected without proper investigations and their accessibility is not acceptable.

Playgrounds in urban environments are so inconsistent with children needs that Cunningham and Jones (1999)

consider playgrounds as confession to failure of urban planning. Perhaps, we need to return to the old concepts of playgrounds such as Children's Garden or common ideas of the 1970s, according to which parts or the whole city should be converted into a playground. In order to achieve an appropriate result and a comprehensive study of the relationship between spatial and behavioral characteristics of children, it is necessary to consider all the requirements at the same time.

The present study tends to determine whether playgrounds can become a proper space for playing and interactions and whether this can be enhanced and modified by using green space. This study considers the following general hypotheses:

- 1) There is a relationship between spatial quality and behaviors of children in playgrounds.
- 2) This relationship is strong and can be enhanced using green space.
- 3) All effective factors on quality of a playground are influenced by green space.

To properly understand these hypotheses, it is necessary to focus on children behaviors in playgrounds and consider the relationship between certain characteristics, spatial design and playing patterns in playgrounds. This study tends to present frameworks for effect of green space in enhancing social interactions of children in playgrounds by reviewing the results of previous studies in different parts of the world.

2 LITERATURE REVIEW

2.1 Natural Environment

Previous studies show that the relationship between the type of playground and children behavior indicates some general features as well as specific details that can well show the correlation between certain activities and quality of the space. One of the first studies conducted on children behavior in contemporary and traditional playgrounds and emergence of curiosity was done by Howard et al. [35]. This study showed that traditional playgrounds limit the activities of children to functional games (based on physical activity), while contemporary playgrounds induce new ways of playing and emerge new creativities for children.

Other studies show that children behavior during playing may not be affected by playground and its elements. While it cannot be ignored that some general features of space can affect children behavior and lead them to a certain game.

To explain conditions of outdoor playgrounds with different growth aspects of children, Herrington and Lesmeister list seven factors, including character, context, connectivity, change, chance, clarity and challenge [10]. These seven factors have also been widely used in other studies.

General features related to playground equipment include game finish challenge, primary needs for children activity and space flexibility for children [11].

Nicholson [12] further developed his theory of loose (non-rigid) parts, in which he addresses the relationship between creativity and the type of variables in the

environment. These results were confirmed by Herrington and Studtmann [4] and it was observed that children on the playground initially go to loose objects such as sand, water, and colorful pebbles; thus, it is recommended to create water games and sandboxes in playgrounds.

The effect of children learning is much higher in natural environments; in such environments, children learn more motor abilities and social skills [13]. The playground should support physical, social, emotional, and cognitive development of children. In addition, it is necessary to consider several important features in any playground:

- Appearance: Playgrounds should be attractive and unique.
- Application: Playgrounds should have gradual challenges, the required equipment and supplies, and flexible objects and materials to provide children with the opportunity to test and experience.
- Layout: Playgrounds should be clearly defined by individual applications and spaces, which should also be correctly connected. This will invite children to playgrounds. It is necessary to combine the playground with complex surroundings correctly and gently to have good performance [6].

2.2 Vegetation (Green Space)

Some apparent features of playgrounds such as natural [14] and unique elements and gardens and special spaces [15] have been considered as stimuli for increasing children activity. Human-designed playgrounds (without vegetation) encourage children to understand and discover the complexities within, while in natural spaces, children mostly on fantasy games and social activities [4].

Many authors emphasize the natural environment as a suitable space for children and playing. Some suggest that playgrounds should allow the development of green space and vegetation. This is associated with strong evidence of healing and therapeutic effects of natural environments on children [16].

Good results and insights were obtained in a comprehensive study on parks and playgrounds in Dhaka, Bangladesh. This study considered rankings and reasons for unwillingness of people to use public spaces and parks. Because of tropical climate [36], there are abundant and various vegetation in this country. Therefore, other effective factors are mentioned below. The results address different aspects of the quality of parks and playgrounds (Work for a Better Bangladesh, 2015).

According to objectives of this study, two important factors, which can be considered in design of suitable playgrounds, are layout and tree shadows.

In studies conducted in the South East Asia, park users addressed the following issues:

- 1) Promoting health and hygiene in the parks
- 2) Improving security of the parks
- 3) Increasing facilities and equipment in the parks
- 4) Improving access to the parks and available spaces
- 5) Others.

Therefore, it can be concluded that the main priority of people is health and hygiene of the parks. This will be significant given user culture as well as services provided by Urban Services Department of the Bandar-e Anzali Municipality. Considering the objectives of this study, the third factor to consider for increasing popularity of the parks is to increase available facilities and equipment. This, which itself includes many subsets, will also include children playground, which we tend to improve considering the user needs.

Direct participation of children in creating their space not only leads to a greater tendency for presence and interaction in the environment, but also provides physical, mental, social, talent, and creativity growth [17].

To secure parks while implementing the integrated safety, health, and environmental management, basic and standard data should be used to design playgrounds around the world, because if parents are uncertain about safety and security of the parks, this uncertainty can affect the visit to the parks and playgrounds [18]. The requirements of the 6346 National Standard Series are not considered for provision of park equipment, and there is no coherent system for supplying goods, installing based on the standards, and establishing a maintenance system in municipalities [19].

Risk assessment of existing playgrounds shows that all parks are in semi-critical range and necessarily require proactive actions; 15.4% of the cases are normal, 71.8% are semi-critical and 12.8% are critical and require urgent action in terms of safety [20].

2.3 Access Route

Another favorite activity of children, which is manipulation and alteration of existing objects, requires opportunities to access various objects [21]. It is very important to consider functions of the playground components in terms of their cost-effectiveness in order to describe proper understanding of the environment and its usefulness [22].

Frost [23] asserts that an appropriate playground should be located in an enclosed, compact, and efficient area to be more attractive to children. Attractiveness of the playground is one of the effective factors on creativity of children.

Fakouri Moridani and Hossein Safari [24] conducted a study in Institute for Intellectual Development of Children and Young Adults. In spite of its proximity to downtown, this institute has not been noticed by citizens. In order to measure the status quo (present situation) and survey operations, observation was used in the gate, directional splits, and people following. The findings indicate that given that motor potential and access are based on the simulation to reach the convenient center, but gate observation showed that traffic and traffic levels of pedestrians were very low in this path [24].

2.3 Interactions

The type, quality, and diversity of playgrounds directly affect the type, quality, and diversity of their play [13], as

well as behavior of children [25, 37]. This implies that an appropriate playground can provide good opportunities and qualities for children to develop their collaboration and interaction [26]. Different games can be considered as effective strategy for identifying playgrounds [27].

Some other studies confirmed that traditional playgrounds stimulate competition in children, while appropriate playgrounds stimulate the sense of adventure and curiosity in children. Campbell and Frost [28] found that most children playing in traditional playgrounds tend to be physically and functionally active and small number of people are engaged in creative activities. They recommend creative playground, which includes contemporary elements along with some adventure facilities. In this situation, creativity of children has increased significantly.

The extent of children participation in different games was addressed in many subsequent studies, which showed that children tend to play games and show violent behaviors in traditional grassed playgrounds [29]. Brown and Burger [30] studied the relationship between the type of playground and the form of the game. Their study did not show a significant relationship between these two factors. In other words, they could not find a relationship between the type of playground and activities of children.

Moore et al. [13] suggested that multiple stimulation is a key factor for children who experience sensory disorders, so that they can further understand the environment. Various studies on arrangement of equipment in playgrounds also show that creation of limited and semi-private spaces for children to provide a limited sense of peace and solitude has a great influence on creativity growth, so that they tend to include other children after dominating a certain part from the playground [31].

Brown and Burger [30] suggested playground zoning, which encourages sandboxes, typical protected areas, and car-free spaces. They believe that landscaping and provision of suitable materials for playing in this space are the most important effective factors on behavior of preschool children.

3 METHODOLOGY

Design of methodology begins with primary scientific presumptions. One of the important components in design of playgrounds is their potential facilities as well as social features, which, given the different equipment available in them, can further show their positive and negative points to users. In this way, a contemporary playground can be a real space and a social environment for children given their specific needs. Sociology of a playground focuses on social capabilities of the space; that is, it considers the ability of children to engage in social games, apply social skills such as cooperation and understand rules and their participation. Similarly, the ability to play in one place can be considered the potential to play in the space. For example, potential child interaction in real activities (not only games) is considered in this section.

The data was collected through observations, interactions, and questionnaires. Next, descriptive data, including general profile of participants, was collected and

analyzed. This method provides an overview of the discovered data. This data analysis method creates a triangle between the data collected in observations (using notes and wallpapers) and surveys. This procedure can provide the considered results for the studied site.

The obtained data is directed according to the information obtained from the literature and previous studies, so that they can be generalized properly. In order to better understand the procedure, note the methodology as outlined below.

- **Problem statement:** In this section, the information and general reviews of previous studies are presented and the gaps are explained accordingly. It is determined that the parks existing in different cities can be improved. This improvement can be defined in different areas; this study focuses on playgrounds.
- **Identification of objectives and questions:** Focus on a certain point of a public park in Bandar-e Anzali requires consideration of user needs. This can be defined in different ways. A certain part of this park is dedicated to children and special items will be determined to enhance the quality of their activities. These can be addressed in the form of research questions as follows:
 - Is green space involve in quality of playground?
 - Can green space be associated with certain goals such as increased curiosity of children?
 - Are parents encouraged to use green space in the public park?
 - Does green space have a positive effect on creativity of children and users of this park?
- **Methodology:** The choice of methodology is an important and significant approach that paves the way for determining the overall framework and strategy of the study. There are three types of scientific studies, qualitative, quantitative, and mixed (qualitative-quantitative). In this study, considering the relevant literature, the best and fastest way to achieve relevant and reliable results is to use quantitative method with questionnaire. This method can provide reliable results by using proper statistical techniques, if proper number of users are available.
- **Data collection:** In this study, field databases are formed by using questionnaires. These questionnaires are designed on a five-point Likert scale and have a good statistical balance. The collected data can be analyzed and the results can be obtained.
- **Analysis and findings:** This step includes two parts. The first part involves statistical studies using a software. This section uses SPSS software version 19; initial statistical results can be achieved by appropriate statistical design for data analysis. The second part involves explanation of statistics to translate mathematics to practical language. In this section, the considered results can be approached and the items defined by the questions can be evaluated.
- **Discussion and conclusion:** This section is according to the previous section; in other words, it is the final argument or inference. Given the literature and the data collected in the previous section, logical implications are

made which will help other authors and provide others with the plan considered by the author given the findings.

Observations, surveys, and interactions can provide a deep understanding of people and their demands. The literature indicates that authors tend to use other methods such as questionnaires because of limitations of interview [32, 33]. Taylor and Bogdan also argue that people have stable attitudes and respond differently in different situations; thus, distribution of questions should be well controlled and it is recommended to eliminate the existing barriers to interviews. Considering the development of statistics science and the ability to use its capabilities, however, it is no longer recommended. In general, the methodology and its steps can be seen in the Fig. 1.

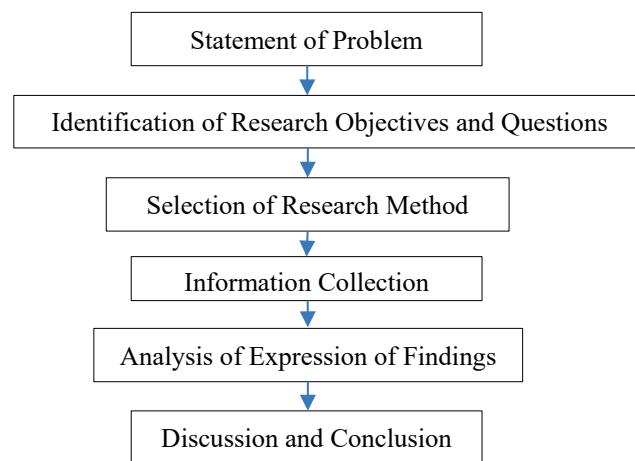


Figure 1 Steps of methodology (Author)

3.1 Case Study

This park has an area of 23490 m² and is located in Pasdaran Street (Fig. 2 and Fig. 3). Vegetation of the park include trees such as pine, oak, mulberry, and shrubs such as magnolia, privet (*Ligustrum vulgare*), crepe myrtles, as well as hedge plants such as Buxus and Japanese barberry (*Berberis thunbergii*). The park was opened in 2002 with furniture, play equipment, toilets, drinking water, lighting fixtures, security guards, and stores. There is also a gardener. The park is also classified as Beach Park [34].



Figure 2 Right: available spaces and vegetation used in the park; left: interior and playground (Author)

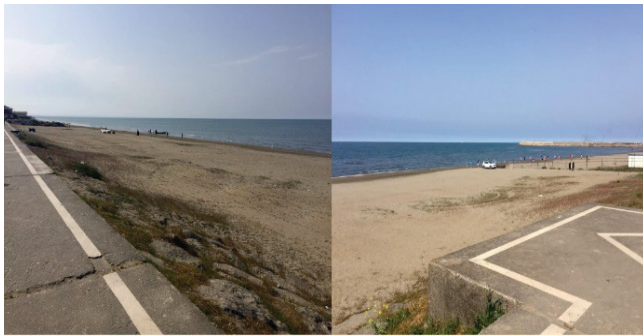


Figure 3 right: the park in vicinity of the beach; left: limited green space and empty spaces in different parts of the park (Author)

Table 1 Suggested solutions for designing spaces suitable for children (Author)

No.	Solution	Example	Result
1	Safety	Increasing resident interactions, monitoring the commuting, controlling spaces by residents, reducing uncontrolled spaces, arranging the components, controlling the number of people living in the area	Increasing the social and psychosocial security of children
2	Usable open spaces	Zoning children-specific open spaces, suitable combination of indoors and outdoors	Responsiveness to development of communications and game in children
3	Connection with nature	Designing green spaces in which children can play	Healthy growth of senses, learning and creativity, and respect for nature
4	Playgrounds	Spaces for water games, slides, swing, special playgrounds, and local match fields on a larger scale	Fun and happiness of children and their growth
5	Semi-open spaces	Providing semi-open and green spaces between buildings, designing balconies of houses	Leisure time supervised by parents
6	Multipurpose indoor spaces for children	Providing salons with flexible layout, like kindergarten, for group playing	Learning social behaviors and collective activities
7	Roof	Providing green roof, outdoor gathering place for all people	Increasing creativity, improving the quality of life and creating a healthy environment
8	Spatial proportions suitable for children	Using children-specific furniture, balancing the spaces used for children	Sense of attachment to space, satisfaction with childhood, health promotion
9	Children involvement in design and planning	Organizing painting workshops on bio spaces, using children views in furnishing playgrounds	Guaranteeing social sustainability, involving children in decision making

4 RESULTS

Different needs of children, as well as their age and behavioral requirements, are associated with the need to build playgrounds, which generally should be considered before building a new space. Some of these features can be listed as shown in Tab. 1.

To obtain practical results, it is necessary to perform field studies. To this end, the parameters and factors defined in the previous section were used as a questionnaire with 18 questions on a five-point Likert scale. People and visitors ($N = 50$) were questioned and the results were extracted using this information. General profile of these people can be presented as descriptive statistics as follows: 36% of people were 20-30 years old, 30% were 10-15 years old, 24% were older than 30, and only 14% were 15-20 years old. Participants were 60% male and 40% female. Of them, 50% had high school diploma and under, 28% had master's degree, 18% had bachelor's degree, and only 4% had PhD; 72% of participants were Gilak and 28% were Turk; 90% of participants lived in Bandar-e Anzali and 10% did not; 72% lived in other neighborhoods and 28% lived in Pasdaran neighborhood.

The results obtained from the questionnaire were imported to SPSS software and analyzed by one-way t-test. To compare the means obtained from the used questionnaire, it is necessary to define a value as a reference for significance. In this study, the significance value at 5% is 3.5 and the significance value at 1% is 4. The results can be seen in Tabs. 2 ÷ 5.

Table 2 Statistical analysis of environment questionnaire by SPSS (author)

No.	Question	Degree of freedom	Value
1	Fresh air can have a positive effect on playing.	49	4.7**
2	Natural environment encourages practice and effort.	49	4.6**
3	Proper playground saves money.	49	3.4
4	Diversity of games increases by vegetation.	49	3.5*

* $P < 0.05$; ** $P < 0.01$

In this section, all questions get high score from respondents. Thus, the positive effect of fresh air on playing and increased practice and effort in the natural environment was significant ($P < 0.01$) and increased diversity of games by vegetation was significant ($P < 0.05$). However, the effect of financial savings when using the playground is not acceptable.

Table 3 Statistical analysis of vegetation questionnaire by SPSS (author)

No.	Question	Degree of freedom	Value
1	Vegetation in the playground increases children interactions.	49	3.86*
2	Vegetation in the playground increases children health awareness.	49	4.42**
3	Vegetation in the playground increases environmental awareness.	49	4.5**
4	Vegetation increases the presence of children in playgrounds.	49	3.92*
5	Vegetation around the playground increases creativity of children.	49	3.76*

* $P < 0.05$; ** $P < 0.01$

All questions related to vegetation were significant. Increasing children health and environmental awareness with vegetation was significant ($P < 0.01$) and increasing interactions, children presence and creativity with vegetation was also significant ($P < 0.05$).

Table 4 Statistical analysis of access route questionnaire by SPSS (author)

No.	Question	Degree of freedom	Value
1	Green routes encourage more children to playgrounds.	49	4.5**
2	Increasing number of green access routes can encourage parents to use spaces.	49	4.48**

* $P < 0.05$; ** $P < 0.01$

In terms of access routes, majority of respondents evaluate both questions effective and both questions are significant ($P < 0.01$).

Table 5 Statistical analysis of interactions questionnaire by SPSS (author)

No.	Question	Degree of freedom	Value
1	Vegetation in playground can lead to their curiosity.	49	3.76*
2	Peacefulness of green playgrounds leads to longer presence.	49	4.7**
3	Risky behaviours in playgrounds are reduced by children contact with the green space.	49	4.2**
4	Green playground lead to stronger communications of children with each other.	49	3.36
5	Children find more friends in green playgrounds.	49	3.46
6	Playing in green spaces increases vitality and happiness of children.	49	4.7**
7	Peace of children and parents in the playground explains the presence of vegetation and improves children interactions.	49	4.52**

* $P < 0.05$; ** $P < 0.01$

Peace of children resulting from vegetation will lead to more presence and reduces their risky behaviors. Their vitality and happiness increase and peace of children and their parents increase their interactions. These questions are all significant ($P < 0.01$); they also show significant results regarding creativity ($P < 0.05$). Two questions regarding more communication between children in green playground and increasing number of friends are not significant.

5 CONCLUSION

According to the results, it can be argued that the most important parts with the ability to increase satisfaction and meet user needs are:

- 1) Providing green access routes and modifying the existing routes.
- 2) Increasing children interactions in green playgrounds.
- 3) Using vegetation in the environment.
- 4) Modifying the environment and designing according to user needs.

It should be noted that all of these cases are presented as mean of responses. According to the studies conducted in different parts of the world, emphasis on access routes is reported as the most important factor for the first time. Increasing interactions and vegetation have also been considered in other studies. Special need of users is also one of the useful factors in design, which can have effective results in improving the quality of the environment.

Using the results of this study, the process of improving and redesigning the Anzali Pasdaran Beach Park can be done to achieve a desirable and efficient park in accordance with demands of users and environmental resources, as well as emphasis on green space and its effects. This space can increase the public interest in the park and entertain children in a special playground. Children entertainment along with increasing social capabilities and interactions and creativity is one of the most important factors considered in various design processes and can be a sign of design success.

6 REFERENCES

- [1] Safari, H. & Moridani, F. F. (2017). Syntactical analysis of the accessibility and sociability of a square in the Kuala Lumpur City Center. *Frontiers of Architectural Research*, 6(4), 456-468. <https://doi.org/10.1016/j.foar.2017.06.005>
- [2] Askarizad, R., Safari, H., & Pourimanparast, M. (2017). The Influence of Organizing Historical Textures on Citizenry Satisfaction in the Old Texture Neighbourhoods of Rasht. *Emerging Science Journal*, 1(3), 57-68. <https://doi.org/10.28991/ijse-01114>
- [3] Aghajanbegloo, S. (2013). *Story Therapy*. Varjevand Publications, First Printing, Tehran.
- [4] Herrington, S. & Studtmann, K. (1998). Landscape interventions: new directions for the design of children's outdoor play environments. *Landscape and urban planning*, 42(2-4), 191-205. [https://doi.org/10.1016/S0169-2046\(98\)00087-5](https://doi.org/10.1016/S0169-2046(98)00087-5)
- [5] Khotbehsara, E. M. & Safari, H. (2018). A Systematic Review of Affective Factors on Locating Specialized Hospitals. *Civil Engineering Journal*, 4(9), 2210-2217. <https://doi.org/10.28991/cej-03091151>
- [6] Czalczyńska-Podolska, M. (2014). The impact of playground spatial features on children's play and activity forms: An evaluation of contemporary playgrounds' play and social value. *Journal of environmental psychology*, 38, 132-142. <https://doi.org/10.1016/j.jenvp.2014.01.006>
- [7] Aghajanbegloo, S. (2015). The Effect of Wall Painting on the Transfer of Educational Concepts in Preschool Children. *Proceedings of the 5th Congress of the Iranian Psychological Association, Contemporary Psychology*, 11-13.
- [8] Ebrahimi, N. (2017). The Effect of Architecture (Residential) on Children's Behavior. *International Conference on Architecture and Urban Planning, Tehran*.
- [9] Erfanifar, F., Khoshbakht, F., Shirvani, D., & Alborzi, M. (2015). Children's Play and Characteristics. *Congress of Iranian Psychological Association, Tehran*.
- [10] Herrington, S. & Lesmeister, C. (2006). The design of landscapes at child-care centres: Seven Cs. *Landscape Research*, 31(1), 63-82. <https://doi.org/10.1080/01426390500448575>
- [11] Sando, O. J. (2021). Places for Children (Doctoral dissertation, Norwegian University of Science).

- [12] Nicholson, S. (1974). *The theory of loose parts*. In G. Coates (Ed.), *Alternative learning environments* (pp. 370-381). Stroudsbrough, PA: Dowden, Hutchinson, and Ross.
- [13] Moore, R. C., Goltsman, S. M., & Iacofano, D. S. (1997). *Play for all guidelines: Planning, design and management of outdoor play settings for all children*. MIG Communications, 800 Hearst Ave., Berkeley, CA 94710.
- [14] Mason, J. (1982). *The environment of play*. Leisure Press.
- [15] Hartle, L. (1996). Effects of additional materials on preschool children's outdoor play behaviors. *Journal of Research in Childhood Education*, 11(1), 68-81. <https://doi.org/10.1080/02568549609594696>
- [16] Kim, S. (2021). Environmental Design Factors for Open Space in Healing Facilities-Focused on Ann & Robert H. Lurie Children's Hospital and the Nationwide Children's Hospital. *International Journal of Advanced Culture Technology*, 9(1), 7-15.
- [17] Najafi, M., Daviran, A., & Nouralshahi, J. (2013). Strengthening the Creative Sense of Children in Playgrounds Case Study of Zanjan Flower Park, *The First Conference on Architecture and Sustainable Urban Spaces*, Mashhad, Parham Applied Research Group.
- [18] Saraci, M. H., Jamshidi, P., & Surrender, S. (2013). Children's Playground Safety Assessment in Parks (Case Study: Yazd Town Parks), *First National Conference on Geography, Urban Development and Sustainable Development*, Tehran, Koomesh Environmental Society, Aviation University.
- [19] Soltan Nayeri, M. (2016). Assessment of the safety improvement of children's play land equipment in Tehran's parks, *the first National Conference on the Development of Health, Safety and Environment in the field of recreational, sporting, religious, and cultural places of the city with the approach of protecting citizens and cities*, Tehran, Amirkabir University of Technology.
- [20] Rezvani, S., Heidari, S., Roustapisheh, N., & Dokhanian, S. (2022). The effectiveness of system quality, habit, and effort expectation on library application use intention: the mediating role of perceived usefulness, perceived ease of use, and user satisfaction. *International Journal of Business Information Systems*, 1-18. <https://doi.org/10.1504/IJBIS.2022.10049515>
- [21] De Monchaux, S. & De Monchaux, S. (1981). *Planning with children in mind: A notebook for local planners and policy makers on children in the city environment*. NSW Department of Environment and Planning.
- [22] Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA, US.
- [23] Frost, J. L. (1992). Reflections on Research and Practice in Outdoor Play Environments. *Dimensions of Early Childhood*, 20(4), 6-10.
- [24] Moridani, F. F. & Safari, H. (2018). Evaluation of the vitality of Lahijan's institute for the intellectual development of children and young adults using space syntax.
- [25] Kritchinsky, S. (1969). *Planning environments for young children: Physical space*.
- [26] Wardle, F. (2000). Supporting constructive play in the wild. *Child Care Information Exchange*, 5, 00-26.
- [27] Shin, D. & Frost, J. L. (1995). Preschool children's symbolic play indoors and outdoors. *International Play Journal*, 3(2), 83-96.
- [28] Campbell, S. D. & Frost, J. L. (1985). The effects of playground type on the cognitive and social play behaviors of grade two children. *When children play*, 81-88.
- [29] Humphreys, A. P. (1984). Rough-and-tumble in preschool and playground. *Play: in animals and humans*.
- [30] Brown, J. G. & Burger, C. (1984). Playground designs and preschool children's behaviors. *Environment and Behavior*, 16(5), 599-626. <https://doi.org/10.1177/0013916584165004>
- [31] Frost, J. L., Wortham, S. C., & Reifel, R. S. (2001). *Play and child development*. Merrill, Prentice Hall.
- [32] Zalnezhad, K., Esteghamati, M., & Hoseini, S. F. (2016). Examining the Role of Renovation in Reducing Crime and Increasing the Safety of Urban Decline Areas, Case Study: Tehran's 5th District. *Armanshahr Architecture & Urban Development*, 9(16), 181-192.
- [33] Zalnejad, K., Hosseini, S. F., & Alipour, Y. (2019). The Impact of Livable City's Principles on Improving Satisfaction Level of Citizens; Case Study: District 4 of Region 4 of Tehran Municipality. *Armanshahr Architecture & Urban Development*, 12(28), 171-183.
- [34] Alipour, M. (2017). *Spatial Analysis of Urban Green Space in Bandar Anzali (Case Study of District 1 Gardens)*. Master's Thesis for Geography and Urban Planning, Faculty of Humanities, Department of Geography, Islamic Azad University of Rasht.
- [35] Hayward, D. G., Rothenberg, M., & Beasley, R. R. (1974). Children's play and urban playground. A comparison of traditional, contemporary and adventure playground types. *Environment and Behavior*, 6(2), 131-167. <https://doi.org/10.1177/001391657400600201>
- [36] Ghiasi, F., & Fashi, F. M. (2019). Ethical Climate of Operating Room of Educational Hospitals Affiliated To Ilam University of Medical Sciences from the Point of View of Anesthesia and the Operating Room Students. *Indian Journal of Forensic Medicine & Toxicology*, 13(4), 1812-1816. <https://doi.org/10.5958/0973-9130.2019.00573.5>
- [37] Sharifi, N., Rezaei, N., Fathnezhad-Kazemi, A., & Ghiasi, F. (2021). Association between Fear of COVID-19 with Self-care Behaviors in Elderly: A Cross-Sectional Study. *Social Work in Public Health*, 36(5), 606-614. <https://doi.org/10.1080/19371918.2021.1937435>

Authors' contacts:

Hossein Safari, PhD
Department of Architecture,
Rasht Branch, Islamic Azad University,
Rasht, Iran

Parisa Sami Khaftani, PhD, Assist. Prof.
(Corresponding author)
Department of Architecture,
Rasht Branch, Islamic Azad University,
Rasht, Iran
parisasami@chmail.ir

Reducing Short Circuit Current with a Novel Approach Model in Wind Power Plants

Mehmet Şen*, Muciz Özcan

Abstract: Wind energy from renewable energy sources, the share of electric power system is increasing faster than other sources. This requires the analysis of the effect of the wind power plants have on the power system in connection with the network. The wind produces energy-dashed and variable power, which can cause some disruptive effects. In this study, a short circuit analysis was performed using the DlgSILENT PowerFactory program. Minimum short circuit analysis calculations have been performed for 15 wind turbines and 19 bus-bar wind power plants. The error current has been reduced to a minimum by changing the transformer connection and ground resistance for the most commonly encountered phase-to-ground short circuit analysis. This application has been developed against more complex and costly methods. This prevents the generators from switching to island mode after a phase-to-ground short circuit current is affected by the fault.

Keywords: Bus-Bar; DlgSILENT; Island mode; Short circuit; Wind power

1 INTRODUCTION

The history of taking advantage of wind power dates back five thousand years. The use of wind has only been aimed at obtaining mechanical power in a significant part of human history [1]. Initial irrigation practices in Mesopotamia were conducted in Babylon in 2800 B.C. [2]. Windmills were first found in Alexandria. It is known that Turks and Persians use windmills in 7 [3]. In European countries, windmills started to be used in the 12th century [4]. Until the early 20th century, the wind was used to provide the mechanical power needed to pom-palate or grind grandees. Due to the industrial revolution, the start of industrial progress and the high efficiency of fossil fuel (oil, coal, etc.) consumption and electricity production have increased the interest in wind energy.

Wind energy among renewable energy sources has improved greatly in recent years in terms of the share of use in electricity production [5]. This increase in the number of wind plants has made it necessary to increase work on network integration in particular. National and international IEEE standards must be adhered to when integrating wind energy distributed production systems into the network [6]. These standards are important to ensure high power quality and system reliability of the network.

A significant factor in ensuring network quality and reliability is the limitation of short-circuit currents. Failure in power systems is unpredictable. Therefore, faults that may occur in the system must be limited to a certain level [7]. The most common short circuit fault in an electrical system is a phase to ground short circuit [8]. The method of application used in our country to protect the system is to limit the phase-to-ground short circuit current of the 34.5/154 kV power transformer to 995 A. [9]. For literature scans, the world application has limited this current to 360 A for overhead lines and 800 A for underground cables [10].

A short circuit fault that will occur on the low voltage (LV) side of a power plant connected to the mains causes an increase in the fault current value [11]. This results in complexity and increased cost in relay coordination. Therefore, the fault current is limited to the minimum level

before the relay is coordinated. In this way, relay coordination can be more effective.

Italy, Japan, Ireland, Russia, Peru, In Spain [12] the unearthed system, which is widely used, is defined as a system where conductors of energy pass through have no direct connection to the ground [13]. Such systems are actually ground-induced with distributed capacitances, transformers and motor windings. When there is no earth fault, the phase ground voltage is approximately equal to each three phases due to the system's off-sitive effect on the 3-phase system [14].

In the case of the first failure of the unearthed system, there is a possibility that the voltage will increase [15]. This high voltage causes a second fault at the weakest point of the system and a more harmful fault current is generated. Over-current protection devices clear the fault. Therefore, surface resistance between phase grounds causes a large rear cause. The size of the fault current may not be sufficient for protection equipment, so that both compensations can result in costly damage and the system can remain unenergized until the fault is repaired. It is very important to locate the first fault and to rectify the fault.

In northern and South Europe, the Peterson coil, which is widely used in China and Israel [16], is used to limit the arching in a 3-phase ground-plunging system that is not grounded during a soil fault [17]. This coil was originally developed by W Petersen in 1916 [18]. This solution is sometimes used in high voltage networks and industrial distribution systems. The sensitivity of the active components' residual currents and operating protection relays and current transformers are used to provide selectivity [19].

When phase ground failure occurs, the inductance of the Peterson coil changes immediately, the zero-component current of the faulty seedling changes very quickly, the others remain the same [20]. The zero-component current of the faulty seedling is maximum in phase ground failure [21]. The equivalent soil capacitance of the network can be measured instantaneously, compensating the capacity of the Peterson coil and the system [22]. The reactor is used between neutral and ground, compensating for the capacitive currents of the system [23]. This grounding method is mainly used in 110

kV systems. The Peterson coil is a reactor used in the system to clear the fault current.

Our aim in this study is to prevent the general from being affected by this fault and going into island mode in the phase ground short circuit that will occur in the system when the generator is operating in parallel with the network. Therefore, the phase ground fault current that may occur in the system must be limited to a certain value, either by neutral ground resistance or by using the Peterson coil. The short circuit current has an active and a reagent component. Therefore, the active component of the short circuit current can be limited to the resistance and the reagent component can be limited to the Peterson coil to be connected in series to the resistance. It's more effective to use this kind of neutral ground in high-voltage networks where generators are connected. The faulty area must then be isolated from the system, depending on the system specification, by means of the relay and protection circuits to be used in relation to the radial or ring mains condition.

2 SHORT CIRCUIT

Most of the faults that may occur in alternating current power systems are unstable, and few are stable. Normal operation of the system is balanced, but a short circuit fault in the system temporarily impair normal conditions [24]. The frequency of occurrence is usually the same, although faults are practically different to the point of occurrence of the fault and the structure of the system. According to a study, these frequency power systems are:

- 5% for three-phase symmetrical short circuit
- 10% for dual phase to ground short circuit
- 15% for phase short circuit
- The phase-to-ground short circuit is 70% [25].

This data shows that the short circuit fault in three-phase systems is less than the other types of fault. Short circuit failures in three-phase systems are often caused by operational errors. However, faults in power systems may initially have a different type of short circuit fault, but depending on the duration of the short circuit that has occurred, they may eventually become a three-phase short circuit. Although calculations of three-phase symmetric short circuit faults can only be carried out easily by knowing the correct component value, calculation is required by determining all component impedances according to the type of fault to calculate other short circuit fault variations [26].

2.1 Short Circuit between Phase and Ground

The phase-to-ground short circuit occurs when any of the phase conductors touch the ground, the neutral conductor or a metallic body with ground contact, as a result of an isolation puncture or jump during incorrect maneuvers. The most common short circuit fault when examining transmission systems is a phase-to-ground short circuit fault [27].

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & r & r^2 \\ 1 & r^2 & r \end{bmatrix} \begin{bmatrix} I_R \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} I_R \\ I_R \\ I_R \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & r & r^2 \\ 1 & r^2 & r \end{bmatrix} \begin{bmatrix} V_R \\ V_S \\ V_T \end{bmatrix} \quad (2)$$

As described in the formulas, all component currents are equal and the sum of these three components current reaches the current. This allows the calculation of a phase-to-ground short circuit fault current to be carried out using Eq. (1) and Eq. (2).

2.2 Short Circuit between Phase and Phase

The phase-to-phase short-circuit condition, which results from two conductors touching each other or incorrect maneuvers, is the second common short-circuit fault after the phase-to-ground short-circuit condition in practice. In the phase-to-phase short circuit, the current of the two phases is equal to each other as absolute magnitude, while the current directions are reversed. In the phase not included in the fault, the current value is zero (Eq. (3) and Eq. (4)).

$$I_S = -I_T, I_R = 0 \quad (3)$$

$$V_{Sg} - V_{Tg} = Z_f I_S \quad (4)$$

Symmetrical components can be calculated with Eq. (5) using phase-to-phase short circuit failure impedance. The current and voltage values in the system can then be checked.

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & r & r^2 \\ 1 & r^2 & r \end{bmatrix} \begin{bmatrix} 0 \\ I_S \\ -I_S \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ \frac{1}{3}(r - r^2)I_S \\ \frac{1}{3}(r^2 - r)I_S \end{bmatrix} \quad (5)$$

2.3 Short Circuit between Two Phase and Ground

Two phase-to-ground short-circuit events occur when the two conductors touch the conductors connected to the ground or earth. It's very rare in practice. For example, S and T phases are a short-circuit condition that can result in contact with each other and the ground. In phases S and T, the voltage to earth is equal to each other in the event of a short circuit fault. This allows all component currents of the two phase-to-ground short circuits to be calculated as described in Eq. (6) and Eq. (7).

$$V_{Sg} = V_{Tg} = Z_f (I_S + I_T) \quad (6)$$

$$I_R = I_0 + I_1 + I_2 = 0 \quad (7)$$

2.4 Short Circuit between Three Phases

A three-phase short circuit fault with a symmetrical fault occurs in practice as a short circuit variant that is least encountered. However, some short-circuit statuses are initially starting in a different short-circuit type, resulting in a three-phase short-circuit fault. The highest short circuit current value occurs in three phase short circuit. Phase voltages are equal and zero in three phase short circuit failure. In this way, component currents can be calculated as described in Eq. (8) and Eq. (9).

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & r & r^2 \\ 1 & r^2 & r \end{bmatrix} \begin{bmatrix} V_R \\ V_S \\ V_T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (8)$$

$$I_R = \frac{V_f}{Z_1} \quad (9)$$

The internal phase in the I_R phase conductor is the value of the short circuit current that can occur and the absolute values of the short circuit current that occurs in all phases are equal ($I_R = I_S = I_T$).

3 MATERIAL AND METOD

The study performed short circuit analysis of the power system on the wind power plant. Phase-to-ground short circuit calculations have been carried out by performing different scenarios in the system. The DigSILENT PowerFactory program uses the IEC 60909 standard for short-circuit calculation. This standard can allow a short circuit to be calculated in rated frequency 50 Hz and 60 Hz for high voltage networks up to 550 kV.

The necessary protection measures must be taken to prevent the transformer from entering island mode. Island mode is the operating state in conjunction with the loads it supplies independently by leaving the network to prevent the transformer from affecting or affecting as a result of a fault on the network. Protection methods are important for this type of situation and for continuous energy transmission to the network.

In renewable energy sources that are connected to the network, they may be affected by any errors that may occur in the system. The system can switch to island mode for the error to occur. In this case, the earth makes it necessary to carry out fault protection. This means that the phase-to-ground fault current in the system must first be reduced using resistance or the Peterson coil. It is a more effective method to take such a protection measure on the transformer first. This has been identified and explained as the most effective method.

The modeled system is a radial system. The load flow in the radial system is unidirectional [28]. In this case, the current of the fault is not only supplied by the network but also by wind turbines. In this context, the effects of wind turbines on the bar and other bars at the common connection

have been analyzed. Phase-to-ground short circuit analysis, when wind turbines are in the circuit and wind turbines are out, phase-to-ground short circuit is created in the system and output results are recorded. The calculation results in pulse current, initial symmetric short circuit current, short circuit thermal current and short circuit power values. The generated short circuit has been performed for the maximum condition according to IEC 60909 standard. The most important factor affecting phase-to-ground short circuit analysis is the connection and ground resistance of the transformer [29]. In this context, different scenarios have been carried out for the transformer link assembly and earth-to-ground resistance.

3.1 Dyn Transformer Connection and 0 Ohm Grounding Resistance

When the output results are analyzed with transformer connection section Dyn, ground resistance 0 ohms, very high short circuit currents are observed on the side where wind turbines are connected. This can cause a disruptive effect on the protection elements and can cause transformers to overheat. Tab. 1 shows the values that occur when turbines are engaged and turbines are not engaged in all bars.

Table 1 Phase-to-ground short circuit values for Dyn transformer connection in the modeled system

Fault Point	WPP On		WPP Off	
	I_k'' (kA)	S_k'' (MVA)	I_k'' (kA)	S_k'' (MVA)
Dyn Connection				
WPP 34.5 kV	4.38	87.67	4.26	84.12
Bus-Bar-1	3.72	74.89	3.61	71.22
Bus-Bar-2	4.19	83.41	4.17	81.81
Bus-Bar-3	2.61	50.02	2.42	48.47
Bus-Bar-4	2.71	53.61	2.74	52.63
Bus-Bar-5	3.02	59.45	2.97	58.74
Bus-Bar-6	3.16	63.45	3.08	62.60
Bus-Bar-7	3.72	74.15	3.68	73.01
Bus-Bar-8	3.98	78.02	3.86	76.45
Bus-Bar-9	4.06	80.72	3.97	79.05
Bus-Bar-10	30.22	73.15	28.56	71.86
Bus-Bar-11	31.45	74.78	29.74	72.63
Bus-Bar-12	31.89	75.02	29.98	73.33
Bus-Bar-13	32.72	76.15	30.53	74.32
Bus-Bar-14	32.86	76.89	30.74	75.09
Bus-Bar-15	33.32	77.62	31.40	76.02
Bus-Bar-16	34.23	79.75	32.45	77.60
Bus-Bar-17	35.12	81.63	33.42	80.03
Bus-Bar-18	37.85	83.84	35.76	81.09
Bus-Bar-19	38.48	85.47	36.23	83.78

3.2 YNyn Transformer Connection and 0 Ohm Grounding Resistance

The phase-to-ground short circuit current was analyzed, which was used by changing the transformer connection pattern, but was kept constant at 0 ohms. The fault will be supplied by both the mains and wind turbines as the connection is changed to. This is an important role for the protection elements on both sides. Phase-to-ground short circuit analysis of all the barges of the modeled power system has been analyzed as 0 ohm transformer connection and ground resistance of the YNyn.

Table 2 Phase-to-ground short circuit values for YNyn transformer connection and 0 ohm grounding resistance in the modeled system

Fault Point	WPP On		WPP Off	
	I_k'' (kA)	S_k'' (MVA)	I_k'' (kA)	S_k'' (MVA)
YNyn Connection	4.68	87.91	4.56	85.12
WPP 34.5 kV				
Bus-Bar-1	3.79	75.85	3.71	73.23
Bus-Bar-2	4.39	84.42	4.27	82.85
Bus-Bar-3	2.81	52.12	2.72	50.45
Bus-Bar-4	2.91	54.71	2.84	53.29
Bus-Bar-5	3.22	60.45	2.99	58.74
Bus-Bar-6	3.23	64.55	3.18	62.64
Bus-Bar-7	3.98	75.85	3.78	73.06
Bus-Bar-8	4.09	79.92	3.98	76.47
Bus-Bar-9	4.15	81.52	4.06	79.14
Bus-Bar-10	33.22	74.35	30.56	72.87
Bus-Bar-11	34.55	74.68	31.74	72.73
Bus-Bar-12	35.89	75.12	33.38	73.13
Bus-Bar-13	36.72	78.55	34.63	76.22
Bus-Bar-14	36.86	78.69	34.44	76.19
Bus-Bar-15	37.52	79.52	36.50	77.12
Bus-Bar-16	37.83	80.65	36.65	78.69
Bus-Bar-17	39.12	82.53	38.52	80.23
Bus-Bar-18	39.85	84.74	38.86	82.19
Bus-Bar-19	40.48	86.37	40.23	84.98

When the output results were analyzed, similar results were obtained from the Dyn connection, but the short-circuit currents on the network side increased, even though wind plants were not activated. The activation of wind turbines has also seen the increase in short-circuit currents on both the network and the wind turbine side. This shows the values that occur in Tab. 2 when turbines are engaged and turbines are not engaged in all bars.

Table 3 Phase-to-ground short circuit values for YNyn transformer connection and 20 ohm grounding resistance in the modeled system

Fault Point	WPP On		WPP Off	
	I_k'' (kA)	S_k'' (MVA)	I_k'' (kA)	S_k'' (MVA)
YNyn Connection	4.68	87.91	4.56	85.12
WPP 34.5 kV				
Bus-Bar-1	3.79	75.85	3.71	73.23
Bus-Bar-2	4.39	84.42	4.27	82.85
Bus-Bar-3	2.81	52.12	2.72	50.45
Bus-Bar-4	2.91	54.71	2.84	53.29
Bus-Bar-5	3.22	60.45	2.99	58.74
Bus-Bar-6	3.23	64.55	3.18	62.64
Bus-Bar-7	3.98	75.85	3.78	73.06
Bus-Bar-8	4.09	79.92	3.98	76.47
Bus-Bar-9	4.15	81.52	4.06	79.14
Bus-Bar-10	0.03	0.01	0.03	0.01
Bus-Bar-11	0.03	0.01	0.03	0.01
Bus-Bar-12	0.03	0.01	0.03	0.01
Bus-Bar-13	0.03	0.01	0.03	0.01
Bus-Bar-14	0.03	0.01	0.03	0.01
Bus-Bar-15	0.03	0.01	0.03	0.01
Bus-Bar-16	0.03	0.01	0.03	0.01
Bus-Bar-17	0.03	0.01	0.03	0.01
Bus-Bar-18	0.03	0.01	0.03	0.01
Bus-Bar-19	0.03	0.01	0.03	0.01

3.3 YNyn Transformer Connection and 20 Ohm Grounding Resistance

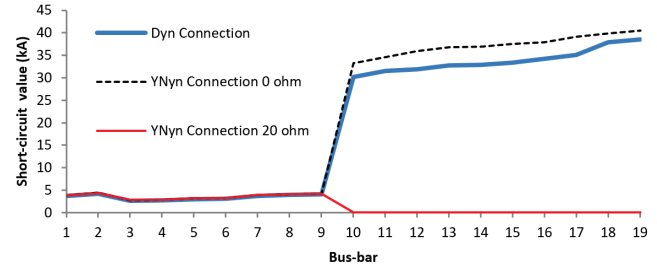
The phase-ground short-circuit current formed in the case when the transformer connection form is YNyn and the grounding resistance is 20 ohms has been studied. The three-phase short circuit analysis performed on all busbars of the

modeled power system was performed on the DigSILENT PowerFactory diagram.

The importance of grounding resistance has been revealed here and the short-circuit fault currents have decreased with the grounding resistance being 20 Ohms. In Tab. 3, this situation shows the values that occur when the turbines are in operation at all busbars and the turbines are not in operation.

4 FINDINGS

By using the grounding resistor in these three different cases made in the phase-ground short circuit, a serious decrease in fault currents was observed in the busbars connected to the wind turbines. It has become clear how necessary the use of grounding resistance in wind turbines included in the system is for protection. A graph showing how the connection method and the grounding resistance change during a phase-ground short-circuit failure, as well as the fault current and short-circuit power change, is shown in Fig. 1.

**Figure 1** Phase-to-ground short-circuit values according to transformer connection method

In Fig. 1, the vertical axis shows the short circuit values (kA) and the horizontal axis shows the bus-bar numbers. In order to ensure uninterrupted and quality energy, it is likely that phase-to-ground short circuit faults are limited by this method, and the most practical encounter will be avoided.

In the analyzes carried out the DigSILENT program, the short-circuit current of 4.68 kA in the WPP 34.5kV busbar with the highest short-circuit current was found when the secondary star point of the 250 MVA transformer was directly connected.

5 DISCUSSION AND CONCLUSION

The ohmic value of the resistor to be used for neutral grounding is determined by the voltage of the system and the magnitude of the default ground fault current. The star point is grounded through a 20 Ω resistor in the phase grounding short circuit of the transformer, the value of the current that will flow to the ground through the resistor is 995 A. If 40 Ω or 60 Ω is used, the phase grounding short circuit can be limited to lower current values. However, we are forced to limit it to this value, since the current that occurs in the primer is still not felt by the mechanical relays that continue to be used in our systems.

Since digital relays are used in developed countries, the short-circuit current is limited to 150 A and even up to 10 A with a Peterson coil in some countries. Thus, the opening times can be adjusted to a long time such as 2 hours. In the phase ground short circuit that will occur with advanced relay coordination, the breakers closest to the fault are activated and the defective region is isolated from the system. Thus, there will be no sudden stops in the system due to a malfunction that will be solved in a short time. In fact, the Peterson coil is the most effective of the neutral earths to be made against the phase ground short circuit. Because the Peterson coil compensates for the capacitive currents of the line and clears the fault. That is, the power factor of the system and the power factor of the malfunction are the same, and the malfunction is easily eliminated at the time of switching the current from zero. Especially in systems with a high concentration of small powerful generators, a resistor is connected to the Peterson coil in series. The purpose of this is to help the generators return to their old regimes after the fault has been cleared in the phase ground short circuit that has occurred.

In the phase ground short circuit that will occur in the system when the generator is operating parallel with the network, it must prevent the generator from being affected by this fault and going into island mode so we first limited the phase ground fault current that could occur in the system to a certain value using neutral ground resistance or the Peterson coil. The short circuit current has an active and a reagent component. Therefore, the active component of the short circuit current can be limited to the resistance and the reagent component Peterson coil to be connected in series. We have noticed that using this kind of neutral ground is more effective in the medium voltage networks that the Generators are connected to. The faulty area was then compressed and isolated from the system, thanks to the relay and protection circuits (50 N or 51 N), which will then be used according to the system specification, depending on whether radial or ring is a main.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

6 REFERENCES

- [1] Sun, T., Chen, Z., & Blaabjerg, F. (2005). Transient stability of DFIG wind turbines at an external short-circuit fault. *Wind Energy*, 8(3), 345-360. <https://doi.org/10.1002/we.164>
- [2] Siano, P., Chen, P., & Chen, Z. (2006). Optimal allocation of wind turbines in active distribution networks by using multi-period optimal power flow and genetic algorithms. *Wang L (ed) Modeling and control of sustainable power systems, green energy and technology*, 249-268. Berlin: Springer. https://doi.org/10.1007/978-3-642-22904-6_9
- [3] Muljadi, E., Butterfield, C., Chacon, J., & Romanowitz, H. (2006). Power quality aspects in a wind power plant. *IEEE power engineering society general meeting*. <https://doi.org/10.1109/PES.2006.1709244>
- [4] Li, J., Samaan, N., & Williams, S. (2008). Modeling of large wind farm systems for dynamic and harmonics analysis. *Proceedings of the IEEE/PES transmission and distribution conference and exposition*, 21-24. <https://doi.org/10.1109/TDC.2008.4517193>
- [5] Jiang, H., Zhang, C., Zhou, T., Zhang, Y., & Zhang, F. (2019). An adaptive control strategy of crowbar for the low voltage ride-through capability enhancement of DFIG. *Energy Procedia*, 601-606. <https://doi.org/10.1016/j.egypro.2019.01.161>
- [6] Beltran-Pulido, A., Cortes-Romero, J., & Coral-Enriquez, H. (2018). Robust active disturbance rejection control for LVRT capability enhancement of DFIG-based wind turbines. *Control Engineering Practice*, 77, 174-189. <https://doi.org/10.1016/j.conengprac.2018.06.001>
- [7] Shah, R., Mithulananthan, N., & Lee, K. (2013). Large-scale PV plant with a robust controller considering power oscillation damping. *IEEE Transactions on Energy Conversion*, 28(1), 106-116. <https://doi.org/10.1109/TEC.2012.2230328>
- [8] Hasanvand, H., Arvan, M., Mozafari, B., & Amraee, T. (2016). Coordinated design of PSS and TCSC to mitigate interarea oscillations. *International Journal of Electrical Power & Energy Systems*, 78(1), 194-206. <https://doi.org/10.1016/j.ijepes.2015.11.097>
- [9] Kerahroudi, S., Alamuti, M., Li, F., Taylor, G., & Bradley, M. (2014). Application and requirement of DlgSILENT PowerFactory to MATLAB/Simulink interface. *Gonzalez-Longatt FM. Powerfactory Applications for Power System Analysis*, 297-322. https://doi.org/10.1007/978-3-319-12958-7_13
- [10] Pourbeik, P., Sanchez-Gasca, J., Senthil, J., Weber, J., & Zadehkhosht, P. (2017). Generic dynamic models for modeling wind power plants and other renewable technologies in large-scale power system studies. *IEEE Transactions on Energy Conversion*, 32(3), 1108-1116. <https://doi.org/10.1109/TEC.2016.2639050>
- [11] Wang, S., Sun, Y., Huang, Z., & Mu, S. (2018). Analysis of stator internal phase-to-phase short circuit in the 12 phase synchronous generator with rectifier-load system. *IEEE Transactions on Energy Conversion*, 33, 299-311. <https://doi.org/10.1109/TEC.2017.2748147>
- [12] Rajveer, M. (2009). Low Voltage Ride-Through (LVRT) of Grid Interfaced Wind Driven PMSG. *ARPJ Journal of Engineering and Applied Sciences*, 4, 73-83.
- [13] Xin, Y. (2010). Application of the Interface Technique Between PSCAD/EMTDC and Matlab in Power System Simulation. *Journal of Electric Power*, 25, 214-217.
- [14] Berger, R. (2017). Global consulting. fuel cells and hydrogen applications for regions and cities. *Consolidated Technol Introduction Dossiers*, 1, 108-123.
- [15] IRENA, L. M. (2014). Ocean energy: technology readiness, patents, deployment status and outlook. *Germany: Int. Renew. Energy Agency IRENA*.
- [16] Chitchyan, R. (2016). Sustainability design in requirements engineering: state of practice. *The 38th IEEE International Conference on Software Engineering Companion*, 533-542. <https://doi.org/10.1145/2889160.2889217>
- [17] Abu-Rub, H., Malinowski, M., & Al-Haddad, K. (2014). Power Electronics for Renewable Energy Systems. *Transportation and Industrial Applications*, 534-535. <https://doi.org/10.1002/9781118755525>
- [18] Petersen, L., Lov, F., Shahid, K., Olsen, R., Altin, M., & Hansen, A. (2016). Voltage control support and coordination between ReGen plants in distribution systems. *WP2 Deliverable D2*.

- [19] Obikwelu, C., & Meliopoulos, S. (2019). CT instrumentation channel error correction using dynamic state estimation. *North American Power Symposium (NAPS)*. <https://doi.org/10.1109/NAPS46351.2019.9000191>
- [20] Chen, Z., Zhang, Z., Zhao, J., Wu, B., & Huang, X. (2018). An analysis of the charging characteristics of electric vehicles based on measured data and its application. *IEEE Access* 6, 24475-24487. <https://doi.org/10.1109/ACCESS.2018.2835825>
- [21] Develder, C., Sadeghianpourhamami, N., Strobbe, M., & Refa, N. (2016). Quantifying flexibility in EV charging as DR potential: analysis of two real-world data sets. *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 600-605. <https://doi.org/10.1109/SmartGridComm.2016.7778827>
- [22] Ringkjøb, H., Haugan, P., Solbrekke, I., Zürich, E., & Pfenninger, S. (2018). A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renew Sustain Energy*, 96, 440-459. <https://doi.org/10.1016/j.rser.2018.08.002>
- [23] Deason, W. (2018). Comparison of 100% renewable energy system scenarios with a focus on flexibility and cost. *Renew Sustain Energy*, 82, 3168-3178. <https://doi.org/10.1016/j.rser.2017.10.026>
- [24] Brown, T., Bischofniemz, T., Blok, K., Breyer, C., Lund, H., & Mathiesen, B. (2018). Response to Burden of proof: a comprehensive review of the feasibility of 100% renewable electricity systems. *Renew Sustain Energy*, 92, 834-847. <https://doi.org/10.1016/j.rser.2018.04.113>
- [25] Subudhi, B., & Pradhan, R. (2013). A comparative study on maximum power point tracking techniques for photovoltaic power systems. *IEEE Trans. Sustain. Energy*, 4(1), 89-98. <https://doi.org/10.1109/TSTE.2012.2202294>
- [26] Kakilli, A. (2013). System analysis with the MVA Method for symmetrical three-phase faults. *TEM Journal*, 51-56.
- [27] Sera, D., Teodorescu, R., Hantschel, J., & Knoll, M. (2008). Optimized maximum power point tracker for fast-changing environmental conditions. *IEEE Trans. Ind. Electron*, 55(7), 2629-2638. <https://doi.org/10.1109/TIE.2008.924036>
- [28] Şen, M., & Özcan, M. (2020). Implementation of Simulation of Possible Short Circuit Fault Situations in Wind Energy Plants by Power Analysis Program. *Avrupa Bilim ve Teknoloji Dergisi*, 196-201.
- [29] Li, R., Wang, W., Chen, Z., Jiang, J., & Zhang, W. (2017). A Review of Optimal Planning Active Distribution System: Models, Methods, and Future Researches. *Energies*, 10(11), 1715. <https://doi.org/10.3390/en10111715>

Authors' contacts:

Mehmet Şen, PhD Student of Electric Electronic Engineering,
(Corresponding author)
Necmettin Erbakan University,
Faculty of Engineering,
Department of Electric Electronic Engineering,
Konya, Turkey
mehmet.sen@asbu.edu.tr

Muciz Özcan, Prof. Dr.
Necmettin Erbakan University,
Faculty of Engineering,
Department of Electric Electronic Engineering,
Konya, Turkey
mozcan@erbakan.edu.tr

A Lévy Flight Based BAT Optimization Algorithm for Block-based Image Compression

Ilker Kilic

Abstract: Many metaheuristics have been adopted to solve the codebook generation problem in image processing. In this paper, the Bat algorithm is combined by the Lévy flight distribution to find out the global optimum codebook. The Lévy flight distribution is combined by the local search procedure. Therefore most of the time the bat concentrate on the local area for specific food while it rarely flies to the different parts of the field for better food opportunities. This process strongly guides the bat on the global minimum way and offers better food, then the bat flies to that direction. Consequently, if a bat is captured by a local minimum point accidentally, the Lévy flight step provides a chance to escape from it easily. Numerical results suggest that the proposed Lévy flight based Bat algorithm is better than the classical ones and provides the global optimum codebook for image compression.

Keywords: Bat algorithm; image compression; Lévy distribution; metaheuristic algorithm; vector quantization

1 INTRODUCTION

Vector quantization (VQ) is a block-based lossy image compression technique combined with a codebook generation algorithm [1-4]. A codebook contains vector blocks that represent the image best. The LBG [5] technique is a well-known codebook generation algorithm that satisfies local minimum solutions.

On the other hand, different metaheuristic algorithms have been developed and combined by VQ to overcome the optimum codebook generation problem. The optimum codebook is produced by the Genetic algorithm (GA) based on Principle Component Analysis (PCA) conversion [6]. The Particle Swarm Optimization (PSO) technique is improved by the advantages of the adaptive fuzzy inference method for codebook generation [7]. Honey Bee Mating Optimization (HBO) [8] algorithm is adapted to obtain the near-global optimal codebook. Improved Differential Evolution (IDE) [9] technique, which is an improved version of the traditional DE with modifications in the scaling factor and the boundary control mechanism, is used for codebook design. Ant Colony Optimization is redesigned for VQ by decreasing both computation number and the time required [10]. Firefly algorithm (FRA) is adapted to LBG and achieved faster and higher quality reconstructed images than PSO, HBO and LBG [11]. Recently Cuckoo search algorithm (CSA) is generated for optimizing the LBG codebook [12]. The Cuckoo search algorithm spends 25% of convergence time for local search while it spends 75% of convergence time for the global optimum solution. The Bat algorithm (BAT) is a quite new optimization technique that is used for optimum solutions [13-18]. Lévy flight distribution is combined with many optimization algorithms for global optimization [19-35].

In this paper, the classical BAT metaheuristic technique is improved by Lévy flight distribution to solve the codebook generation problem. When the bats are searching for food by locally near to the global solution, the Lévy flight mechanism helps the bat both focus on the local and random search at the same time. Lévy flight food search deals with the local search mostly and random search rarely.

2 RELATED ALGORITHMS

In this section, LBG algorithm of VQ and the Lévy algorithm is presented briefly.

2.1 LBG Algorithm and Vector Quantization

Vector quantization (VQ) is a block-based lossy image compression technique. The codebook production is the fundamental process of VQ. Let input image $Y = \{x_{ij}\}$ be $N \times N$ pixels size. The original image is formed by sub-blocks with the size of $m \times m$ pixels. The dimension of the sub-blocks are defined by $N_b = \left(\frac{N}{m} \times \frac{N}{m}\right)$ and represented by a collection of original image vectors expressed as $X = \{x_i, i = 1, 2, \dots, N_b\}$. Let the variable L be assumed as $m \times m$ pixel size. Then the original sub-blocks x_i are determined as L dimensional Euclidean space $x_i \in \mathcal{R}_L$. The number of N_c codewords belong to codebook named $C = \{C_1, C_2, \dots, C_{N_c}\}$, $j = 1, 2, \dots, N_c$. The original image vectors are determined by row vectors of $x_i = (x_{i1}, x_{i2}, \dots, x_{iL})$ and the i^{th} code word is defined as $c_i = (c_{i1}, c_{i2}, \dots, c_{iL})$. The codeword which has the minimum error is assigned to each original image block. So, the block number from the codebook list is used instead of the real image block to obtain a lossy image compression. Composing the vectors of C is achieved by minimizing the mean square error (MSE) formulation Eq. (1) to Eq. (3).

$$MSE(C) = \frac{1}{N_b} \sum_{j=1}^{N_c} \sum_{i=1}^{N_b} \mu_{ij} \|x_i - c_j\|^2 \quad (1)$$

$$\sum_{i=1}^{N_c} \mu_{ij} = 1, i \in \{1, 2, \dots, N_b\} \quad (2)$$

$$\mu_{ij} = \begin{cases} 1, & \text{if } x_i \text{ is in the } j^{\text{th}} \text{ cluster} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The Euclidean distance between the x_i image block and codeword c_j is defined by $\|x_i - x_j\|$. The LBG algorithm executes the two rules given by Eq. (4) and Eq. (5). The group of image vectors ($R_j, j = 1, 2, \dots, N_c$) must satisfy

$$R_j \supset \{x \in X : d(x, c_j) < d(x, c_k), \forall k \neq j\} \quad (4)$$

The center of the named as c_j is calculated as

$$c_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i, \quad x_i \in R_j \quad (5)$$

Here, N_j is the total number of elements that belong to R_j . Assume that the image vectors are $x_i, i = 1, 2, \dots, N_b$. Euclidean distance is d , and initial codewords are $c_j(0), j = 1, 2, \dots, N_c$. Then the LBG technique achieves the following three steps to obtain the local optimal codebook;

Divide the original image blocks into several clusters using Euclidean distance equation. The center of clusters is saved in an indicator matrix named U size of $N_b \times N_c$ pixels.

$$\mu_{ij} = \begin{cases} 1, & \text{if } d(x_i, c_j(k)) = \min(x_i, c_j(k)) \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

Determine the new cluster centers by Eq. (7).

$$c_j(k+1) = \frac{\sum_{i=1}^{N_b} \mu_{ij} x_i}{\sum_{i=1}^{N_b} \mu_{ij}}, \quad j = 1, 2, \dots, N_c \quad (7)$$

Until there is no change of centroid c_j value execute the Eq. (6) and Eq. (7) sequentially.

2.2 BAT Algorithm

The bat-inspired metaheuristic algorithm is produced by Xin-She Yang [13] based on imitation of the echolocation feature of microbats. The microbats use echolocation between 25 kHz to 125 kHz, which means 2 mm to 14 mm wavelengths in the air. Microbats searches prey using this echolocation feature. In the search phase microbats emit short pulses, but when they approach a potential prey their pulse emitting frequencies increase, wavelengths of echolocation become shorter. Therefore, the microbats increase the accuracy of the prey location. The standard bat algorithm uses three rules; i. All microbats use echolocation and detect the distance to the potential food. ii. Microbats fly

randomly with the parameters of V_i velocity, X_i position, F_i frequency, λ_i wavelength and A_i loudness. They can adjust the rate of pulse emission $r \in [0, 1]$ according to the distance to the food. iii. The loudness varies from $A_{\min}$ to A_0 a microbat has two flying types which are local search or generation of or random search.

Generation of new solutions;

When a bat begins to search for prey it flies randomly following the three steps defined by Eq. (8) to Eq. (10). Here the frequency, velocity and position of the bats are updated.

$$F_i^t = F_{\min} + (F_{\max} - F_{\min})u_i \quad (8)$$

$$V_i^t = V_i^{t-1} + (X_i^{t-1} - X_g)F_i^t \quad (9)$$

$$X_i^t = X_i^{t-1} + V_i^t \quad (10)$$

In the equations above, the frequency boundaries are defined as $(F_{\min}, F_{\max}) = (0 \text{ Hz}, 1 \text{ Hz})$, the uniform random variable $u_i \in [0, 1]$. $F_{\min}$ and $F_{\max}$ are the user-defined parameters and can be changed according to optimization problem type. X_i and V_i are the positions and the velocities of a bat, where the X_g denotes the global best position. The frequencies of the microbats are determined randomly at the beginning of the algorithm.

Local flying;

When a bat detects a prey specifically then the local flying strategy is used. This procedure is determined by Eq. (11) and Eq. (12).

$$\text{if } (U_k \geq R_k), \text{ then } X_i^t = X_i^{t-1} + \varepsilon_k A_{ort}^{t-1} \quad (11)$$

$$\text{if } (U_k < R_k), \text{ then } X_i^t = X_i^{t-1} \quad (12)$$

Here $U_k \in [0, 1]$ and $\varepsilon_k \in [-1, 1]$ are uniform random variables, A_{ort} is the average loudness of all the bats and R_k is the pulse rate of the k^{th} bat. Once a bat found its prey, the loudness usually decreases and the rate of pulse emission increases in order to determine true location of the prey. Therefore Eq. (12) is used more frequently than Eq. (11) when the bat is close to its prey.

Parameter updating;

After a local or random search flying is completed the pulse rate R_k and the loudness A_i parameters of the bats are updated using Eq. (13) and Eq. (14).

$$A_i^{t+1} = \alpha A_i^t \quad (13)$$

$$R_i^{t+1} = R_{\max} (1 - 0.5\gamma^t) \quad (14)$$

Here α is similar to the cooling factor of the cooling schedule in the simulated annealing [16]. Therefore both alpha and gamma parameters are changed slowly depending on the iteration numbering which is shown in Fig. 1. Therefore both of them are determined as 0.997 for all iterations of the codebook generation method. The loudness reaches zero and the pulse rate variable reaches the final value of one by the end of the iteration number as shown in Fig. 1. The procedure of updating velocities and positions of

the BAT algorithm is similar to the PSO algorithm. The frequency f_i controls the step and range of the bat movement. The local search is controlled by the loudness and the pulse rate variables.

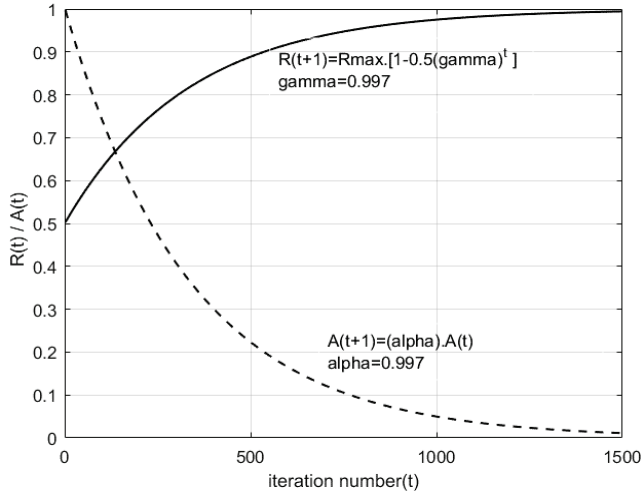


Figure 1 Loudness and pulse rate distribution according to the iteration number t

3 LEVY FLIGHT BASED BAT ALGORITHM

The BAT optimization contains two search mechanisms. The first one is the generation of the solution step. The other one is the local search focusing on the prey. If the bat is not near to the global optimum solution when it is on the random search, the equations of Eq. (9) and Eq. (10) may cause the bat captured by a local minimum. The local search performed by Eq. (11) may not provide an approximation step in the correct scale [17]. Therefore the standard BAT formulations are reorganized and simplified [18]. The first modification is a simplification of the bat location. The Eq. (15) is generated instead of Eq. (9) and Eq. (10). In Eq. (10) the velocity of the bat is disregarded.

$$X_i^{t+1} = X_i^t + (X_g - X_i^t)F_i^t \quad (15)$$

Adoption of Eq. (15) for updating the bat location instead of Eq. (9) and Eq. (10) allows the bats to fly closer to the prey. On the other hand Eq. (15) decreases the ability of the bat population to visit the around effectively. Therefore a Lévy flight distribution is added to the local search procedure which is given by Eq. (16).

$$X_i^{t+1} = X_i^t + (X_g - X_i^t)F_L(X) \quad (16)$$

Here $F_L(X)$ represents the Levy function, whose values are small at big X values. On the contrary the magnitude of $F_L(X)$ is relatively big at small X values. Therefore the bats search is close to global solution in most of the local search iterations. On the other hand the bats rarely fly to far away with the help of Levy flight in order to escape from possible local minimum location. Let $F(X)$ be a normal distribution.

$$F(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (17)$$

Let $\mu = 0$ and standard deviation σ be a function of x .

$$\sigma = \sqrt{\frac{x^3}{c}} \quad (18)$$

Replacing Eq. (18) into Eq. (17), normal distribution is converted to the Lévy distribution which is given by Eq. (19).

$$F_L(X) = \sqrt{\frac{c}{2\pi}} \cdot \left(\frac{e^{-\frac{c}{2x}}}{x^{3/2}} \right) \quad (19)$$

Exponential term decreases to one by increasing x , and the Lévy distribution becomes like a power law function with long tail, which has an infinite variance with an infinite mean.

$$F_L(X) \cong k \cdot x^{-\lambda} \quad (20)$$

The approximated Lévy function which is expressed by Eq. (20) is plotted in Fig. 2. The approximated Lévy distributions are shown for the values of $\lambda_1 = -1.0$, $\lambda_2 = -1.25$, $\lambda_3 = -1.5$, $\lambda_4 = -2.0$. The maximum value of $F_L(X)$ is determined by $k = 2.0$ value.

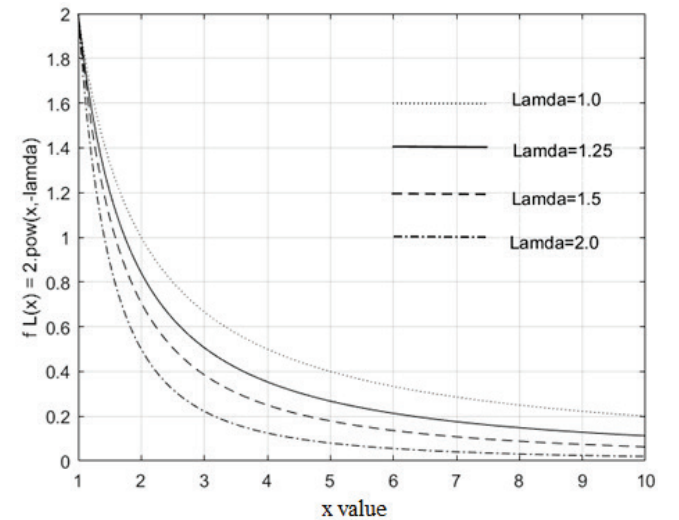


Figure 2 Approximated Lévy functions for different λ values

Lévy Flight based Bat algorithm (LBAT) can be obtained by using Lévy distribution function $F_L(X)$ in the bat random search procedure which is defined by Eq. (21). Therefore, if Eq. (20) is embedded into Eq. (16) the bat movement is determined as

$$X_i^t = X_i^{t-1} + k(X_g - X_i^{t-1})(\text{random } x)^{-\lambda} \quad (21)$$

where the *random* x is uniformly distributed in $[0, 1]$, k and λ are the user defined coefficients. Since the term $k(\text{random } x)^{-\lambda}$ is power law function, in most of the iterations its value is small whereas the value is rarely relatively big. Consequently, the second term of Eq. (21) is called Lévy Flight step guides to the bat escape from local minimum points. The variable λ is quite sensitive in Eq. (21) and must be properly tuned. The λ is selected as $\lambda = 2.0$ for searching the best codebook. The pseudo code of the proposed LBAT is given below.

Other modifications of LBAT are about the pulse rate R_i and loudness A_i . The original pulse rate equation of Eq. (14) and loudness equation of Eq. (13) are redesigned as in Eq. (22) and Eq. (23) respectively [18]. Instead of exponential function the R_i^t and A_i^t variables are changed into linear ones depending on the iteration number.

$$R_i^t = (R_0 - R_\infty) \left[\frac{t - t_{\max}}{1 - t_{\max}} \right] + R_\infty \quad (22)$$

$$A_i^t = (A_0 - A_\infty) \left[\frac{t - t_{\max}}{1 - t_{\max}} \right] + A_\infty \quad (23)$$

Here, $t_{\max}$ is the maximum iteration number while the notations 0 and ∞ corresponds the initial and the last values respectively. As previously mentioned, the resulting algorithm is named as LBAT after all these modifications. The pseudo code of the proposed LBAT is given below.

begin

Objective function $MSE(C)$ by Eq. (1)

Generate initial bat population $C = \{C_1, C_2, \dots, C_{N_C}\}$.

$C_j \in \mathbb{R}$, $j = 1, 2, \dots, N_C$

Define the Lévy Flight variables λ, k

Define the initial and the end values of pulse rate R_i and loudness A_i

Define iteration number n

while ($n < \text{Max Iteration Number}$)

 Define uniform random numbers $U_i \in [0, 1]$

 If ($U_i < R_i$) Local search by Levy Flight distribution as in Eq. (21)

 else Randomly search around by Eq. (11)

 Update the parameter of pulse rate R_i^t by Eq. (22)

 Update the parameter of loudness A_i^t by Eq. (23)

end

3.1 Parameters

In this paper, the LBG algorithm is adapted to the LBAT then each bat is represented by a codebook defined as $C = \{C_1, C_2, \dots, C_{N_C}\}$, $C_j \in \mathbb{R}$, $j = 1, 2, \dots, N_C$. The number of codewords are selected as 8, 16, 32, 64 and 128 for the 256 gray level and 256×256 pixel sized standard images. C-

Means variable of ε is defined as 0.001 for VQ. The user defined parameters k_w , k_p , and k_g of PSO are defined as 0.9, 0.9 and 0.9 respectively when the random_1 and random_2 are the uniform randoms change in $[0, 1]$. The α and γ parameters of BAT technique defined as 0.997 and 0.996 for image codebook generation. Objective functions of BAT and LBAT are directly proportional to inverse codebook error, which is determined as $MSE^{-1}(C)$. The variables α , γ and β_0 of FRA are determined as 0.5, 0.01 and 1.0 respectively. The coefficients of LBAT are selected as $\lambda = 1.5$ when $\gamma = 0.01$ and $k = 2.0$ for optimization. The iteration number of LBAT is determined as 1500 which is sufficient for LBAT algorithm to reach the global minimum point.

4 SIMULATIONS AND RESULTS

The codebook producing is achieved by using 256×256 pixel size of standard images, which have 256 gray levels. The standard images are divided into 4×4 pixel sized of 4096 subvectors. The standard FRA, BAT, PSO, VQ and LBAT algorithms are used for codebook generation. The variables of k and λ are important parameters of approximated Lévy Flight function $F_L(X)$. They are responsible for the convergence speed. Therefore they are fine-tuned as $k = 2.0$, $\lambda = 2.0$. The k determines the maximum Lévy Flight step value whereas the λ defines the percentile of the long Lévy Flight step rate as seen in Fig. 2. Small values of k and λ cause to be caught by a local minimum whereas the high values cause delay in the convergence time duration. When the bat is focusing on the local search the Lévy flight random step provides an opportunity of food search at the different fields of the search area. Therefore most of the time the bat consantrate on the local area for specific food while it rarely flies to the different part of the field for better food opportunity. The x variable interval is determined as $x \in [1, 7]$. Consequently, Lévy Flight steps change from $F_L(1) = 2.0$ to $F_L(7) = 0.04$. The bat steps of LBAT algorithm change generally smaller than 0.5 but rarely bigger than 0.5. The long steps provide the bat easily escaping from possible local minimum whereas small steps help bat focusing to the food. The mean square error $MSE(C)$ performance values for the codebooks of different and proposed techniques are given in Tabs. 1-3. The number of 8, 16, 32, 64 and 128 codewords are used in the simulation. In addition to these results the error convergence graphics of the algorithms according to MSE versus iteration number are shown in Fig. 3 and Fig. 5 for the high contrast 256×256 pixels size Barbara and low contrast 256×256 pixel size Clock standard images.

Table 1 MSE performances of codebooks of different techniques for Barbara image

	Codebook Size	VQ	PSO	FRA	BAT	LBAT
Barbara	8	357.9	343.3	341.7	351.0	338.6
	16	259.9	250.0	247.1	256.0	244.0
	32	197.5	189.8	186.6	197.1	183.2
	64	157.1	149.5	145.5	153.3	141.9
	128	125.8	120.7	118.8	127.7	115.2

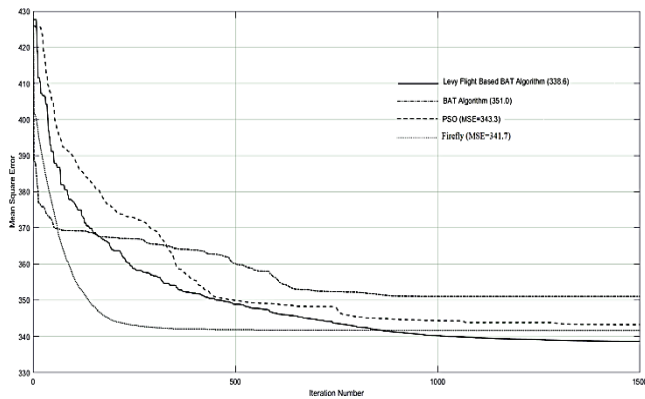


Figure 3 Convergence results of PSO, FRA, BAT and LBAT algorithms for BARBARA image and 8 codewords of codebook



Figure 4 Visual results of LBAT, BAT, FRA, PSO and VQ algorithms for BARBARA image for 8 codewords of codebook

Table 2 MSE performances of different techniques for Clock image

	Codebook Size	VQ	PSO	FRA	BAT	LBAT
Clock	8	291.3	287.4	280.0	304.4	271.2
	16	198.9	195.0	189.7	205.5	179.7
	32	143.5	141.0	138.8	149.9	133.3
	64	125.7	123.8	120.1	131.0	117.0
	128	94.8	92.5	90.2	96.7	88.8

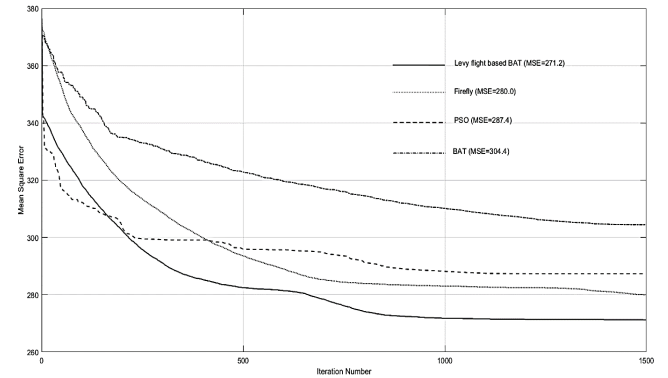


Figure 5 Convergence results of PSO, FRA, BAT and LBAT algorithms for CLOCK image and 8 codewords of codebook

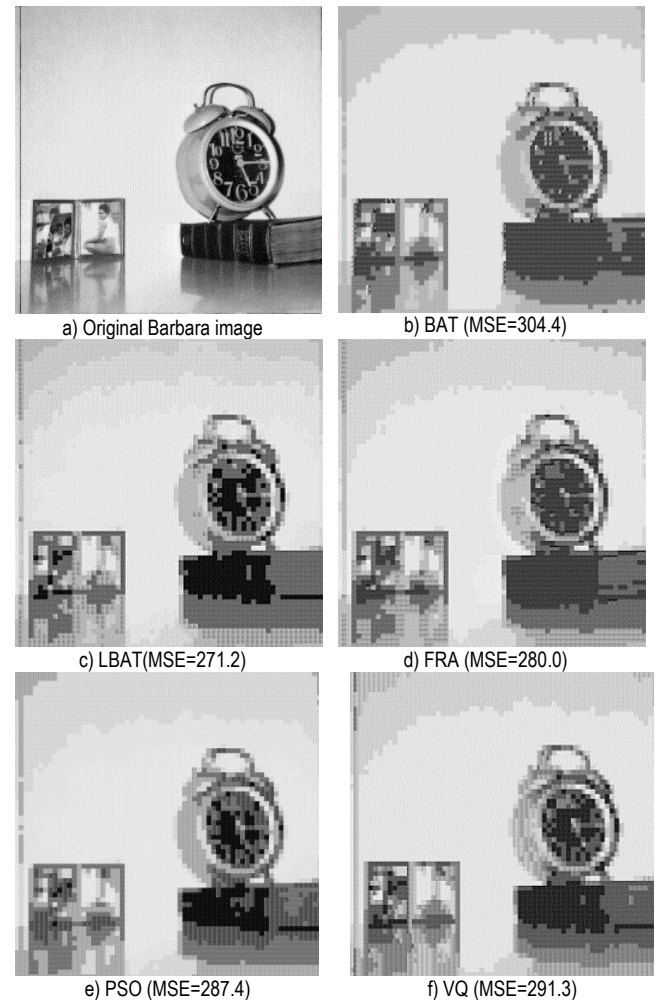


Figure 6 Visual results of LBAT, FRA, BAT and PSO algorithms for CLOCK image for 8 codewords of codebook

The results of other standard images of 256×256 pixel size are given in Tab. 3.

It is seen from Tabs. 1-3 and Figs. 3-6 that the proposed LBAT codebook MSE performances are better than both classical techniques and metaheuristic algorithms. On the other hand LBAT, FRA, BAT and PSO classical metaheuristic algorithms are superior to the VQ technique. The proposed LBAT algorithm performs better and reaches

to the global optimum codebook whereas VQ technique is seem to be captured local minimums. Consequently in the proposed LBAT, when the bats are on the local search they rarely search around by Lévy Flight, which is a small random step in most of the iterations but rarely a big random step. Therefore this mechanism protects a bat to be captured by a possible local minimum and also leads to it on the way of global minimum solution.

Table 3 MSE performances of different techniques for standard images

Compression Ratio 0.122 bpp	VQ	PSO	FRA	BAT	LBAT
Peppers	332.4	330.1	323.9	339.0	320.2
Clock	291.3	287.4	280.0	304.4	271.2
Airplane	122.8	122.5	121.1	131.0	118.5
Boat	376.7	372.8	373.3	383.1	363.2
Chemical P.	369.9	362.2	364.0	376.8	355.9
Cameraman	416.6	410.0	411.3	422.7	405.4
Barbara	357.9	343.3	341.7	351.0	338.6
Einstein	242.1	238.8	240.1	249.6	229.2
Couple	312.8	308.1	309.7	319.2	303.0
Lena	296.3	293.6	322.6	303.3	286.3
Moon	147.7	142.2	145.6	154.0	137.4
Baboon	433.3	427.7	432.3	440.4	418.7
Aerial	532.2	526.6	528.5	540.3	518.4

5 CONCLUSION

In this paper, the BAT metaheuristic algorithm is improved by using Lévy Flight distribution for codebook generation in image processing. In the proposed LBAT, besides the local search for the global solution, a special random Lévy Flight step opportunity is given to the bat. Lévy Flight satisfies to a bat an opportunity of special random movement, which is generally small random steps in most of the iterations but rarely a big one. Therefore this mechanism protects a bat to be captured by a possible local minimum and also leads to it on the way of global minimum solution. Numerical results suggest that the proposed LBAT is better than both the classical techniques and the metaheuristic algorithms and provides the global optimum codebook.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

6 REFERENCES

- [1] Gray, R. M. (1984). Vector quantization. *IEEE Assp Magazine*, 4-29. <https://doi.org/10.1109/MASSP.1984.1162229>
- [2] Hu, Y. C., Chen, W. L., et al. (2012). Improved vector quantization scheme for grayscale image compression. *Opto Electronics Review*, 20(2), 187-193. <https://doi.org/10.2478/s11772-012-0016-z>
- [3] Chuang, J. C., Hu, Y. C., et al. (2013). Improved Mean-Removed Vector Quantization Scheme for Grayscale Image Coding. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(5), 315-332. <https://doi.org/10.14257/ijsp.2013.6.5.28>

- [4] Zhang, L., Zhang, M., et al. (2014). An Effective Image Coding Method using Lattice Vector Quantization in Wavelet Domain. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 7(2), 305-316. <https://doi.org/10.14257/ijsp.2014.7.2.28>
- [5] Chang, C. C. & Hu, Y. C. (1998). A fast codebook training algorithm for vector quantization. *IEEE Transactions on Consumer Electronics*, 44, 1201-1208. <https://doi.org/10.1109/30.735818>
- [6] Sun, H., Kwok-Yan Lam, et al. (2005). Efficient vector quantization using genetic algorithm. *Neural Computing and Application*, 14, 203-211. <https://doi.org/10.1007/s00521-004-0455-7>
- [7] Feng, H. M., Chen, C. Y., et al. (2007). Evolutionary fuzzy particle swarm optimization vector quantization learning scheme in image compression. *Expert Systems with Applications*, 32, 213-222. <https://doi.org/10.1016/j.eswa.2005.11.012>
- [8] Horng, M. H. & Jiang, T. W. (2011). Image vector quantization algorithm via honey bee mating optimization. *Expert Systems with Applications*, 38, 1382-1392. <https://doi.org/10.1016/j.eswa.2010.07.037>
- [9] Nag, S. (2019). Vector quantization using the improved differential evolution algorithm for image compression. *Genetic Programming and Evolvable Machines*, 20, 187-212. <https://doi.org/10.1007/s10710-019-09342-8>
- [10] Tsaia, C. W., Tsengb, S. P., et al. (2013). PREACO: A fast ant colony optimization for codebook generation. *Applied Soft Computing*, 13, 3008-3020. <https://doi.org/10.1016/j.asoc.2013.01.017>
- [11] Rani, M. L. P., Rao, G. S., et al. (2020). An efficient codebook generation using firefly algorithm for optimum medical image compression. *Journal of Ambient Intelligence and Humanized Computing*, 1-13. <https://doi.org/10.1007/s12652-020-01782-w>
- [12] Chiranjeevi, K. & Jena, U. R. (2018). Image compression based on vector quantization using cuckoo search optimization technique. *Ain Shams Engineering Journal*, 9, 1417-1431. <https://doi.org/10.1016/j.asej.2016.09.009>
- [13] Yang, X. S. & Gandomi, A.H. (2012). Bat algorithm: a novel approach for global engineering *Engineering Computations: International Journal for Computer-Aided Engineering and Software optimization*, 29(5), 464-483. <https://doi.org/10.1108/02644401211235834>
- [14] Adarsh, B. R., Raghunathan, T., et al. (2016). Economic dispatch using chaotic bat algorithm. *Energy*, 96, 666-675. <https://doi.org/10.1016/j.energy.2015.12.096>
- [15] Jordehi, A. R. (2015). Chaotic bat swarm optimisation (CBSO). *Applied Soft Computing*, 26, 523-530. <https://doi.org/10.1016/j.asoc.2014.10.010>
- [16] Yang, X. S. (2014). *Nature-inspired optimization algorithms*. Elsevier. <https://doi.org/10.1016/B978-0-12-416743-8.00010-5>
- [17] Chakri, A., Khelif, R., et al. (2017). New directional bat algorithm for continuous optimization problems. *Expert Systems with Applications*, 69, 159-75. <https://doi.org/10.1016/j.eswa.2016.10.050>
- [18] Karri, C. & Jena, U. (2016). Fast vector quantization using a Bat algorithm for image compression. *Engineering Science and Technology, an International Journal*, 19, 769-781. <https://doi.org/10.1016/j.jestch.2015.11.003>
- [19] Wang, G., Guo, L., et al. (2013). Lévy-Flight Krill Herd Algorithm. *Mathematical Problems in Engineering*, 1-14, Article ID: 682073. <https://doi.org/10.1155/2013/682073>
- [20] Hakli, H. & Uguz, H. (2014). A novel particle swarm optimization algorithm with Levy flight. *Applied Soft Computing*, 23, 333-345. <https://doi.org/10.1016/j.asoc.2014.06.034>

- [21] Jensi, R. & Jiji, G. W. (2016). An enhanced particle swarm optimization with levy flight for global optimization. *Applied Soft Computing*, 43, 248-261. <https://doi.org/10.1016/j.asoc.2016.02.018>
- [22] Sharmaa, H., Bansal, J. C., et al. (2016). Lévy flight artificial bee colony algorithm. *International Journal of Systems Science*, 47(11), 2652-2670. <https://doi.org/10.1080/00207721.2015.1010748>
- [23] Xian, S., Kang, L., et al. (2016). Modified Bat Algorithm Based on Lévy Flight and Opposition Based Learning. *Hindawi Publishing Corporation Scientific Programming*, 1-13, Article ID 8031560. <https://doi.org/10.1155/2016/8031560>
- [24] Ling, Y., Zhou, Y., et al. (2017). Lévy Flight Trajectory-Based Whale Optimization Algorithm for Global Optimization. *IEEE Access*, 5, 6118-6186. <https://doi.org/10.1109/ACCESS.2017.2695498>
- [25] Amirsadri, S., Mousavirad, S. J., et al. (2018). A Levy flight-based grey wolf optimizer combined with back propagation algorithm for neural network training. *Neural Computing & Applications*, 30, 3707-3720. <https://doi.org/10.1007/s00521-017-2952-5>
- [26] Li, Y., Li, X., et al. (2019). An Improved Bat algorithm based on Lévy Flights and adjustment factors. *Symmetry*, 11, 925-944. <https://doi.org/10.3390/sym11070925>
- [27] Liu, Y. & Cao, B. (2020). A Novel ant colony optimization algorithm with Lévy Flight. *IEEE Access*, 8, 67205-67213. <https://doi.org/10.1109/ACCESS.2020.2985498>
- [28] Abdulwahab, H. A., Noraziah, A., et al. (2019). An enhanced version of Black Hole algorithm via Levy Flight for optimization and data clustering problems. *IEEE Access*, 7, 142085-142096. <https://doi.org/10.1109/ACCESS.2019.2937021>
- [29] Zhou, Y., Ling, Y., et al. (2018). Lévy flight trajectory-based whale optimization algorithm for engineering optimization. *Engineering Computations*, 35(7), 2406-2428. <https://doi.org/10.1108/EC-07-2017-0264>
- [30] Zhang, J. & Wang, J. S. (2020). Improved Salp Swarm Algorithm Based on Levy Flight and Sine Cosine Operator. *IEEE Access*, 8, 99740-99771. <https://doi.org/10.1109/ACCESS.2020.2997783>
- [31] Deotti, L. M. P., Pereira, J. L. R., et al. (2020). Parameter extraction of photovoltaic models using an enhanced Lévy flight bat algorithm. *Energy Conversion and Management*, 221, 113114. <https://doi.org/10.1016/j.enconman.2020.113114>
- [32] Zhao, J., Gao, Z. M., et al. (2020). The improved slime mould algorithm with Levy flight. *Journal of Physics: Conference Series*, 1617, 012033. <https://doi.org/10.1088/1742-6596/1617/1/012033>
- [33] Ingle, K. K. & Jatoth, R. K. (2020). An efficient JAYA algorithm with Lévy Flight for non-linear channel equalization. *Expert Systems with Applications*, 145, 112970. <https://doi.org/10.1016/j.eswa.2019.112970>
- [34] Iacca, G., Junior, V. C. S., et al. (2021). An improved Jaya optimization algorithm with Lévy flight. *Expert Systems with Applications*, 165, 113902. <https://doi.org/10.1016/j.eswa.2020.113902>

Author's contacts:

Ilker Kilic
Pamukkale University,
Electrical and Electronics Department,
Kinkli Campus, 20160, Denizli, Turkey
Tel: +90-0258-2963010
E-mail: ilkerk@pau.edu.tr

Automated Semantic Segmentation for Autonomous Railway Vehicles

Oğuzhan Katar*, Erkan Duman

Abstract: With the development of computer vision methods, the number of areas where autonomous systems are used has also increased. Among these areas is the transportation sector. Autonomous systems in the transportation sector are mostly developed for road vehicles, but highway rules and standards different between countries. In this study, models capable of semantic segmentation have been developed for autonomous railway vehicles with the help of the public dataset. Four different U-Net models were trained with 8500 images for four different scenarios. The model trained for binary semantic segmentation reached mean Intersection over Union (mIoU) value of 89.1%, while the models trained for multi-class semantic segmentation reached 83.2% mIoU, 79.7% mIoU and 29.6% mIoU. Information about the inclusion of high-resolution images in model training and performance metrics in semantic segmentation studies shared.

Keywords: autonomous systems; deep learning; railway vehicles; semantic segmentation; U-Net

1 INTRODUCTION

As a result of the increase in the processing capacity of computer hardware, there have been significant developments in artificial intelligence and computer vision [1]. Thanks to these developments, autonomous systems have become widespread. The main purpose of autonomous systems is to minimize the human factor. Due to the advanced decision-making mechanisms, autonomous systems do not need any direction during movement [2]. One of the most basic benefits of autonomous technologies is the opportunity to save time and energy in the area where it is used. One of the areas where time and energy saving is most needed is the transportation sector. For this reason, the use of autonomous systems in the field of transportation is inevitable [3].

Transportation is one of the important components of the world economy [4]. Because in order for the production to take place, first the raw material must be transported and then the product must be shipped after the production is completed [5]. It is desired that all these processes take place in a low cost and reliable way. Therefore, rail transport is preferred. Railway transportation is used not only for freight transport but also for passenger transfer. With all these features, railway transportation has an important potential for autonomous systems, but most studies are carried out by researchers for highways. Among the reasons for this is that the scope of public datasets is limited to highways.

The most vital function of autonomous systems in transportation is the ability to vision. With the help of a camera, the ability to vision by using pixel-based segmentation in the processing of the image can be brought to the system. In the segmentation process, instead of using classical methods, deep learning-based approaches should be used [6]. In this way, the system will ensure that the image is interpreted and the relevant action is taken in return, without the need for any human contribution. With such approaches, undesirable events that occur due to human factors such as inattention, fatigue, and thoughtlessness can be prevented.

The main purpose of this study is to provide vision capability to autonomous railway vehicles with deep

learning-based segmentation method. With its implementation in the field, it is aimed to increase efficiency and safety in railway transportation. The rest of this paper is organized as follows. Section 2 includes other studies in the literature on deep learning techniques for image segmentation. The details of the model, dataset and performance metrics is described in Section 3. The analysis of test results and experimental results of U-Net model are given in Section 4. Conclusion part of the study is in Section 5.

2 RELATED WORKS

In this section, some studies from computer vision in railway are examined. The main idea behind these studies is to contribute to railway transportation with various methods.

Edge detection methods are frequently used in computer vision studies for railway transportation due to the texture and shape of the rail-tracks [7] [8]. In addition to the edges extracted with edge detection, the Hough transform is applied [9]. In edge detection methods used with classical image processing, there may be factors such as the angle of light that will affect the removal of the relevant part and the determination of its area. In order to overcome such problems, deep learning-based approaches are used.

In [10], a model capable of deep learning-based railroad track segmentation has been proposed. This model, called RailNet, consists of a feature extraction network and a segmentation network. In order to train the specified model, a private dataset consisting of 3000 images was created by the researchers. This model, which was designed only for binary segmentation, reached a mIoU value of 89.8% as a result of the tests performed.

In [11], rail track regions are determined from the images taken with the help of unmanned aerial vehicles. The U-Net based model is trained for the binary segmentation task. The model, called Rail-UNet, reached 95.9% mIoU.

3 MATERIAL AND METHOD

3.1 Dataset

In this study, the dataset RailSem19 [12] was used. Most of the studies on gaining autonomous vision features of transportation vehicles cover road vehicles. A general reason for this is that the content of the datasets that researchers can use consists of only highway images. In order to overcome this deficiency in railway transportation studies, RailSem19 was made publicly available to researchers by the Austrian Institute of Technology employees in 2019.

RailSem19 dataset consists of 8500 unique high-resolution images taken from the machinist point of view of a railway vehicle and mask images prepared to be used in segmentation studies of these images. Samples of images and masks in the dataset are given in Fig. 1.

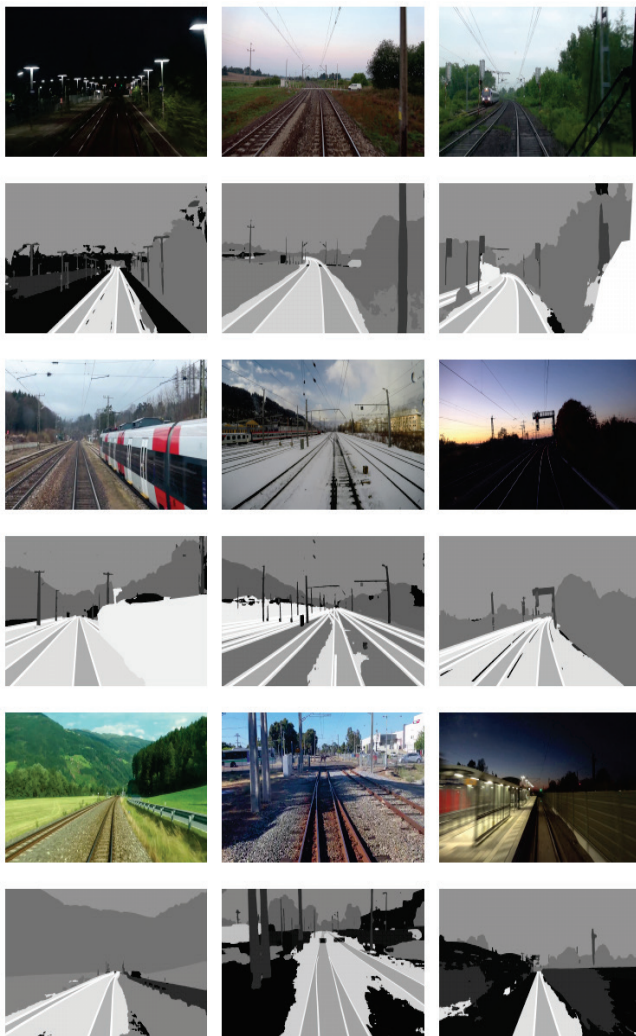


Figure 1 Samples of images and masks in RailSem19

There are 19 different labels in the mask images in the RailSem19 dataset. Each label represents a class, so class distribution calculations in the dataset can be made. These labels are represented by various pixel values in 8-bit and single-channel '.png' images. Pixel values, labels and

percentage of frames containing pixels with the label information are given in Tab. 1.

Table 1 Dataset pixel details

Label Name	Pixel Value	In Frames (%)
road	0	48.1
sidewalk	1	57.2
construction	2	72.1
tram-track	3	51.7
fence	4	48.9
pole	5	60.0
traffic-light	6	37.1
traffic-sign	7	32.3
vegetation	8	83.3
terrain	9	61.2
sky	10	94.5
human	11	6.0
rail-track	12	86.2
car	13	13.8
truck	14	4.6
trackbed	15	87.6
on-rails	16	15.4
rail-raised	17	87.2
rail-embedded	18	14.6

3.2 Pre-processing

Before defining and training the model, various pre-processes were applied to the RailSem-19 dataset samples for this study. The common objectives of the pre-processes applied are to use our resources more efficiently, to increase the training accuracy and time efficiency.

3.2.1 Creating Sub-datasets

Firstly, the dataset consisting of 8500 images and 8500 mask images was divided into four sub-datasets by preserving the image and mask unity. While performing the division process, it is important to determine the appropriate visuals for the scenarios they will represent, rather than separating them in equal numbers. The sub-datasets are named with the 'SubDS' tag and the detail of the sub-datasets is shown in Tab. 2.

Table 2 Sub-datasets details

Dataset Name	Image Count	Mask Count	Resolution
SubDS1	1870	1870	1920×1080 px
SubDS2	2850	2850	1920×1080 px
SubDS3	2250	2250	1920×1080 px
SubDS4	1530	1530	1920×1080 px

No random selection was made while creating the sub-datasets. This is because each dataset represents a scenario. Samples are separated according to the labels and distribution of pixel values that each scenario needs. Detailed information about the scenarios is as follows.

SubDS1 represents the scenario in which railway vehicles are able to segment only rail-tracks, and the dataset is designed for input to the binary segmentation model. In the relevant dataset, the pixels with the rail-track label in the mask images are given a value of '1'. Pixel values for all other labels are assigned '0'. SubDS1 sample and its mask image are given in Fig. 2.



Figure 2 SubDS1 dataset sample

SubDS2 is designed for the segmentation of rail-raised areas in addition to rail-track during the movement of railway vehicles. It is an example of multi-class segmentation scenarios that contains the least number of class labels. Because it consists of three labels in total. In this dataset, as in binary segmentation pixels with rail-track label are given the value '1', pixels with rail-raised label are given the value '2' and all other labels are given the value '0'. SubDS2 sample and its mask image are given in Fig. 3.

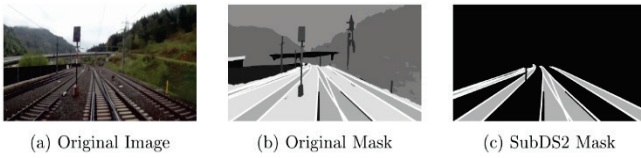


Figure 3 SubDS2 dataset sample

SubDS3 dataset is designed for scenarios where it is necessary to detect whether there is a person on the rail-track or rail-raised in autonomous driving. There are four different pixel values and labels in the dataset. As in SubDS2, pixels with rail-track label are represented by a value of '1', pixels with a rail-raised label are represented by a value of '2'. In addition, the pixels in which people are located are assigned the value of '3' and the value of the ones other than the specified labels is '0'. SubDS3 sample and its mask image are given in Fig. 4.

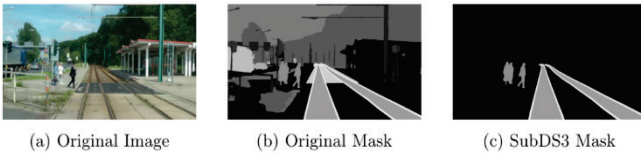


Figure 4 SubDS3 dataset sample

SubDS4 dataset has the default labels and pixel values of the RailSem-19 dataset. It is the most challenging multi-class segmentation example among the scenarios designed. SubDS4 sample and its mask image are given in Fig. 5.



Figure 5 SubDS4 dataset sample

3.2.2 Data Splitting

Considering the unity of the images and mask images in the datasets, they were randomly divided into 70% train, 20% validation and 10% test. Data splitting is shown in Fig. 6.



Figure 6 Data splitting method

3.2.3 Image Cropping

As it is known, the size of the dataset samples is 1920×1080 px. It is not a logical approach to input such high-resolution images directly into the model, considering our hardware features. Since the input size of the default U-Net model is 256×256 , resizing our images to this value may be a solution, but resize is not recommended in segmentation studies [13]. The details of the algorithm we applied to solve the related problem are as follows.

First, the reference image to be cropped is determined. 1920×1080 px image is cropped in a square format, starting from the first pixel, with edges on the x -axis of 256 pixels and the y -axis of 256 pixels. This process is repeated until it covers the size values of the image. A total of 40 256×256 cropped images can be created for 1920×1080 px images. In Fig. 7, the cropping process and numbering of rows and columns are given.



Figure 7 Image cropping method

When the first step is completed, the squares shown in white in Fig. 7. are smoothly cropped sub-images. The squares indicated in red there are pixels that cannot complete the 256×256 size. In order not to lose the information in these pixels, the values (0, 0, 0) were added with the padding methods. The result of the padding method is given in Fig. 8.

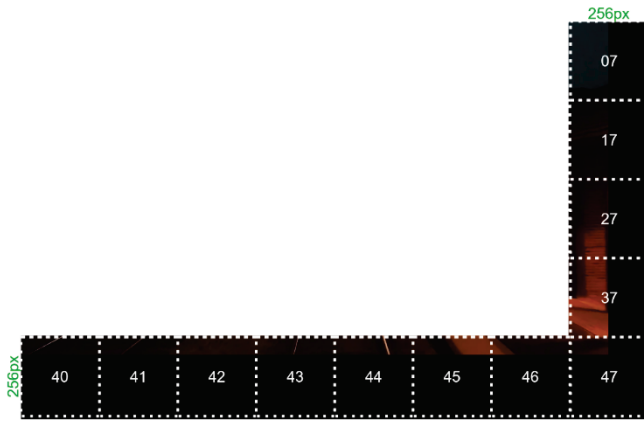


Figure 8 Result of padding method

The same steps applied for the images were applied for the mask images, and a sample mask image cropping output is given in Fig. 9.

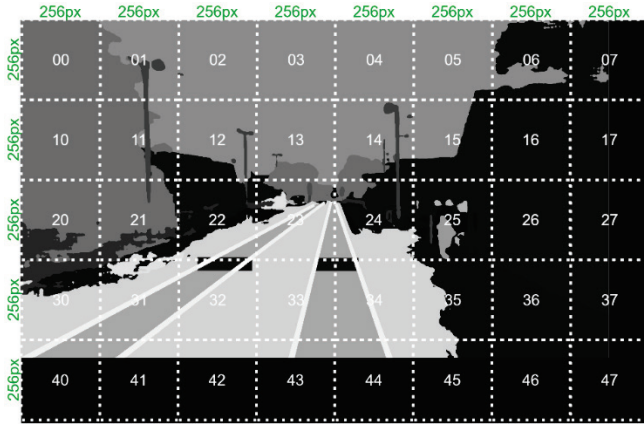


Figure 9 Image cropping method for masks

3.3 U-Net Model

U-Net is a kind of artificial neural network that contains a series of convolutional layers and non-convolutional layers to perform the image segmentation task. It is one of the most popularly used approaches to any semantic partitioning task today. U-Net gets its name from its U-like architecture, as seen in Fig. 10.

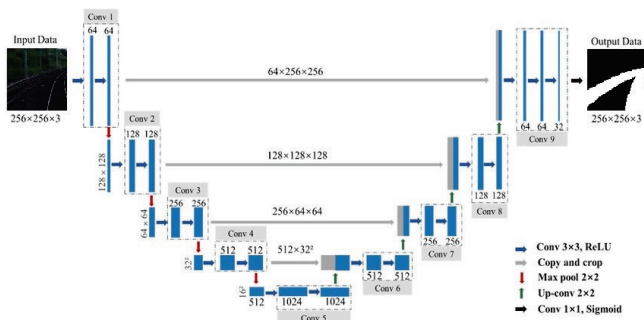


Figure 10 U-Net architecture

U-Net is an encoder-decoder network architecture consisting of four encoder blocks and four decoder blocks

connected via bridge. The encoder is designed to extract spatial features from the original image.

3.3.1 U-Net Encoder Network

The encoder network acts as a feature extractor and learns an abstract representation of the input image through a sequence of encoder blocks. Each encoder block consists of two 3×3 convolutions where each convolution is followed by a Rectified Linear Unit (ReLU) activation function. ReLU introduces non-linearity to the network, which helps to better generalize the training data. Then comes the 2×2 maximum pooling, in which the height and width of the feature maps are reduced by half. This reduces the calculation cost by reducing the number of trainable parameters.

3.3.2 U-Net Decoder Network

Decoder network is used to take the abstract representation and create a semantic segmentation mask. The decoder block starts with a 2×2 transposed convolution. It is then combined with the corresponding feature map from the encoder block. These links provide features from previous layers that are sometimes lost due to the depth of the network. Two 3×3 convolutions are then used, where each convolution is followed by a ReLU activation function. At the output of the final decoder, sigmoid is used for binary segmentation scenarios, while softmax activation function is used for multi-class segmentation scenarios.

3.4 Model Training

Model training was carried out with the help of the workstation computer, the details of which are given in Tab. 3.

Table 3 Computer specifications

CPU	GPU	RAM	Operating System
Intel Xeon E5-1603	Quadro P1000	16 GB	Windows 10

Python was chosen as the programming language. The Keras library was used for the implementations of the models. Epoch number is 500, batch size is 16, learning rate is 0.001, Adam is used as optimizer function.

3.5 Performance Metrics

In this study, pixel accuracy and jaccard index evaluation criteria were used to measure the prediction success of the segmentation model (code available at <https://github.com/OguzhanKATAR23>). Predictions in machine learning studies are examined under four categories: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). These categories are also valid for segmentation studies, but they are more difficult to distinguish compared to standard classification studies.

3.5.1 Pixel Accuracy

Pixel Accuracy (PA) is the ratio of the number of pixels known to be correctly classified because of the pixel-based comparison of the ground truth mask and the predicted mask to the total number of pixels. It is calculated according to the mathematical equation specified in Eq. (1).

$$\text{Pixel Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Pixel Accuracy seems useful due to its easy computation and complexity of performance metrics, but it can produce deceptive results. High resolution images may cause deceptive results as stated. For example, an image with 1920×1080 size has a total of 2073600 pixels. Assuming 90% of pixels with background value in ground truth masks, the remaining 10% for other labels. Such cases mean that the class pixel distribution is extremely uneven. Even if the model fails to predict the mask and only returns an image with pixel values of zeros, the pixel accuracy value will be 90% due to the large true negatives value. However, the model could not produce results for the areas to be segmented in this scenario. Therefore, pixel accuracy evaluation should be used for images with balanced distribution instead of images with uneven class distribution. The extraction of false values, which are important in the calculation of the pixel accuracy metric is shown in Fig. 11.

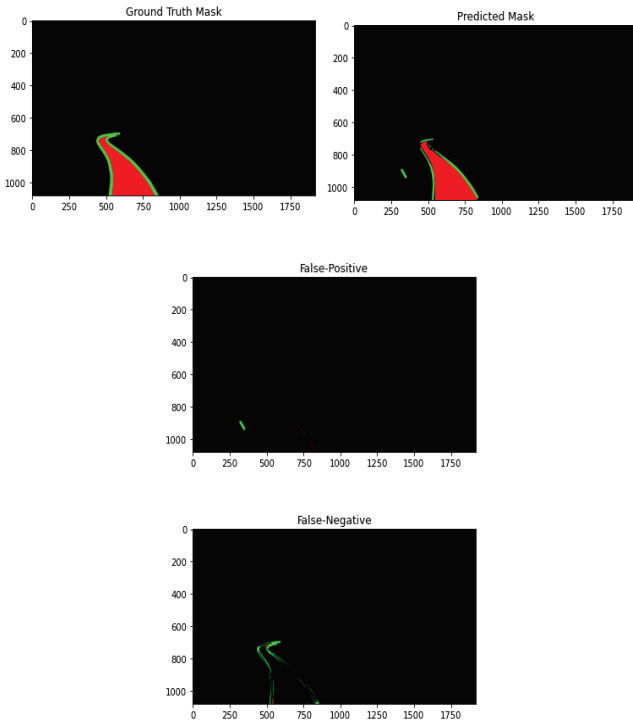


Figure 11 False values in pixel accuracy calculation

Various additional evaluation criteria can be calculated using the four categories required to reveal the pixel accuracy value. These evaluation criteria and mathematical Eq. (2), Eq. (3) and Eq. (4) are as follows.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

3.5.2 Jaccard Index

The jaccard index, which indicates the similarity between the two given images, is also called Intersection over Union (*IoU*). *IoU* value is calculated by the mathematical equations given in Eq. (5) at what rate the estimated mask pixels overlap with the original mask.

$$\text{IoU}(A, B) = \frac{A \cap B}{A \cup B} \quad (5)$$

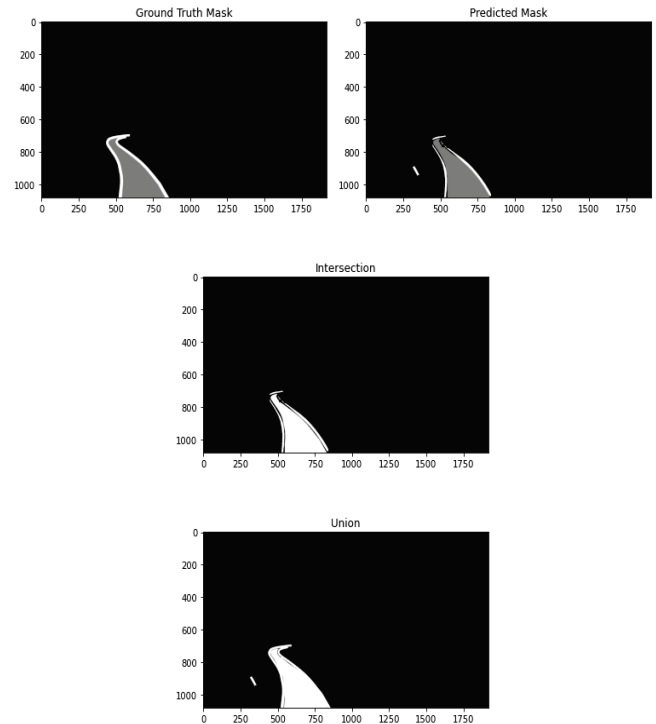


Figure 12 Intersection and Union values in IoU calculation

It gives more reliable results than pixel accuracy. Therefore, it is one of the most important performance metrics considered in segmentation studies. Examples of intersection and union required for calculating the IoU value are given in Fig. 12.

In order to compare the pixel accuracy and jaccard index methods more clearly, the same ground truth mask and predicted mask were used in both methods. In Tab. 4, the numerical values of the related methods are given in detail.

Table 4 Pixel accuracy vs jaccard index

Metric	Score	Equation with pixel counts
PA	99.6%	$(67227+1998187) / (67227+1998187+1361+6825)$
IoU	87.4%	$65576 / 75002$

4 EXPERIMENTAL RESULTS

In this study, we developed a fully automatic method to provide vision ability to railway vehicles. Our U-Net models were trained for 500 epochs with the help of the four sub-datasets. The mean Intersection over Union (mIoU) values obtained as a results of the relevant trainings are given in Tab. 5.

Table 5 Results of the trainings

Dataset Name	Model Name	Number of Class	mIoU (%)
SubDS1	U-Net	2	89.1
SubDS2	U-Net	3	83.2
SubDS3	U-Net	4	79.7
SubDS4	U-Net	19	29.6

The loss graph of the model, which was designed to detect the rail trail of autonomous railway vehicles with the multi-class segmentation method and trained with the SubDS2 dataset, is shown in Fig. 13.

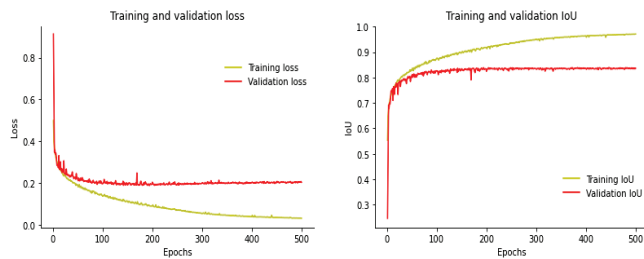


Figure 13 Loss and IoU graphics

In the SubDS1 dataset, the mask was estimated with an *IoU* value of 95.2% for the randomly selected image among the samples reserved for testing. Intersection pixel count is 153935, union pixel count is 161549. The prediction mask is given in Fig. 14.

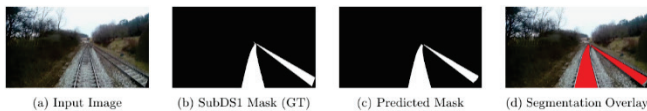


Figure 14 Prediction of the model trained with SubDS1

In the SubDS2 dataset, the mask was estimated with an *IoU* value of 91.9% for the randomly selected image among the samples reserved for testing. Intersection pixel count is 141705, union pixel count is 154038. The prediction mask is given in Fig. 15.

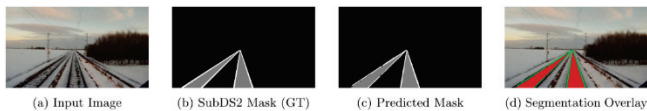


Figure 15 Prediction of the model trained with SubDS2

In the SubDS3 dataset, the mask was estimated with an *IoU* value of 83.6% for the randomly selected image among the samples reserved for testing. Intersection pixel count is 222384, union pixel count is 265786. The prediction mask is given in Fig. 16.

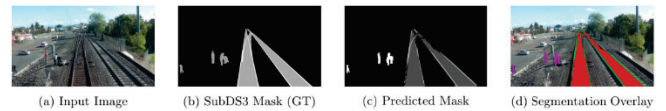


Figure 16 Prediction of the model trained with SubDS3

In the SubDS4 dataset, the mask was estimated with an *IoU* value of 45.4% for the randomly selected image among the samples reserved for testing. Intersection pixel count is 941415, union pixel count is 2073600. The prediction mask is given in Fig. 17.

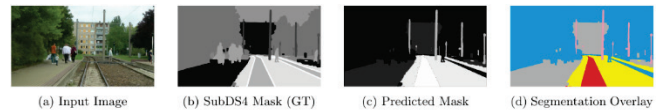


Figure 17 Prediction of the model trained with SubDS4

5 CONCLUSION

The increase in the usage areas of autonomous systems has also led to the realization of various studies in the transportation sector. In this study, an artificial intelligence-based system was proposed to improve the vision features of autonomous driving vehicles in railway transportation. A public railway dataset was customized for this study. It was aimed to reach high *IoU* values by training the U-Net models with the specified dataset. In order to increase the prediction accuracy values of our models, which are capable of multi-class segmentation and binary segmentation, the number of samples used in training and the scope of the samples should be increased. For autonomous railway vehicles that will perform vital tasks such as passenger transport, the margin of error should be close zero. In addition to the operations that need to be done on the datasets to reduce this margin of error it is necessary to develop deep learning-based models used. The proposed system can be easily integrated into the technology used in the field of railway transportation today. Designing such artificial intelligence-based studies and using them in the field will create the future of the transportation sector. Revealing similar studies can help researchers working on rail systems and smart transportation.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

6 REFERENCES

- [1] Hager, G. D., Bryant, R., Horvitz, E., Mataric, M., & Honavar, V. (2017). Advances in artificial intelligence require progress across all of computer science. arXiv preprint arXiv:1707.04352.
- [2] Chen, H., Wen, Y., Zhu, M., Huang, Y., Xiao, C., Wei, T., & Hahn, A. (2021). From automation system to autonomous system: An architecture perspective. *Journal of Marine Science and Engineering*, 9(6), 645. <https://doi.org/10.3390/jmse9060645>

- [3] Hutchins, N. & Hook, L. (2017). Technology acceptance model for safety critical autonomous transportation systems. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, 1-5. <https://doi.org/10.1109/DASC.2017.8102010>
- [4] Banister, D. (Ed.). (1995). *Transport and urban development*. Taylor & Francis.
- [5] Lababidi, H. M., Ahmed, M. A., Alatiqi, I. M., & Al-Enzi, A. F. (2004). Optimizing the supply chain of a petrochemical company under uncertain operating and economic conditions. *Industrial & Engineering Chemistry Research*, 43(1), 63-73. <https://doi.org/10.1021/ie030555d>
- [6] Tabernik, D., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, 31(3), 759-776. <https://doi.org/10.1007/s10845-019-01476-x>
- [7] Gschwandtner, M., Pree, W., & Uhl, A. (2010). Track detection for autonomous trains. In *International Symposium on Visual Computing*, Springer, 19-28. https://doi.org/10.1007/978-3-642-17277-9_3
- [8] Singh, A. K., Swarup, A., Agarwal, A., & Singh, D. (2019). Vision based rail track extraction and monitoring through drone imagery. *Ict Express*, 5(4), 250-255. <https://doi.org/10.1016/j.icte.2017.11.010>
- [9] Aly, M. (2008). Real time detection of lane markers in urban streets. In *2008 IEEE Intelligent Vehicles Symposium*, 7-12. <https://doi.org/10.1109/IVS.2008.4621152>
- [10] Wang, Y., Wang, L., Hu, Y. H., & Qiu, J. (2019). RailNet: a segmentation network for railroad detection. *IEEE Access*, 7, 143772-143779. <https://doi.org/10.1109/ACCESS.2019.2945633>
- [11] Mammeri, A., Siddiqui, A. J., & Zhao, Y. (2021). UAV-assisted Railway Track Segmentation based on Convolutional Neural Networks. In *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 1-7. <https://doi.org/10.1109/VTC2021-Spring51267.2021.9448887>
- [12] Zendel, O., Murschitz, M., Zeilinger, M., Steininger, D., Abbasi, S., & Beleznaï, C. (2019). Railsem19: A dataset for semantic rail scene understanding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. <https://doi.org/10.1109/CVPRW.2019.00161>
- [13] Liu, Y., Ren, Q., Geng, J., Ding, M., & Li, J. (2018). Efficient patch-wise semantic segmentation for large-scale remote sensing images. *Sensors*, 18(10), 3232. <https://doi.org/10.3390/s18103232>

Authors' contacts:

Oğuzhan Katar, Research Assistant
(Corresponding author)
Firat University,
Department of Software Engineering,
23119 Elazığ/Turkey
okatar@firat.edu.tr

Erkan Duman, Assistant Prof. Dr
Firat University,
Department of Computer Engineering,
23119 Elazığ/Turkey
erkanduman@firat.edu.tr

Performance Optimisation of Single-Phase Variable Reluctance Shaded-Pole Motor

Adem Dalcali*, Mehmet Akbaba

Abstract: Shaded pole induction motors' theoretical analysis is quite difficult because of the presence of strong harmonic components and the elliptical rotating field. The change of air gap reluctance is affecting the air gap flux distribution and thus the performance of the motor. In the study, the effect of the change in the SPIM's stator, rotor and squirrel cage on the performance of the motor is analysed by the Finite Element Method. The flux distributions and torque values of the closed-interpoles type and the stator with different interpoles clearances are obtained. The effects of single and double cage structures in the rotor part and the effects of copper and aluminium cage materials in the squirrel cage were analysed. According to the analyses, it was determined that the single cage motor with aluminium cage material exhibited the best torque performance with a torque value of 132.78 mNm.

Keywords: double cage; elliptical rotating field; finite element method; shaded-pole motor; single cage

1 INTRODUCTION

When voltage is applied to the stator of three-phase induction motors, a rotating field is produced, and the motor starts to rotate. The resulting rotating field circulates current in the rotor, which creates a closed circuit. The interaction between the field formed by these currents and the rotating field of the stator ensures the rotation of the rotor. In three-phase induction machines, the placement of the windings with a 120° phase difference provides the formation of the rotating field. However, in single-phase induction machines, the rotating field cannot be created since a single phase is applied to the stator. Therefore, the machine should be manufactured with auxiliary winding or shaded pole to prevent this situation in single-phase induction motors [1-3]. Shaded pole induction motors are preferred, especially in low-power applications, due to their simple structure, low cost, and direct operation from a single-phase source [4, 5]. However, these motors cannot compete with three-phase motors, especially in industrial applications that require high power.

Although Shaded pole induction motors (SPIMs) are structurally simple, they are among the electrical machines that are theoretically the most difficult to analyse. This difficulty is due to the fact that SPIMs have an elliptical rotating magnetic field. When studies in the literature was reviewed, it was observed they focused on the finite element analysis (FEA) of the shaded pole motors. Four different motor models were designed, and 2D FEAs were performed under different load conditions. The magnetic flux density and saturation regions in the core were determined by performing the magnetostatic and transient analyses of the models [6]. In another analysis study, the skew effect of the rotor bars was realized by creating n-piece structures. The authors preferred the 2D analysis method since 3D analyses would cause longer analysis time [7]. Unlike this study, Shamlou and Mirsalim preferred 3D analyses since short-circuit resistance, winding leakage inductance, and skew cannot be directly taken into account in 2D analyses [8]. When the performance improvement studies of the SPIM were reviewed, it was observed that the effects of structural

and different materials on the shaded pole ring, stator, and rotor were investigated. In the study investigating the effect of the short-circuit ring on the motor, it was determined that the reverse rotating magnetic field component was less in the double short circuit ring motor, which resulted in the reduction of vibrations in the torque characteristic [9]. The dynamic model of the shaded pole motor can be derived using the step response of the circuit approach. With this method, the effect of core losses, which is generally neglected, can also be added to the model [10]. Another analysis method is the symmetrical components theory. In this method, the accuracy of the results depends on taking the parameters right and complete [11]. In the analysis of the shaded pole motor performed using the symmetrical components theory, high order harmonic components were neglected in the study conducted by considering only the fundamental harmonics. However, the fact that the SPIM has high harmonic components reduces the accuracy of the study in this sense [11]. The design problem of SPIMs can be considered a vector optimization problem. Aluzri used the problem in the analysis of the cost of the motor and the starting torque using a nonlinear programming technique [12].

Finite element analyses are used in the approximate analysis of the quantities which are used in the analysis of electrical machines and are continuous in a certain region such as electric and magnetic fields. In this method, these continuous quantities are divided into a finite number of small regions that can be expressed by partial differential equations. To obtain the solution at a point in the region, the contribution of the elements surrounding that point is taken into account. Thus, a set of linear equations with as many equations as the number of nodes is derived, and the desired quantities are calculated. Performance and design criteria such as core losses and flux distributions of the machine can be obtained by analyses [13-15]. In the SPIM, starting torques are small compared to other single-phase motors with auxiliary winding. SPIMs are low-efficiency motors, although they are widely used [16]. Starting torques can be developed using structural designs and developing material technology.

In this study, analysis of a variable reluctance shaded pole induction motor is carried out using the finite element method. The effects of closed-interpoles (i.e. without clearance between the stator interpoles) type structure and different aperture stator on motor performance are investigated. In addition, the effect of squirrel cage material on the performance is also investigated. The paper has been structured as follows: In section II the operating principle of the shaded pole motors and the characteristics of the motor under investigation subject to study is given. In section III, the torque and magnetic flux distributions of different structures of the motor are investigated. Section 4 draws conclusions.

2 SHADED POLE MOTORS IN GENERAL AND THE PROPOSED MOTOR

While the rotor of SPIMs is squirrel-cage, the stator consists of salient poles. A field is formed as a result of a single-phase alternating voltage applied to the main winding in the salient poles. This field also induces a field in the rotor. However, torque cannot be generated in the motor. There is a need for shaded poles so that the motor can generate torque. Copper rings serve to provide starting. The voltage induced in the shaded pole due to the main winding causes short-circuit currents to flow in the short-circuited slit pole rings. These currents create a flux in the almost backward phase compared to the stator flux. These two alternative fields allow the motor to run by creating an elliptical rotating field in the air gap. The rotor moves from the non-shaded pole part of the stator pole to the shaded pole part of it with the effect of the elliptical rotating field given in Eq. (1). The field created by the pole winding slips toward the commutating field. When the current changes direction, this time the polarity of the poles changes, and again the slip direction of the field remains the same [1].

$$\Phi_s = \Phi_a + \Phi_{sp} \quad (1)$$

The relative states of the stator winding area and slit pole ring areas, and flux phasor diagrams are presented in Fig. 1.

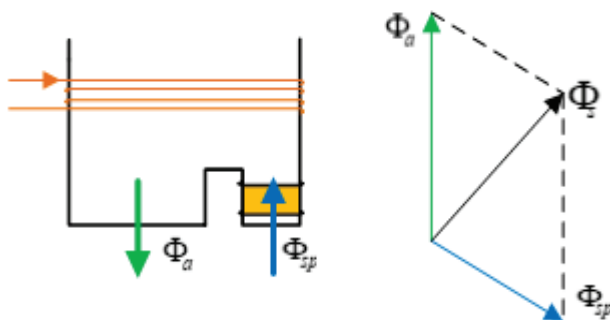


Figure 1 Flux paths and flux diagram of the SPIM

Shaded pole induction motors with variable air gap are electrical machines that are easy to manufacture but difficult to analyse mathematically. The first reason for the simplicity

in their structure is that the auxiliary winding consists of copper rings placed in the slits opened on the stator poles. The second reason is that the stator windings are collectively placed on salient poles. The reasons for the difficulty of the mathematical analysis are the presence of harmonic components in the stator magnetomotive force and the ellipticity of the rotating field [17]. The comparison of single-phase induction motors is summarily presented in Tab. 1 [18, 19]

Table 1 Characteristics of single-phase induction motors

Motor Type	Power (HP)	Efficiency (%)	Application Area
Auxiliary winding	1/20 - 1	55-65	Aspirator, dishwasher etc. That require low starting torque applications
Capacitor-start	0.125 - 1	55-65	Compressor, refrigerator, washing machine, pump
Permanent split capacitor	0.125 - 1	60-70	Low noise applications such as aspirator, pump
Shaded pole	1/200 - 1/20	25-40	Applications requiring low starting torque such as aspirator, hair dryer

The initial model of the motor discussed in the study is presented in Fig. 2 [20].

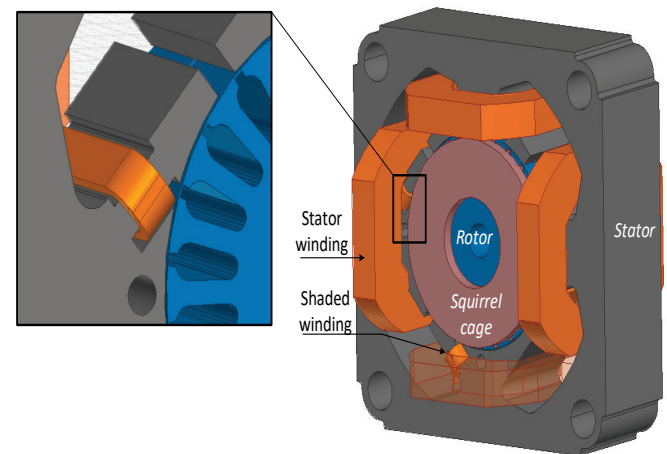


Figure 2 SPIM model

The dimensions and characteristics of the motor used in the study are presented in Tab. 2.

Table 2 Motor characteristics

Property	Value
Voltage (V)	220
Number of poles	4
Number of slot	26
Core size (mm)	82×82×25
Number of winding turns	580
Stator core material	M43
Rotor core material	Stell-1010
Rotor type	Squirrel cage
Power (W)	15

3 PERFORMANCE EVALUATION

The design of electrical machines requires the solution of interrelated problems of electromagnetic, mechanical,

thermal, and material technology. The designer determines input variables and constraints by considering the cost and desired performance criteria. The design of the machine that exhibits the expected characteristics requires long-term analysis and investigation. The performance of the machine is affected by many physical parameters such as the number of stator slots, the number of rotor slots, and the slot structure of the machine [21]. In this section, three different structures of the stator are first examined. In the first case, the closed-interpoles stator structure (Fig. 3a) is investigated. In the second case, the stator interpoles clearance is selected as 4 mm (Fig. 3b), and in the last case, the stator interpoles clearance is selected as 8 mm (Fig. 3c).

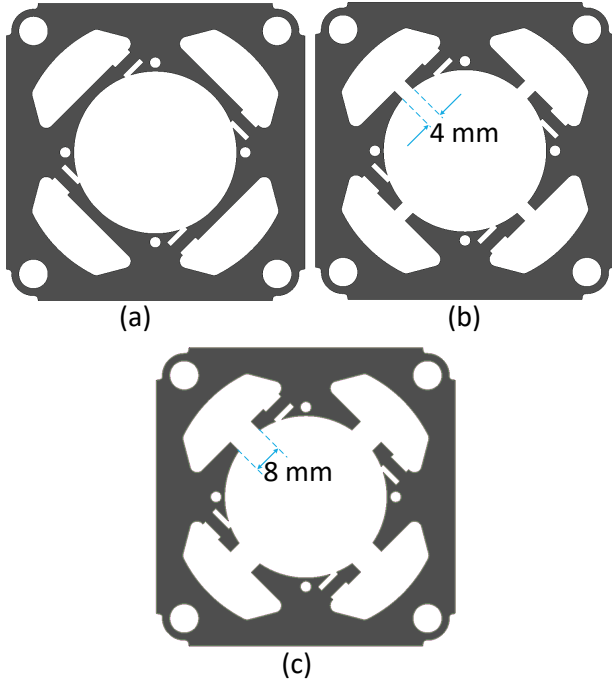


Figure 3 Stator geometry

Except for the specified clearance, the stator parameters were left constant. Thus, only the effect of this parameter is examined. It is important to know the distribution of the magnetic field in the motor in the design and determination of the operating characteristics of the SPIM. The net flux of a closed surface in a magnetic field can be written as follows [22];

$$\int_s \vec{B} \, ds = 0 \quad (2)$$

The linear integral of the magnetic field along a closed circumference in a magnetic field is equal to the sum of the currents enclosed by the circumference and is expressed by Eq. (3).

$$\int \vec{H} \, dl = \sum I \quad (3)$$

By expressing the surface of the closed circumference as s , assuming that the current is uniformly distributed over the surface with a constant current density (J);

$$\int_c \vec{H} \, dl = \int_s \vec{J} \, ds \quad (4)$$

can be written. According to the Stokes theorem.

$$\int_s \text{rot} \vec{H} \, ds = \int_s \vec{J} \, ds \quad (5)$$

and

$$\text{rot} \vec{H} = \vec{J} \quad (6)$$

are obtained. If Eq. (6) is rewritten with the $\vec{B} = \mu \vec{H}$ relation between the magnetic field strength and magnetic flux density.

$$\text{rot} \left(\frac{1}{\mu} \vec{B} \right) = \vec{J} \quad (7)$$

The magnetic field can be derived from a vector function, which we will denote by $\vec{A}$.

$$\vec{B} = \text{rot} \vec{A} \quad (8)$$

If Eq. (7) is arranged with the help of Eq. (8) by using $1/\mu = v$.

$$\text{rot} (v \text{rot} \vec{A}) = \vec{J} \quad (9)$$

is obtained.

$$\text{rot}(\vec{A}) = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & A \end{bmatrix} = \frac{\partial A}{\partial y} \vec{i} - \frac{\partial A}{\partial x} \vec{j} \quad (10)$$

If Eq. (10) is combined with the relation $\text{rot}(v \text{rot} \vec{A}) = \vec{J}$.

$$\frac{\partial}{\partial x} \left(v \frac{\partial A}{\partial x} \right) + \frac{\partial}{\partial y} \left(v \frac{\partial A}{\partial y} \right) = -\vec{J} \quad (11)$$

The solution of the magnetic field distribution in the motor is obtained from the solution of Eq. (11) by FEA. The flux density distribution in the case of the main winding and squirrel cage excitation is obtained in 3 different structures and are presented in Fig. 4.

When the flux distributions are examined, especially in case of closed-interpoles stator structure, a large amount of flux completes its circuit over the stator instead of passing through the high reluctance air gap to the rotor. In this case, it creates a high flux density in the relevant parts. When

rotors are analysed, less flux density is observed in the closed-interpoles structure compared to other structures due to less flux passing to the rotor. It is observed that useful common flux is higher in open-interpoles structures, which is the first indication that it will result in outstanding performance.

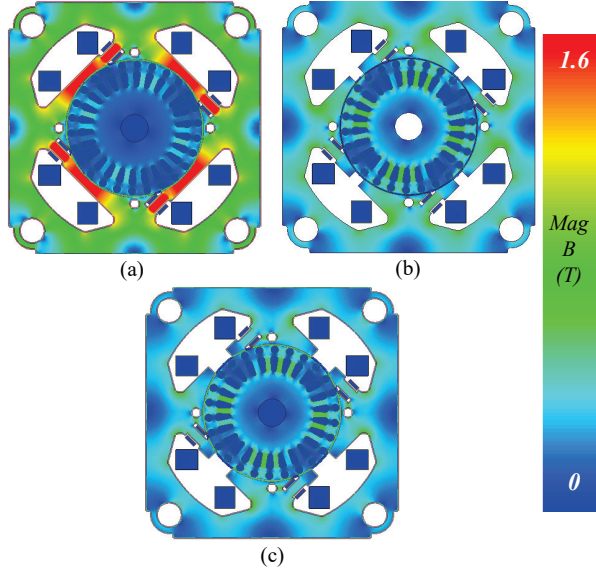


Figure 4 Magnetic flux density distribution; a) closed-interpoles, b) clearance 4 mm, c) clearance 8 mm (both main winding and squirrel cage excited)

The flux density distribution performed only in the main winding excited state was obtained in 3 different structures and is presented in Fig. 5.

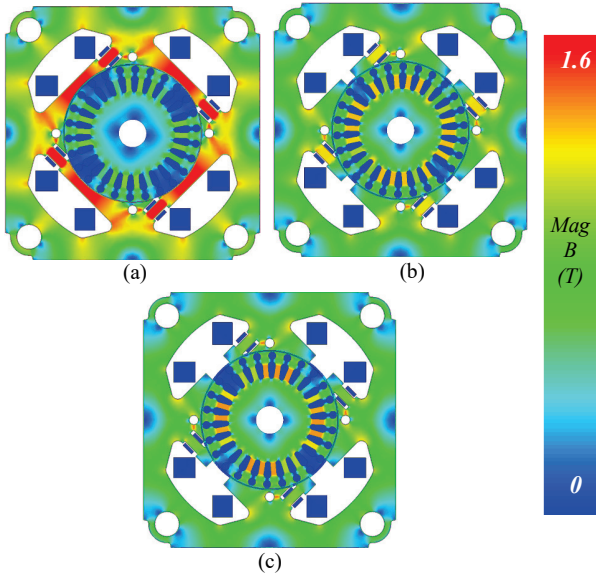


Figure 5 Magnetic flux density distribution; a) closed-interpoles, b) clearance 4 mm, c) clearance 8 mm (only with main winding excited)

3.1 Comparison of Single and Double Cage Rotor Designs

In this section, a squirrel cage rotor is designed as a single and double cage, and their performances are examined.

Copper was used as the squirrel cage material. The structures used in the study and their dimensions are presented in Fig. 6.

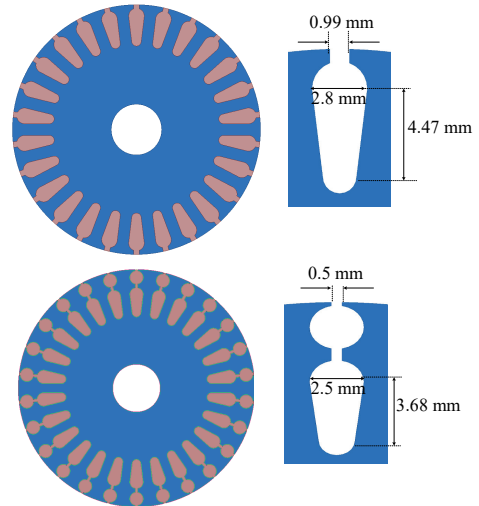


Figure 6 Rotor structure a) single cage, b) double cage

The torque values obtained in analyses performed when the slip value was 0.1 are presented in Tab. 3. Copper was preferred in the squirrel cage material in the analyses.

Table 3 Induced torque values

Slip	Rotor type	0 (closed type)	4 mm	8 mm
$s = 0.1$	Single	84.87 mNm	111.51 mNm	97.46 mNm
	Double	55.16 mNm	85.74 mNm	81.02 mNm

The torque values of the double cage structure are found to be relatively higher. Furthermore, less useful flux transferred to the rotor in the closed-type stator caused the induced torque to be low in this structure. It is observed that torque value is maximum in the structure with a 4 mm interpoles-clearances. Flux distributions in single and double cage rotors are presented in Fig. 7. Appropriate flux distributions are obtained for the slots in the designs. However, the designer should consider the flux density in the selection of the number of slots so that the teeth are not saturated, especially in high numbers of slots.

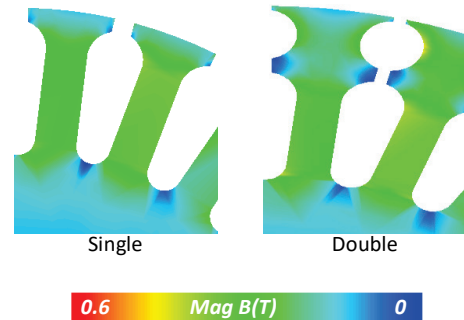


Figure 7 Flux density distribution

3.2 Comparison of Copper and Aluminium Rotor Conductors

The material of the cage is another parameter that can be examined in squirrel cage motors. In this study, as squirrel cage material copper and aluminium are used and resulting

performance are analysed separately. The conductivity of copper is better by 60% compared to aluminium material. However, the high melting point and the high cost of the copper at the production stage constitute a disadvantage [23]. The torque values obtained in analyses when the slip value is 0.1 are presented in Tab. 4. In the analyses, the rotor structure was chosen as a double cage.

Table 4 Induced torque values

Slip	Cage type	0 (closed-interpoles type)	4 mm	8 mm
$s = 0.1$	Copper	55.16 mNm	85.74 mNm	81.02 mNm
	Aluminium	66.18 mNm	99.17 mNm	96.05 mNm

The aluminium cage motor provided higher torque in all three structures compared to the copper cage motor. However, the fact that the resistance of aluminium is higher compared to copper causes high rotor Joule loss. The higher Joule loss will also result in lower efficiency in an aluminium motor compared to a copper motor. Finally, in the analysis of the single cage and aluminium structure, motor, the highest torque rating of 132.78 mNm has been obtained.

4 CONCLUSION

SPIMs are frequently preferred in the industry due to their low cost and simple structure, especially in low-power applications. In this respect, it is important to improve the performance of these low-efficiency machines. In this study, magnetostatic and transient analyses of the variable reluctance SPIM were performed, and magnetic flux distributions and induced torque values are analysed. The stator body is designed in three different structures, and the flux distributions in the core are derived. The flux distributions are derived for the cases when only the main winding is excited and when both the main winding and squirrel cage are excited. Especially in the closed-interpoles structure, the a large portion of the useful flux completes its circuit over the stator instead of passing to the rotor, which resulted in low induced torque and high flux density, which means high saturation and high iron losses, in the closed-interpole areas. In this type of structures, designers should pay attention to the saturation, especially in the bridge parts. In the analyses performed on the rotor part, single and double cage structures are examined. In the single cage structure, the torque value of 111.51 mNm was obtained at the highest value in the structure with a clearance of 4 mm. After this analysis, the material of the cage was changed to aluminium and the induced torques are obtained. Higher torque is obtained in the aluminium cage motor compared to the copper motor. Investigations showed that a highest torque of 132.78 mNm is obtained in the case of single cage rotor with aluminium bars and end rings.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which

took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

5 REFERENCES

- [1] Kurita, N., Ishikawa, T., & Suzuki, G. (2016). Development of the two pole type shaded pole self-bearing motor. *Materials Science Forum*, 856, 196-201. <https://doi.org/10.4028/www.scientific.net/MSF.856.196>
- [2] Enesi, Y. A., Tsado, J., Nwohu, M., Usman, U. A., & Imoru, O. A. (2016). Split-phase motor running as capacitor starts motor and as capacitor run motor. *Leonardo Electronic Journal of Practices and Technologies*, 15(28), 253-272.
- [3] Crăciunescu, A. (2019). The Development of a New Type of Repulsion-Induction Motor. *2019 IEEE International Electric Machines & Drives Conference (IEMDC)*, 1447-1451. <https://doi.org/10.1109/IEMDC.2019.8785265>
- [4] Naku, L. & Spahiu, A. (2018). Saving energy by replacing IM with BLDC MOTOR in fan application. *European Journal of Electrical Engineering and Computer Science*, 2(5). <https://doi.org/10.24018/ejece.2018.2.5.27>
- [5] Alidousti, A., Sadoughi, A., Behbahani, H., & Raiesi, Y. (2018). A new rotor prototype for single phase line start permanent magnet synchronous motor based on amendments to a small industrial shaded pole induction motor. *The 9th Annual Power Electronics, Drives Systems and Technologies Conference (PEDSTC)*, 218-223. <https://doi.org/10.1109/PEDSTC.2018.8343799>
- [6] Sarac, V. & Cundev, D. (2012). Electromagnetic fields calculation at single phase shaded pole motor. *Electrotechnica & Electronica*, 47(7-8/20), 41-45.
- [7] Petkovska, L., Cundev, M., & Sarac, V. (2002). FEM analysis of asymmetrical magnetic field in electrical machines. *The 2nd International Conference on Advanced Computational Methods in Engineering*, 1-10.
- [8] Shamlou, S. & Mirsalim, M. (2015, November). A new restructured shaded pole induction motor-3D finite element analysis and experimental verification. *The 30th International Power System Conference (PSC)*, 136-141. <https://doi.org/10.1109/IPSC.2015.7827739>
- [9] Pessina, G. & Morra, E. (2010, June). Optimization and design of the shaded pole single phase asynchronous motor. *SPEEDAM 2010*, 978-981. <https://doi.org/10.1109/SPEEDAM.2010.5544905>
- [10] Ojaghi, M. & Daliri, S. (2015, October). A detailed dynamic model for single-phase shaded pole induction motors. *The 18th International Conference on Electrical Machines and Systems (ICEMS)*, 1987-1992. <https://doi.org/10.1109/ICEMS.2015.7385366>
- [11] Sarac, V., Petkovska, L., & Cundev, M. (2001). An improved performance analysis of a shaded pole motor. *Proceedings of the Power Electronics Intelligent Motion and Power Quality Conference*, 399-405.
- [12] Aluzri, A. M. & Mohammed, I. A. (2002). Vector optimization design of a shaded-pole induction motor. *IEEE Transactions on Energy Conversion*, 17(3), 374-379. <https://doi.org/10.1109/TEC.2002.801990>
- [13] Dalcı, A. & Akbaba, M. (2017). Comparison of the performance of bridge and bridgeless shaded pole induction motors using FEM. *International Journal of Applied Electromagnetics and Mechanics*, 54(3), 341-350. <https://doi.org/10.3233/JAE-160133>
- [14] Dalcı, A. & Akbaba, M. (2016). Comparison of 2D and 3D magnetic field analysis of single-phase shaded pole induction

- motors. *Engineering Science and Technology, an International Journal*, 19(1), 1-7. <https://doi.org/10.1016/j.jestch.2015.04.013>
- [15] Çetin, O., Dalcalı, A., & Temurtaş, F. (2020). Estimation of Permanent Magnet Synchronous Generator Performance with Artificial Neural Network Models. *Sakarya University Journal of Computer and Information Sciences*, 3(1), 60-73. <https://doi.org/10.35377/saucis.03.01.724976>
- [16] D'Aguanno, D., Marignetti, F., Faginoli, F., & El-Shahat, A. (2018). Single-phase motors for household applications. *Electric Machines for Smart Grids Applications Design, Simulation and Control, IntechOpen*, 416-908. <https://doi.org/10.5772/intechopen.79203>
- [17] Akbaba, M. & Fakhro, S. Q. (1992). Field distribution and iron loss computation in reluctance augmented shaded-pole motors using finite element method. *IEEE Transactions on Energy Conversion*, 7(2), 302-307. <https://doi.org/10.1109/60.136225>
- [18] Bal, G. (2011). *Special Electrical Machines*. 3th Edition, Seçkin Publications. (In Turkish).
- [19] Hajek, V., Mach, M., & Kuchar, L. (2015, September). Special small electric motors. *International Conference on Electrical Drives and Power Electronics (EDPE)*, 296-300. <https://doi.org/10.1109/EDPE.2015.7325309>
- [20] Dalcalı, A. & Akbaba, M. (2017). FEM Study of the Effects of Geometric Changes on the Variable Reluctance Shaded-Pole Motors Performance. *The 6th Paris International Conference on Recent Trends in Engineering and Technology*.
- [21] Murthy, K. V. (2008). Computer-aided design of electrical machines (pp. 223-276). Hyderabad, India: BS Publications.
- [22] Akbaba, M. (1984). A Study on Shaded-Pole Induction Motors with Variable Air Gap. Associate Professorship Thesis, Karadeniz Technical University, Trabzon, Turkey.
- [23] Brush, E. F., Cowie, J. G., Peters, D. T., & Son, D. J. V. (2003). Die-cast copper motor rotors: motor test results, copper compared to aluminium. *Energy efficiency in motor driven SYSTEMS*, Springer, Berlin, Heidelberg, 136-143. https://doi.org/10.1007/978-3-642-55475-9_21

Authors' contacts:

Adem Dalcalı, Assoc. Prof. Dr.

(Corresponding author)

(1) Department of Electrical and Electronics Engineering,
Bandırma Onyedli Eylül University,
Bandırma, Balıkesir, Turkey

(2) National Research University, TIAME,
100000, Tashkent, Uzbekistan

+902667170117, adalcali@bandirma.edu.tr

Mehmet Akbaba, Prof. Dr.

Department of Computer Engineering,

College of Engineering,

Karabük University,

Karabük, Turkey

+905332645060, mehmetakbaba2021@gmail.com

EEG Based Emotion Prediction with Neural Network Models

F. Kebire Bardak*, M. Nuri Seyman, Feyzullah Temurtaş

Abstract: The term "emotion" refers to an individual's response to an event, person, or condition. In recent years, there has been an increase in the number of papers that have studied emotion estimation. In this study, a dataset based on three different emotions, utilized to classify feelings using EEG brainwaves, has been analysed. In the dataset, six film clips have been used to elicit positive and negative emotions from a male and a female. However, there has not been a trigger to elicit a neutral mood. Various classification approaches have been used to classify the dataset, including MLP, SVM, PNN, KNN, and decision tree methods. The Bagged Tree technique which is utilized for the first time has been achieved a 98.60 percent success rate in this study, according to the researchers. In addition, the dataset has been classified using the PNN approach, and achieved a success rate of 94.32 percent.

Keywords: AdaBoost; bagged tree; EEG signals; emotion prediction; multi-layer perceptron; probabilistic neural networks

1 INTRODUCTION

Psychological response to an event, individual, or thing is called emotion. Emotions that play a significant part in a human's daily life, appearing in both psychological and physiological situations, have been studied in various fields, which include virtual reality, e-therapy, health, and brain-machine interfaces. There has, however, been a significant rise in studies predicting emotions recently. Various methods, susceptible to manipulation and interpretation, such as facial expressions, mimicry, and voice signals, have also been used to predict emotions in the literature [1, 2]. Hence, this has led to a decrease in the prediction accuracy rate since the data has been erroneously collected. Nonetheless, several physiological signals that are difficult to control, such as an electrocardiogram (ECG), an electroencephalogram (EEG), an electromyogram (EMG), blood pressure (BP), and heart rate (HR), have acquired relevance for the accuracy of emotion classification and reliability. Across these technologies, EEG data, which is low-cost, wireless, and mobile device-based, has become increasingly more common to predict emotions [3].

The emotion model has been classified into discrete and dimensional models. While the discrete model represents six essential emotions (nervousness, fear, sorrow, anger, joy, and surprise), positive and negative, the dimensional model describes their locations on the arousal-valence plane [4]. Fig. 1 depicts an emotional pattern separated into four distinct zones on the arousal-valence plane.

The main purpose of this study is to determine the best classification method to predict positive, negative, and neutral emotional responses based on EEG signals. The success rate has also been studied in k-nearest neural network (KNN), support vector machine (SVM), probabilistic neural networks (PNN), and in both PCA and non-PCA scenarios using the multi-layer perceptron (MLP). Moreover, various parameters' effects on classification accuracy rates have been examined. Consequently, it has been shown that the bagged tree ensemble method utilized has enhanced accuracy by 98.6 percent.

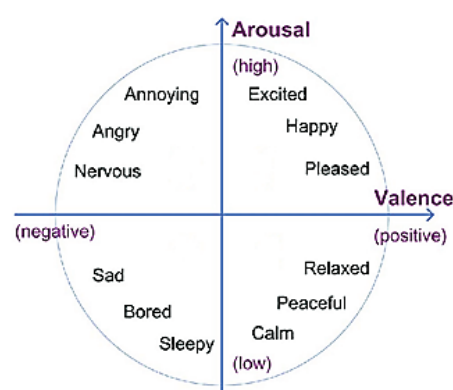


Figure 1 Arousal-valence emotion plane [5]

2 RELATED WORKS

Numerous performance analyses have been published in the literature to predict emotional states based on EEG signals. In [1], the best conclusion has been reached with Random Forest classifiers at 87.16 percent, using OneR to extract the features of the raw dataset. The features of EEG signals have been extracted using OneR, BayesNet, and InfoGain techniques and categorized using classification techniques such as MLP, random forest (RF), and SMO are given in [6]. However, with InfoGain and Random Forest, the most accurate rate has been achieved at 97.89 percent. In [7], SVM, KNN, Lightgbm, Xgboost, RF, Logistic Regression, and two neural network classification models have been used to predict emotion. The classification results have been obtained by removing the most important 150 features by using the PCA method after normalization varied between 90.61 percent and 97.0 percent. Additionally, in [8] which is aimed at classifying positive and negative emotions using visual stimuli, support vector machines have been used to classify the features obtained by particle swarm optimization and genetic screening, and success rates of between 48.85 percent and 57.42 percent have been obtained. In [9], support vector machines have been used to classify EEG collected data with audio-visual stimuli as positive or negative emotions, and the achieved success rate has been

74.55 percent on average. Emotion classification utilizing common spatial patterns has been shown to be exceptionally successful in [10], with 93.5 percent being the overall best outcome. In [11], which is proposed a convolutional neural network (CNN) based model that extracts the EEG signal properties with the discrete wavelet transform using Welch's power spectral density estimation, the best estimation result has been 90.16 percent.

The rest of this paper is organized as follows. Details regarding the dataset, the classification algorithms, and the outcomes of the used methods are thoroughly contained in Section 3 and Section 4, respectively. Finally, the analysis of the achieved results throughout the research is given in Section 5.

3 METHODS

3.1 Dataset

This study analyzed the emotion dataset in [1] and [6] based on EEG signals. The MUSE EEG headband has been used to collect data from four electrodes, TP9, AF7, AF8, and TP10 points, which are shown in Fig 2. Six movie clips have been utilized as stimuli to evoke positive and negative emotions in this dataset, which contains 2132 samples, each with 2549 features, collected from a female and a male, whereas no stimulus has been used to elicit a neutral emotion. However, EEG signals have been collected for 36 minutes in total, three minutes for each emotion. The mean, maximum-

minimum, and Fast Fourier transforms have been used in EEG signal processing at a sampling frequency of 150 Hz. Emotion-related changes in brain waveforms are seen in Fig. 3. The classification of the three main groups identified is as follows: in Class 1, clips from the films Marley and Me, Up, and My Girl elicited negative emotions. In Class 2, positive emotions have been evoked by film clips from La La Land, Slow Life, and Funny Dogs. Besides, no neutral emotions have been elicited in Class 3.

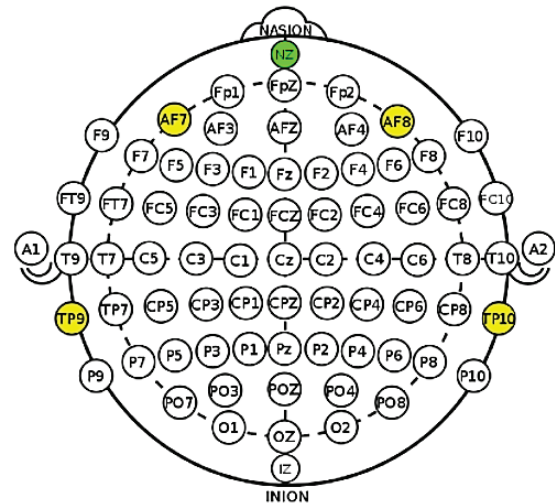


Figure 2 TP9, AF7, AF8 and TP10 EEG sensors of Muse headband in international standard EEG insertion system [5]

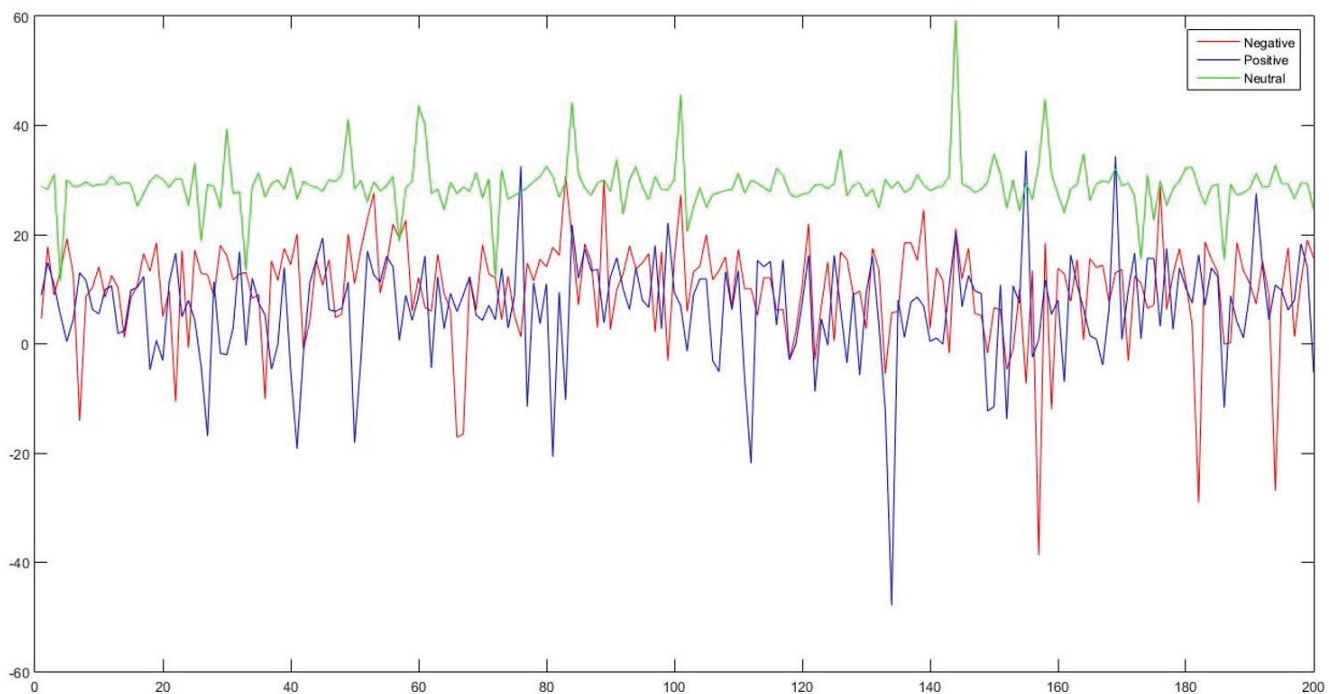


Figure 3 Emotional changes in brainwaves

3.2 Classification Methods

Classification methods, MLP, SVM, PNN, KNN, and decision trees, all of which employed five-fold cross-

validation and an 80 percent training set, are given in this section.

Multi-Layer Neural Networks

A multi-layered neural network employs the Levenberg-Marquardt algorithm, a well-known training approach for

identifying brain signals. The classification performance is determined by an input layer, two hidden layers, and an output layer. However, PCA transformation is carried out, which reduces the dataset's size and removes irrelevant information from the dataset to identify the significant data attributes [12, 13]. The significant advantages of PCA for large-scale datasets include reduced noise sensitivity, decreased memory, and capacity restrictions, and more efficient performance in small-dimensional spaces [14]. Depending on whether PCA is utilized, the system is trained by obtaining the top 100, 150, and 200 most relevant attributes, which uses a five-cross-validation technique.

Support Vector Machine

EEG signal classification techniques frequently use a two-group classifier called a support vector machine whose nonlinear basis functions are used to relocate linearly inseparable planes to a new space [15, 16]. Furthermore, SVM takes advantage of high-dimensional features in the kernel functions that are utilized to increase model performance [17]. In this paper, which employs the quadratic kernel function, 80 percent of the data is utilized for training and 20 percent for testing.

K-Nearest Neighbour

The k-nearest neighbour method is a classification based on the distances and similarities between the data points [12]. When data with an unknown decision class is received, the value of k closest data is used to identify which decision class the data belongs to. The data is then labeled with the class that is similar to the most relevant k value [17]. This study examines the performance of the KNN classifier using Euclidean distance over a variety of k values, with the dataset divided into 80 percent for training and testing.

Probabilistic Neural Network

A pattern recognition system known as a probabilistic neural network considers all possible outcomes before identifying a winner. The network structure includes the use of Parzen Estimators, which by Bayes' theorem increases probability density functions [18]. The link weights forward the features' input vectors to the hidden layer, which has nodes that are determined by the Euclidean distances between the weights and the vectors [16]. This paper is evaluated the classification performance using a five-fold cross-validation approach with a spread parameter of 0.0993, which has been identified as the most critical parameter in reducing errors in PNN.

Table 1 Various classifiers' results

Classifier Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
MLP	96.53	97.88	97.19
KNN	94.70	99.29	93.88
PNN	94.32	98.30	95.30
SVM	96.90	98.66	98.92
AdaBoost	96.0	99.32	96.04
Bagged Trees	98.60	99.33	98.56

Decision Trees

Decision trees resembled tree topologies, are a basic algorithm made of nodes and branches; each branch represents a probability state and, in theory, divides the data into subgroups by branching. Each rule corresponds to a branch produced during the separation process [19].

Nevertheless, decision trees are ensemble algorithms that are prone to overlearn. Bagging decision trees and AdaBoost algorithms, on the other hand, increase generalization, hence mitigating the impacts of over-learning [20]. Therefore, the study has examined the bagged trees and AdaBoost algorithms.

a Bagged Trees

In bagged trees, each new training set is formed by picking from the existing training set. In this case, certain data from the training set is excluded entirely from the bag, whereas others are included more than once. Furthermore, the algorithm randomly picks a subset of estimators to utilize in each decision segment. Thus, the training sets are used to create regression trees, which are not pruned [20]. Hence, by reducing variation and combining the outcomes of several decision trees that enhance generalization, bagging decision trees reduces the consequences of over-learning. In this paper is evaluated the classification performance of 80% for training and 20% for testing.

b AdaBoost

The AdaBoost technique, which constructs a powerful classifier from numerous weak classifiers, operates on the principle of re-concerning the classifier at each step and increasing the weight of the inaccurate predictions made in the previous step. This method is intended to identify and correct inaccurate predictions, thereby increasing the classification accuracy of the trained models [21]. Although AdaBoost is more resistant to overlearning than other classification algorithms, is commonly influenced by noisy data and outliers [20].

Optimizing the algorithm's performance by modifying the learning rate, which is highly proportionate to the system accuracy, is possible. A learning rate greater than one causes oscillation and makes it impossible to find the global optimum. In contrast, choosing an extremely slow learning rate causes an increase in training time due to the highly increased number of steps. In the study, which used 80 percent of the dataset for training, the effect of changing the learning rate from zero to one has been examined.

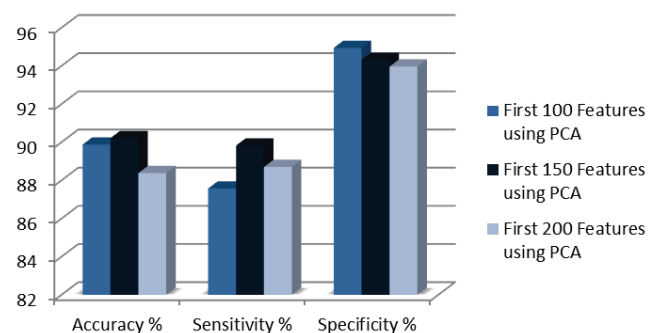


Figure 4 Accuracy, sensitivity, and specificity of MLP training using PCA

4 RESULTS

The section discussed the various methods used to classify emotion prediction datasets based on EEG signals and the implications of modifying certain classifier parameters on the network. Depending on whether or not the

PCA technique is being used, the training is performed on different features of the data. Besides, the accuracy, sensitivity, and specificity results of the classifiers are also examined.

Fig. 4 illustrates the accuracy, sensitivity, and specificity values achieved by combining PCA that extracts the most significant 100, 150, and 200 features with the MLP classification algorithm. The 150 most essential characteristics extracted with PCA have an accuracy rate of 90.21 percent, whereas the rates for 100 and 200 features are 88.86 percent and 88.38 percent, respectively. Data loss occurs when fewer features are extracted by PCA, whereas over-learning can generally occur when more features are extracted. Therefore, classification performance has been evaluated utilizing the top 100, 150, and 200 features.

For each classification, the accuracy, sensitivity, and specificity percentages are shown in Tab. 1. The sensitivity percentage represents the accurate prediction rate across all classes, while the specificity percentage represents the appropriateness prediction rate for negative predictions [22].

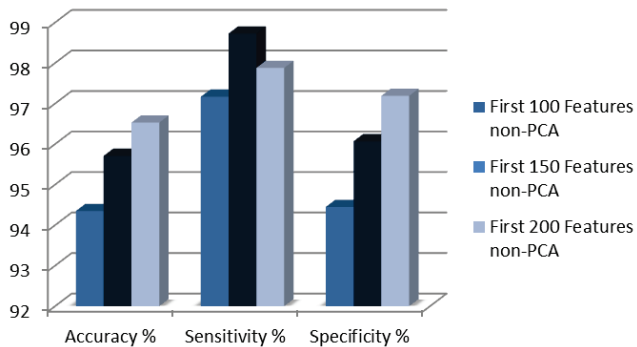


Figure 5 Non-PCA MLP training accuracy, sensitivity, and specificity

It is shown in Fig. 5 that the MLP technique, which trains the first 100 and 150 and 200 features of the dataset without the use of PCA, achieves accuracy, sensitivity, and specificity. MLP classification-based training and testing are more effective when the PCA approach is not included; the system achieves a 96.53% success rate for the first 200 features. In other words, a higher feature amount enhances the likelihood of successful accuracy.

The accuracy curves of the classifiers that have been used in the study are shown in Fig. 6. The best training achievement is 98.6 percent in the bagged trees classification. The PNN, in particular, when compared to the bagged trees algorithm, has the lowest success rate of 94.32 percent.

Table 2 Changes of the learning rate parameter in AdaBoost classification

Learning Rate	Accuracy (%)	Sensitivity (%)	Specificity (%)
0.01	90.70	97.88	97.19
0.02	91.50	97.99	92.45
0.1	92.0	98.30	95.30
0.3	92.60	99.10	92.81
0.5	93.90	98.32	93.17
0.9	96.0	99.32	96.04

Training performance based on changes in learning rate is shown in Tab. 2. The AdaBoost classifier's accuracy rate

increased once the learning rate parameter has been raised. Besides, the sensitivity and specificity are also affected according to the change in the learning rate. In addition, as seen in Fig. 7, learning rate has a significant impact on accuracy results. The algorithm's success rate has risen in direct proportion to the learning rate.

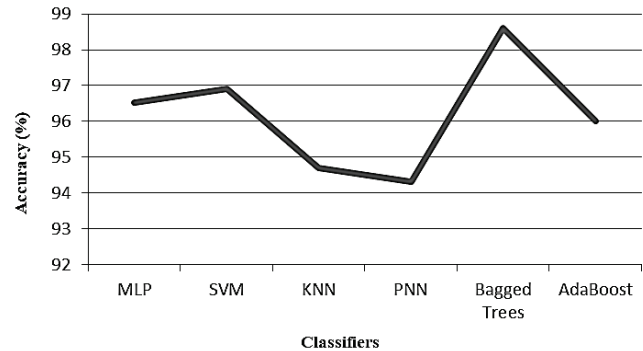


Figure 6 Several classifiers' accuracy performance

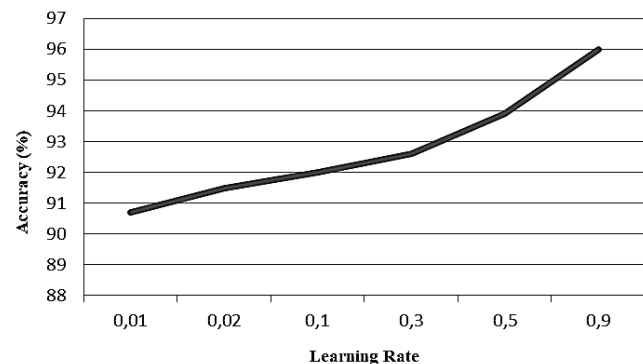


Figure 7 Accuracy curve for the AdaBoost classifier with various learning rates

Fig. 8 illustrates the accuracy curve for each k -value of the emotion prediction dataset trained with the KNN classifier. Selecting the optimal k value impacts system performance. Otherwise, selecting a value for k that is more or less than this specified value increases the mistake rate. As seen in Tab. 3, the dataset classified using KNN has shown that the optimal k value is 5 and that when k is increased to 9, the error rate rises.

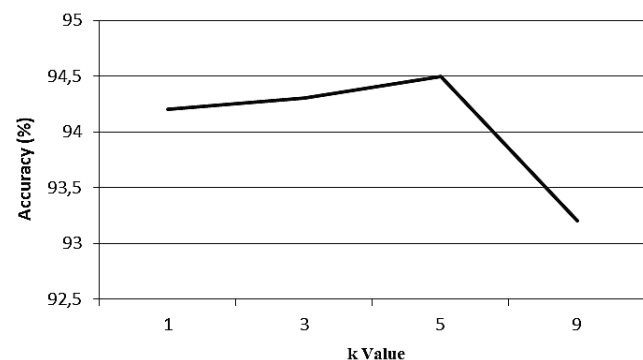


Figure 8 KNN classifier accuracy performance according to k value

Bag trees' test data is depicted in Fig. 9 as a confusion matrix. Labels 1, 2, and 3 express negative emotions, positive

emotions, and neutral emotions, respectively. The greatest misclassification for negative emotions in the bag trees model is 4, 10 for positive emotions, and 1 for neutral emotions. Consequently, compared to the other models studied, the proposed model bagged trees has considerably decreased the number of erroneous classifications.

Table 3 KNN classifier accuracy, sensitivity, and specificity for various k values

k value	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	94.20	97.88	97.19
3	94.30	97.99	96.04
5	94.50	98.30	95.30
9	93.20	99.30	93.88

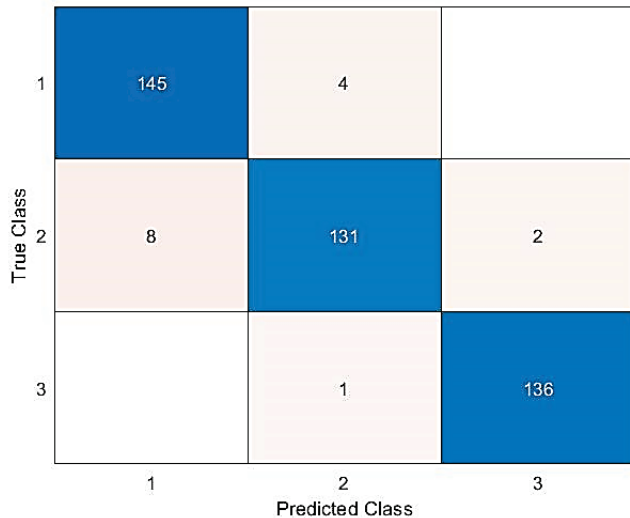


Figure 9 Test set - bagged trees confusion matrix

5 CONCLUSION

This study analyzes the prediction of positive, negative, and neutral emotions based on EEG signals, utilizing MLP, SVM, PNN, KNN, and decision tree classification methods. Depending on whether the PCA approach is utilized or not, the training has been evaluated on various properties of the dataset. The accuracy in the MLP classification, 150 most important extracted features using PCA is achieved to be 90.21 percent. Thereby, when the PCA technique is not used, MLP-based training and testing become more successful. When the learning rate parameter is increased, the accuracy of the AdaBoost-classified emotion prediction has also increased. The proposed model, bagged trees, significantly reduces the occurrence of incorrect classifications.

To the best of the authors' knowledge, the bagged tree ensemble approach has achieved a success rate of 98.60 percent. In addition, the classification of the dataset using PNN yielded a success rate of 94.32 percent. However, the system model trained to use the 100 most essential factors extracted from PCA and classified using the MLP classifier has the lowest success rate of 89.86 percent. The Bagged Trees technique used for classification, improves the performance by 8 percent on the MLP classifier using the PCA approach, while the PNN classifier is used, a 4 percent performance increase is achieved.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

6 REFERENCES

- [1] Bird, J. J., Manso, L. J., Ribiero, E. P., Ekart, A., & Faria, D. R. (2018). A study on mental state classification using EEG-based brain-machine interface. *The 9th International Conference on Intelligent Systems, IEEE*, 1541-1672. <https://doi.org/10.1109/IS.2018.8710576>
- [2] Alakus, T. B. & Turkoglu, I. (2018). EEG Based Emotion Analysis Systems. *TBV Journal of Computer Science and Engineering*, 11(1), 26-39.
- [3] Szwoch, W. (2013). Using Physiological Signals for Emotion. *6th International Conference on Human System Interactions (HSI)*, 556-561.
- [4] Alakus, T. B. & Turkoglu, I. (2019). Pozitif ve Negatif Duyguların Ayrımında Etkili EEG Kanallarının Dalgacık Dönüşümü ve Destek Vektör Makineleri ile Belirlenmesi. *International Journal of Informatics Technologies*, 12(3), 229-237. <https://doi.org/10.17671/gazibtd.482939>
- [5] Ramirez, R. Planas, J., Escude, N., Mercade, J., & Farriols, C. (2018). EEG-Based Analysis of the Emotional Effect of Music Therapy on Palliative Care Cancer Patients. *Front Psychol.* <https://doi.org/10.3389/fpsyg.2018.00254>
- [6] Bird, J. J., Ekart, A., & Faria, D. R. (2019). Mental Emotional Sentiment Classification with an EEG-based Brain-machine Interface. *The International Conference on Digital Image and Signal Processing (DISP'19)*. <https://doi.org/10.1109/IS.2018.8710576>
- [7] Jain, V., Parab, K., Kalgutkar, S., & Sonkusare, R. (2021). EEG Brainwave Emotion Detection Using Stacked Ensembling Method. *The 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE. <https://doi.org/10.1109/ICCCNT.2021.9579818>
- [8] Mehmood, R. M., & Lee, H. J. (2016). Emotion Recognition from EEG Brain Signals Based on Particle Swarm Optimization and Genetic Search. *IEEE International Conference on Multimedia and Expo Workshops*. <https://doi.org/10.1109/ICMEW.2016.7574682>
- [9] Ralekar, C., Gandhi, T. K., & Roy, S. K. (2016). Emotion Classification from EEG Signals. *The 3rd International Conference on Computing for Sustainable Global Development*, 2543-2546.
- [10] Koelstra, S., Yazdani, A., Soleymani, M., Mühl, C., Lee, J. S., Nijholt, A., Pun, T., Ebrahimi, T., & Patras, I. (2010). Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos. *Int. Conf. on Brain Informatics, Springer*, 89-100. https://doi.org/10.1007/978-3-642-15314-3_9
- [11] Ozcan, C. & Cizmeci, H. (2020). EEG Based Emotion Recognition with Convolutional Neural Networks. *The 28th Signal Processing and Communications Applications Conference (SIU)*. <https://doi.org/10.1109/SIU49456.2020.9302498>
- [12] Atasoy, H., Yildirim, S., Yildirim, E., & Kutlu, Y. (2014). Emotion Recognition from EEG Signals Using Fractal Dimension And Wavelet Transform. *Elektrik, Elektronik, Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu (ELECO)*, 650-653.

- [13] Gorur, K., Bozkurt, M. R., Bascil, M. S., & Temurtas, F. (2021). Tongue-Operated Biosignal over EEG and Processing with Decision Tree and kNN. *Academic Platform Journal of Engineering and Science*, 9(1), 112-125. <https://doi.org/10.21541/apjes.583049>
- [14] Yazar, I., Yavuz, H. S., & Cay, M. A. (2009). Temel Bileşen Analizi Yönteminin ve Bazı Klasik ve Robust Uyarlamalarının Yüz Tanıma Uygulamaları. *Journal of Engineering and Architecture Faculty of Eskişehir Osmangazi University*, 22(1), 49-63.
- [15] Bascil, M. S., Tesneli, A. Y., & Temurtas, F. (2016). Spectral feature extraction of EEG signals and pattern recognition during mental tasks of 2-D cursor movements for BCI using SVM and ANN. *Australasian Physical and Engineering Sciences in Medicine*, 39(3), 665-676. <https://doi.org/10.1007/s13246-016-0462-x>
- [16] Gorur, K., Bozkurt, M. R., Bascil, M. S., & Temurtas, F. (2018). Glossokinetic potential based tongue-machine interface for 1-D extraction. *Australasian Physical and Engineering Sciences in Medicine*, 41, 379-391. <https://doi.org/10.1007/s13246-018-0635-x>
- [17] Alan, A. & Karabatak, M. (2020). Veri Seti - Sınıflandırma İlişkisinde Performansa Etki Eden Faktörlerin Değerlendirilmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 32(2), 531-540. (in Turkish) <https://doi.org/10.35234/fumbd.738007>
- [18] Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Mach. Learn*, 20, 273-297. <https://doi.org/10.1007/BF00994018>
- [19] Albayrak, A. S. & Koltan, Y. S. (2009). Veri madenciliği: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(1), 31-52. (in Turkish)
- [20] Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [21] Friedman, J. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [22] Gorur, K., Bozkurt, M. R., Bascil, M. S., & Temurtas, F. (2019). GKP Signal Processing Using Deep CNN and SVM for Tongue-Machine Interface. *Traitement du Signal*, 36(4), 319-329. <https://doi.org/10.18280/ts.360404>

Authors' contacts:

F. Kebire Bardak, Res. Assis.

(Corresponding author)

Department of Electrical and Electronics Engineering,

Bandırma Onyedi Eylül University,

Balıkesir, Bandırma, Turkey

+90 266 606 3554, kbardak@bandirma.edu.tr

M. Nuri Seyman, Assoc. Prof. Dr.

Department of Electrical and Electronics Engineering,

Bandırma Onyedi Eylül University,

Balıkesir, Bandırma, Turkey

+90 266 606 3512, mseyman@bandirma.edu.tr

Feyzullah Temurtas, Prof. Dr.

(1) Department of Electrical and Electronics Engineering,

Bandırma Onyedi Eylül University,

Balıkesir, Bandırma, Turkey

(2) AINTELIA Artificial Intelligence Technologies Company,

16240, Bursa, Turkey

+90 266 606 3501, ftemurtas@bandirma.edu.tr

AFMB-Net: DeepFake Detection Network Using Heart Rate Analysis

A. Vinay, Nipun Bhat, Paras S. Khurana*, Vishruth Lakshminarayanan, Vivek Nagesh, S. Natarajan, T. B. Sudarshan

Abstract: With advances in deepfake generating technology, it is getting increasingly difficult to detect deepfakes. Deepfakes can be used for many malpractices such as blackmail, politics, social media, etc. These can lead to widespread misinformation and can be harmful to an individual or an institution's reputation. It has become important to be able to identify deepfakes effectively, while there exist many machine learning techniques to identify them, these methods are not able to cope up with the rapidly improving GAN technology which is used to generate deepfakes. Our project aims to identify deepfakes successfully using machine learning along with Heart Rate Analysis. The heart rate identified by our model is unique to each individual and cannot be spoofed or imitated by a GAN and is thus susceptible to improving GAN technology. To solve the deepfake detection problem we employ various machine learning models along with heart rate analysis to detect deepfakes.

Keywords: Convolutional Neural Network (CNN); Celeb-DF; Deepfake; Frame Normalisation; Generative Adversarial Networks (GAN); Motion Magnification Spatio-Temporal Map (MMST Map); MBConv Blocks; PhotoPlethysmography (PPG)

1 INTRODUCTION

Deepfakes are fake media in which a person's likeness is replaced with that of some-one else in the media. Though content faking isn't recent, deepfakes use advanced ML and AI generation techniques to influence or produce social media content that can easily deceive the naked eye. Using generative neural network architectures like autoencoders or GANs, is the key machine learning approach used to generate deepfakes (GANs). The neural network used to synthesize deepfakes are autoencoders. Autoencoders comprise of an encoder and a decoder. Encoders reduce dimensions of the images and the decoder uses these reduced dimensions to reconstruct the images. Deepfakes use this autoencoder architecture by encoding a human into the latent space with a universal encoder. Attaching a GAN to the decoder is a standard up-grade to this architecture. In an adversarial relationship, a GAN trains a generator and a discriminator. As a result, the generator produces images that closely resemble reality, as any flaws are detected by the discriminator. Deepfake generation algorithms are constantly evolving making deepfakes difficult to detect. This is because once a flaw is detected in the generation algorithm by the detection algorithms, it is easily fixed. Deepfakes are being used for both positive and negative purposes.

It has become really important to be able to classify if a video is a deepfake or not. While there are many neural network-based methods to identify GAN generated deepfakes, none of these methods can keep up with the improving deepfake generation technology as well as generalize to an unknown domain.

2 RELATED WORK

Ipek Ganiyusufoglu et al. [1] proposes the usage of spatio-temporal features, modelled by 3D CNNs to maximize generalization. These spatio-temporal models can transfer some of their understanding of already learned deepfake methods onto new unseen fake kinds. Xuan Hau Nguyen et al. [2] used 3D convolution kernels to construct deep 3D CNN models that extract spatio-temporal features from

frames sequence for detecting deepfakes. David Güera et al. [3] have used RNN in their approach to detect deepfakes. Ekraam Sabir et al. [4] have used recurrent-convolutional network combinations. ResNet and DenseNet are employed for the CNN component in different experiments, along with unidirectional, bi-directional and multi RNNs on FF++ dataset. Lingzhi Li et al. [5] used HRNet in their paper for classification of deepfakes. It works under the assumption that there is a boundary when deepfake image is formed by blending two images. This leads to the existence of intrinsic image discrepancies across the blending boundary. The face X-ray of the input produces a greyscale image, which reveals if it can be decomposed. Pavel Korshunov et al. [6] have used the VGG network to determine deepfakes from generating their own dataset by using GAN based face-swapping algorithm. Forrest N. Ian-dola et al. [7] have published the SqueezeNet paper, which aims to reduce the number of parameters while retaining the accuracy. Xiangyu Zhang et al. [8] have proposed ShuffleNet, which combines modern convolutional neural networks with ShuffleNet Units stacked in stages, which makes use of sparse connections and grouped convolutions. The ShuffleNet unit introduces channel shuffle operations for information cross-flow between the channel groups in grouped convolutions. The paper aims for efficient computations while maintaining a sturdy performance in detection of deepfakes. Mark Sandler et al. [9] proposed the MB-Conv Blocks in the MobileNet. The MBConv Block performs feature space exploration through the use of its inverted residual blocks which help extract more information from the frames through expansion and squeezing. Peng Zhou et al. [10] propose a method of using a two-stream neural network for deepfake detection where in one stream extracts the facial features from the input and the other stream extracts the steganalysis features. Darius Afchar et al. [11] propose a neural network architecture which is widely used in the domain of deepfakes - MesoNet, which is powerful and compact. Huy H. Nguyen et al. [12] proposes the use of capsule networks to detect deepfakes which is proven to perform better at the video level than at the frame level. Chih-Chung Hsu et al. [13] a discriminator is used in this compared to learning from a binary classifier to effectively and easily detect deepfakes. Lakshmanan

Nataraj et al. [14] published their work where co-occurrence matrices are extracted on three color channels and a CNN Model is trained. Huy H. Nguyen et al. [15] proposed the concept of multitask learning which is used to detect areas of manipulation in the fake video. Run Wang et al. [16] published FakeSpotter where they monitored the behaviour of the neuron, analyzed the pattern in which the neurons are activated to capture features. Andreas Rossler et al. [17] introduces a new deepfake dataset - FaceForensics++ along with Patrick Kwon et al. [18] who introduce the KoDF, a large-scale Korean deepfake dataset. Bojia Zi et al. [19] have also introduced another challenging real-world dataset for deepfake detection - WildDeepfake. Thanh Thi Nguyen et al. [20] summarized of techniques of deepfake creation and detection at that time. Recently many new approaches have been introduced in the deepfake detection domain like Domain adaptation [21], GANs [22, 23], Generalization [24], Dynamic face augmentation [25], N-Shot learning [26, 27], Reverse engineering [28-30], Self-supervised learning [31, 32] and Vision transformers [33, 34].

2.1 Heart Rate Analysis

Lijun Yin et al. [35] performed heart rate analysis to detect deepfakes, their method-ology PPG cells are generated by extracting facial features using face detectors, identifying faces with highest PPG stability and extracting regions of interest, aligning non-linear ROI into a rectangular image so that they form a face video, each image is now divided into 32 squares and the chromaticity of the PPG signal in each square is found, using the 32 PPG chromaticity values, the chromaticity of the entire image is found. Roberto Caldelli et al. [36] work aims to exploit inter-frame correlations which are used as features for classification. J. Hernandez-Ortega et al. [41] proposed the Motion model and Appearance model with the normalized frames as input as published in their work. The paper published by Steven Fernandes et al. [37] proposes a simpler method for extraction of heart rates from deepfakes compared to the existing techniques using a neural ODE. The heart rate of the original videos is learnt by the Neural ODE using which it can predict the heart rate of all manipulated versions of the original video. Creating a deepfake dataset using a commercial website, 320 videos from DeepFake TIMI are also used. Heart rates of the facial videos were extract-ed. This could be performed using any of the methods listed below:

- Measuring changes in skin colour caused by blood flow patterns, this is performed by finding the average RGB values in the regions of interest.
- The heart rate was calculated using the Fourier Fast Transform The average optical intensity in the forehead was found due to its stability across all frames and the heart rate was found.

Colours in the frames are amplified and the intensity of these colours are then used to identify the flow of blood. The colour red is given special importance as it implies more flow of blood in the given region.

2.2 Attention Feature Map

Hyeonseong Jeon et al. [38] proposed FDFTNet which is a mechanism that is employed to increase the accuracy of popular CNN Models. The core of this approach is to introduce an image-based self-attention module called attention feature map that uses only the attention module and the downsampling layer. This module is added to the pre-trained model and fine-tuned on data to search for new features in the feature space to detect fake images.

3 METHODOLOGY

3.1 Pre-processing Stage

Our study uses the heart rate analysis method in order to detect deepfakes as proposed by Hua Qi et al. [42]. The videos used are from the Celeb-DF dataset [39].

3.1.1 Extracting Faces and Alignment of Frames

- Preparing the dlib library for face detection.
- Preparing MTCNN [40] for face extraction.
- After detecting and extracting the faces from each frame, we process the frame by using 81 face landmarks predictor and remove the eyes and the inner mouth regions from the face.
- We save the newly extracted and processed face as a jpg file, as shown in Fig. 1, representative of each frame along with the facial landmarks as a NumPy file. (Each jpg image from a given video are stored in sequential order in each folder representative of that video)



Figure 1 Extracted face after aligning and removing eyes and mouth

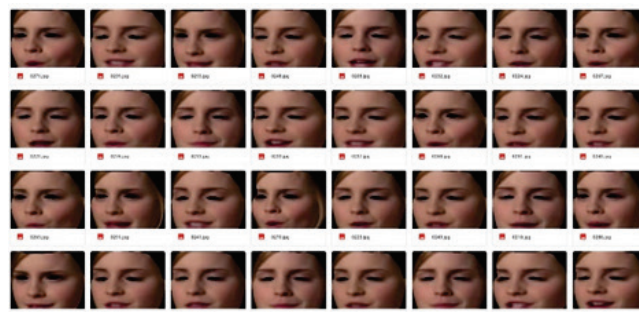


Figure 2 Resized faces

3.1.2 Resizing the Frames

- The output from previous step is the input in this step. These files are all of different dimensions. The aim of this step is to make all the frames of equal dimensions.

- The original dimensions of each jpg file representing a frame are resized to the dimensions 300×300 .
- The resized frames, as shown in Fig. 2, are now saved as jpg files.

3.1.3 Generating Face Video

- We iterate through the folders containing frames of the aligned faces.
- Using the ffmpeg tool, we now combine these resized and aligned frames to generate a face video.

3.1.4 Generating Motion Magnified Video

- For this step, we use an open-source tool called PyEVM. A frame from a motion magnified video is shown in Fig. 3.
- Spatial decomposition is performed on the video using a Laplacian Pyramid.
- The spatially decomposed videos are now passed through a temporal bandpass filter.
- The resulting signal is amplified.
- This amplified signal is now added back so as to generate the motion magnified video.



Figure 3 A frame in motion magnified video

3.1.5 Creating Motion Magnified Spatio-Temporal Map (MMST Map)

- Each frame in the motion magnified video is divided in 5×5 grid of Region of Interest (ROI) blocks.
- Pooling operation is performed for each block and each color channel of the frame.
- The MMST map is generated for each video in which each row represents a particular ROI block and the corresponding columns of the row represents the temporal variation of the block across all the frames present in the video. A visual representation of the MMST Map is shown in Fig. 4.

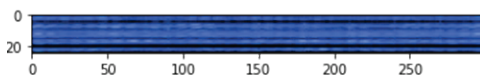


Figure 4 MMST Map

3.1.6 Assigning Spatial and Temporal Weights to the MMST Map

In this step, spatial and temporal weights are assigned to the MMST Map. These weights are used to help with detection of deepfakes.

- The MMST map is passed through a pretrained MesoNet for the spatial weights and the weight matrix W is generated.
- The MMST map is passed through a pretrained LSTM for the temporal weights and the temporal weight matrix T is generated.
- The spatial weight matrix S is multiplied with the temporal weight matrix T , resulting in W , as shown in Eq. (1).

$$W = S * T^T \quad (1)$$

- The MMST Map, represented by X is now multiplied by the product of Step 4 which is used as the input for the detection models, as shown in Eq. (2).

$$A = X * W^T \quad (2)$$

3.2 Detection Stage

3.2.1 AFMB-Net - Attention Feature Map + ResNet-18 + MB-Conv Blocks

We are using an attention feature map along with MB-Conv blocks and ResNet-18 as represented in Fig. 5. The attention feature map is a method to help ameliorate the accuracy of the baseline model along with the MB-Conv Blocks that perform feature space extraction thus further helping us to extract the most relevant features in identification of deepfakes. By using the attention feature map, we can develop long term dependencies through the use of softmax outputs thus helping reduce computations. If the CNN were to be used to develop the same dependencies, it would take a large number of computations as it performs computation only one time and only sees the kernel size.

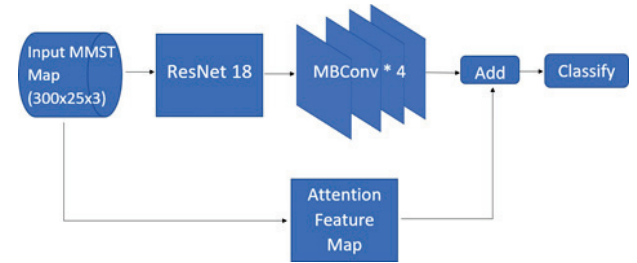


Figure 5 AFMB-Net Architecture

3.2.2 FrameNorm - Frame Normalization Method

The Motion model and Appearance model, shown in Fig. 6, as proposed in the works of J. Hernandez-Ortega et al. [41] are used. This model is slightly different as we have taken the input of this model as the motion magnified videos instead of the MMST Map.

- The difference of frame at time t and at time $t-1$ of motion magnified video is taken and normalization is performed on the difference.
- The frame at time t is normalized and saved simultaneously with the previous step.

- The output of first step is then passed to the motion model and the output of second step is passed to the spatial model.
- The outputs of both are multiplied and passed through a fully connected dense layer.

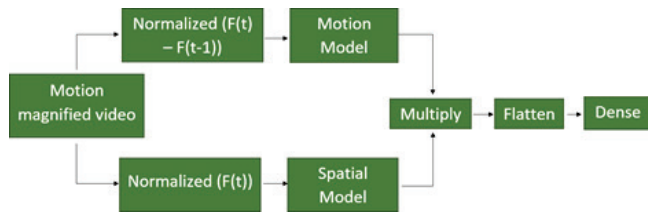


Figure 6 FrameNorm Architecture

4 RESULTS AND DISCUSSION

Our results obtained are depicted through Tab. 1. The AFMB-Net architecture uses the ResNet-18 model as a backbone. The ResNet is a powerful CNN architecture widely used in image classification. This is mainly due to use of deep residual networks in the architecture which reuse the activation functions from the previous layers. The ResNet-18 is a lightweight version of the ResNet that provides high accuracy. The Celeb-DF dataset is a challenging dataset for deepfake detection where most models fall short. In order to enhance the performance of the ResNet for deepfake detection, we have used MBConv blocks and an Attention Feature Map. The Attention Feature Map works by giving special attention to the RGB channels. The MBConv blocks work by performing the squeezing and expansion of the input. This performs feature space extraction on the input and thus further helps extract important features. The AFMB-Net architecture gave us a training accuracy and testing accuracy of 97.91% and 95.19% respectively. Using the above-mentioned intuition, the ResNet-18, MBConv blocks and the Attention Feature Map have helped achieve this accuracy.

Table 1 Models Performance

Model Architecture	Metric	
	Training Accuracy	Testing Accuracy
AFMB-Net	0.9791	0.9519
FrameNorm	0.8538	0.7538

The FrameNorm architecture on the other hand gave us a testing accuracy of 75.38% and training accuracy of 85.38%. By the use of the Motion model, we aimed to extract maximal information between two consecutive frames in order to help us detect any inconsistencies over time (similar to the Heart Rate Analysis). The use of motion magnified videos as input allowed us to further enhance any inconsistencies between consecutive frames along the temporal and spatial domains. Similarly, the Appearance model was used to detect any inconsistencies within the spatial domain of a single frame. The use of motion magnified videos further helped detect any inconsistencies. This combination of Motion and Appearance model helped us classify deepfakes.

5 CONCLUSION

Through the course of our work, we have aimed to show the use of the heart rate analysis method in the detection of deepfakes. We believe that this method is fool-proof as the heart rates of individuals cannot be generated by GANs and are unique to each person. The proposed methodology is to implement a biological signal extractor in the form of PPG cells or spatio-temporal cells which is passed through different Machine Learning Models. The dataset used for our experimentation is the Celeb-DF dataset. This dataset provides high quality deepfake videos that are difficult to tell apart from real videos using the naked eye alone. Using the attention feature map, a machine learning method that is used to increase the accuracy of Convolutional Neural Networks, the MBConv Blocks for feature space extraction and developing long term dependencies, we see an increase in the accuracy achieved on the ResNet-18, 95%. The FrameNorm architecture however gave us a testing accuracy of 75.38%.

Some limitations encountered were:

- Objects in front of the Face: Any objects in front of the face will cause the model to not be able to identify/detect the face thus hindering the creation of MMST Map.
- MMST Map Dimensions: The size of the MMST Map is of the dimensions $300 \times 25 \times 3$, this may have led to lower levels of accuracy than if the model had height and width of same or similar dimensions.
- As illustrated by the work carried out, it is apparent that our intuition on the use of heart rates for the detection of deepfakes is rightly justified and yields good results in the detection of deepfakes.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

6 REFERENCES

- [1] Ganiyusufoglu, I., Ngo, L., M., Savov, N., Karaoglu, S., & Gevers, T. (2020). Spatio-temporal features for generalized detection of deepfake videos. *arXiv*: 2010.11844. <https://doi.org/10.48550/arXiv.2010.11844>
- [2] Nguyen, X. H., Tran, T. S., Nguyen, K. D., Truong, D. T., et al. (2021). Learning spatio-temporal features to detect manipulated facial videos created by the deepfake techniques. *Forensic Science International: Digital Investigation*, 36, 301108. <https://doi.org/10.1016/j.fsdi.2021.301108>
- [3] Güera, D. & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. *The 15th IEEE international conference on advanced video and signal based surveillance (AVSS)*, 1-6. <https://doi.org/10.1109/AVSS.2018.8639163>
- [4] Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., & Natarajan, P. (2019). Recurrent convolutional strategies for face manipulation detection in videos. *arXiv*: 1905.00582. <https://doi.org/10.48550/arXiv.1905.00582>
- [5] Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., & Guo, B. (2020). Face x-ray for more general face forgery detection.

- CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5000-5009.
<https://doi.org/10.1109/CVPR42600.2020.00505>
- [6] Korshunov, P. & Marcel, S. (2018). Deepfakes: a new threat to face recognition? Assessment and detection. *CoRR*.
 - [7] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). Squeezenet: Alexnet-level accuracy with 50× fewer parameters and 0.5 mb model size. *arXiv*: 1602.07360. <https://doi.org/10.48550/arXiv.1602.07360>
 - [8] Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6848-6856.
<https://doi.org/10.1109/CVPR.2018.00716>
 - [9] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4510-4520.
<https://doi.org/10.1109/CVPR.2018.00474>
 - [10] Zhou, P., Han, X., Morariu, V. I., & Davis, L. S. (2017). Two stream neural networks for tampered face detection. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1831-1839.
<https://doi.org/10.1109/CVPRW.2017.229>
 - [11] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). Mesonet: a compact facial video forgery detection network. *IEEE International Workshop on Information Forensics and Security (WIFS)*, 1-7. <https://doi.org/10.1109/WIFS.2018.8630761>
 - [12] Nguyen, H. H., Yamagishi, J., & Echizen, I. (2019). Using capsule networks to detect forged images and videos. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2307-2311.
<https://doi.org/10.1109/ICASSP.2019.8682602>
 - [13] Hsu, C.-C., Lee, C.-Y., & Zhuang, Y.-X. (2018). Learning to detect fake face images in the wild. *International Symposium on Computer, Consumer and Control (IS3C)*, 388-391.
<https://doi.org/10.1109/IS3C.2018.00104>
 - [14] Nataraj, L., Mohammed, T. M., Manjunath, B., Chandrasekaran, S., Flenner, A., Bap-py, J. H., & Roy-Chowdhury, A. K. (2019). Detecting GAN generated fake images using co-occurrence matrices. *Electronic Imaging*, 2019(5), 532-1-532-7.
<https://doi.org/10.2352/ISSN.2470-1173.2019.5.MWSF-532>
 - [15] Nguyen, H. H., Fang, F., Yamagishi, J., & Echizen, I. (2019). Multi-task learning for detecting and segmenting manipulated facial images and videos. *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2019, 1-8.
<https://doi.org/10.1109/BTAS46853.2019.9185974>
 - [16] Wang, R., Juefei-Xu, F., Ma, L., Xie, X., Huang, Y., Wang, J., & Liu, Y. (2019). FakeSpotter: A simple yet robust baseline for spotting AI-synthesized fake faces. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, (IJCAI-20)*, 3444-3451. <https://doi.org/10.24963/ijcai.2020/476>
 - [17] Rossler, D., Cozzolino, L., Verdoliva, C., Riess, J., Thies, M., & Niessner, M. (2019). *FaceForensics++*: Learning to detect manipulated facial images. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1-11.
<https://doi.org/10.1109/ICCV.2019.00009>
 - [18] Kwon, P., You, J., Nam, G., Park, S., & Chae, G. (2021). KoDF: A large-scale Korean deepfake detection dataset. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 10724-10733.
<https://doi.org/10.1109/ICCV48922.2021.01057>
 - [19] Zi, B., Chang, M., Chen, J., Ma, X., & Jiang, Y.-G. (2020). Wilddeepfake: A challenging real-world dataset for deepfake detection. *Proceedings of the 28th ACM International Conference on Multimedia*, 2382-2390.
<https://doi.org/10.1145/3394171.3413769>
 - [20] Nguyen, T. T., Nguyen, C. M., Nguyen, D. T., Nahavandi, S. et al. (2019). Deep learning for deepfakes creation and detection: A survey. *arXiv*: 1909.11573.
<https://doi.org/10.48550/arXiv.1909.11573>
 - [21] Chen, B. & Tan, S. (2021). Featuretransfer: Unsupervised domain adaptation for cross-domain deep-fake detection. *Security and Communication Networks*, 2021.
<https://doi.org/10.1155/2021/9942754>
 - [22] Borji, A. (2021). Pros and cons of GAN evaluation measures: New developments. *arXiv*: 2103.09396.
<https://doi.org/10.48550/arXiv.2103.09396>
 - [23] Xuan, X., Peng, B., Wang, W., & Dong, J. (2019). On the generalization of GAN image forensics. *Chinese conference on biometric recognition*, 134-141.
https://doi.org/10.1007/978-3-030-31456-9_15
 - [24] Kim, M., Tariq, S., & Woo, S. S. (2021). Fretal: Generalizing deepfake detection using knowledge distillation and representation learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1001-1012. <https://doi.org/10.1109/CVPRW53098.2021.00111>
 - [25] Das, S., Seferbekov, S., Datta, A., Islam, M., Amin, M. et al. (2021). Towards solving the deepfake problem: An analysis on improving deepfake detection using dynamic face augmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3776-3785.
<https://doi.org/10.1109/ICCVW54120.2021.00421>
 - [26] Aneja, S. & Niessner, M. (2020). Generalized zero and fewshot transfer for facial forgery detection. *arXiv*: 2006.11863.
<https://doi.org/10.48550/arXiv.2006.11863>
 - [27] Liu, Y., Stehouwer, J., Jourabloo, A., & Liu, X. (2019). Deep tree learning for zero-shot face anti-spoofing. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4680-4689.
<https://doi.org/10.1109/CVPR.2019.00481>
 - [28] Skibba, R. (2020). Accuracy eludes competitors in facebook deepfake detection challenge. *Engineering*, 6(12), 1339-1340.
<https://doi.org/10.1016/j.eng.2020.10.008>
 - [29] Tariq, S., Jeon, S., & Woo, S. S. (2021). Am I a real or fake celebrity? Measuring commercial face recognition web APIs under deepfake impersonation attack. *arXiv*: 2103.00847.
<https://doi.org/10.48550/arXiv.2103.00847>
 - [30] Asnani, V., Yin, X., Hassner, T., & Liu, X. (2021). Reverse engineering of generative models: Inferring model hyperparameters from generated images. *arXiv*: 2106.07873.
<https://doi.org/10.48550/arXiv.2106.07873>
 - [31] Aneja, S., Bregler, C., & Niessner, M. (2021). COSMOS: Catching out-of-context misinformation with self-supervised learning. *arXiv*: 2101.06278.
<https://doi.org/10.48550/arXiv.2101.06278>
 - [32] R  thlingsh  fer, V., Sharma, V., & Stiefelhagen, R. (2019). Self-supervised face-grouping on graphs. *Proceedings of the 27th ACM International Conference on Multimedia*, 247-256.
<https://doi.org/10.1145/3343031.3351071>
 - [33] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S. et al. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv*: 2010.11929. <https://doi.org/10.48550/arXiv.2010.11929>

- [34] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, 9992-10002. <https://doi.org/10.1109/ICCV48922.2021.00986>
- [35] Ciftci, U. A., Demir, I., & Yin, L. (2020). How do the hearts of deep fakes beat? Deep fake source detection via interpreting residuals with biological signals. *IEEE International Joint Conference on Biometrics (IJCB)*, 1-10. <https://doi.org/10.1109/IJCB48548.2020.9304909>
- [36] Caldelli, R., Galteri, L., Amerini, I., & del Bimbo, A. (2021). Optical flow based CNN for detection of unlearned deepfake manipulations. *Pattern Recognition Letters*, 146, 31-37. <https://doi.org/10.1016/j.patrec.2021.03.005>
- [37] Fernandes, S., Raj, S., Ortiz, E., Vintila, I., Salter, M., Urosevic, G., & Jha, S. (2019). Predicting heart rate variations of deepfake videos using neural ode. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. <https://doi.org/10.1109/ICCVW.2019.00213>
- [38] Jeon, H., Bang, Y., & Woo, S. S. (2020). FDFTnet: Facing off fake images using fake detection fine-tuning network. *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, 416-430. https://doi.org/10.1007/978-3-030-58201-2_28
- [39] Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2020). Celeb-df: A large-scale challenging dataset for deepfake forensics. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3207-3216. <https://doi.org/10.1109/CVPR42600.2020.00327>
- [40] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503. <https://doi.org/10.1109/LSP.2016.2603342>
- [41] Hernandez-Ortega, J., Tolosana, R., Fierrez, J., & Morales, A. (2020). DeepFakesON-Phys: Deepfakes detection based on heart rate estimation. *arXiv*: 2010.00400. <https://doi.org/10.48550/arXiv.2010.00400>
- [42] Qi, H., Guo, Q., Juefei-Xu, F., Xie, X., Ma, L., Feng, W., Liu, Y., & Zhao, J. (2020). DeepRhythm: Exposing deepfakes with attentional visual heartbeat rhythms. *Proceedings of the 28th ACM International Conference on Multimedia*, 4318-4327. <https://doi.org/10.1145/3394171.3413707>

Authors' contacts:

A. Vinay, Professor
PES University,
Ring Road Campus, Bangalore, Karnataka, India
a.vinay@pes.edu

Nipun Bhat
PES University,
Ring Road Campus, Bangalore, Karnataka, India
bhatnipun29@gmail.com

Paras S. Khurana
(Corresponding author)
PES University,
100 Feet Ring Road, BSK III Stage, Ring Road Campus,
Bangalore-560085, Karnataka, India
paras.s.khurana@gmail.com

Vishruth Lakshminarayanan
PES University,
Ring Road Campus, Bangalore, Karnataka, India
vishruth2270@gmail.com

Vivek Nagesh
PES University,
Ring Road Campus, Bangalore, Karnataka, India
vivek.nagesh1@gmail.com

S. Natarajan, Doctor
PES University,
Ring Road Campus, Bangalore, Karnataka, India
natarajan@pes.edu

T. B. Sudarshan, Doctor
PES University,
Ring Road Campus, Bangalore, Karnataka, India
sudarshan@pes.edu

Performance Comparison of Open Source and Commercial Computing Tools in Educational and Other Use — Scilab vs. MATLAB

Matija Mikac*, Robert Logožar, Miroslav Horvatić

Abstract: In this paper, the authors compare the features and the overall performance of the two high-level numerical computing and modeling software environments: the freeware Scilab and commercially available industry-standard MATLAB. The motivation for the work emanated from the educational use of these tools at the college and university level, but with a perspective to their professional and scientific use as well. Their performance is tested by measuring the execution times of several combined-task benchmarks implemented as test functions, built upon nine common numerical tasks that are often found in programs for solving standard engineering problems. They include basic algebra and matrix calculations, signal generation, signal analysis, and storing and retrieving data to and from the hard disk drive. Although MATLAB outperforms Scilab in all the benchmarks except the disk file manipulations, in the presumed vectorization versions of the benchmarks, it is not for much. The overall performance of the freeware rival is very satisfactory, making it a good choice not only for educational use but also for scientific and professional purposes, especially when funding is critical.

Keywords: calculation benchmarks; mathematical and modeling software; MATLAB; programming; performance comparison; Scilab

1 INTRODUCTION

In contemporary education oriented to science, technology, engineering, and mathematics—now covered by the widely popular acronym STEM—there are many courses in which numerical, modeling, and simulation problems are solved by computers running specialized, high-level numerical programming software. One of the most known software packages of that kind is MATLAB® [1]. Although there are discounted and student versions of this software, MATLAB is primarily a commercial product with a stern licensing policy. Thus, unless the special fees are paid, the students and teachers will not be able to use this software on their computers, outside the classrooms. This notably limits the use of MATLAB in the educational process.

The answer to the high-cost and licensing renewal problems with a commercial product can be in turning to free and open source substitutes, such as Scilab, GNU Octave, Sage FreeMat, and Maxima [2, 3, 4, 5, 6]. However, they are all mainly considered as the "MATLAB alternatives". Their functionality and performance are standardly compared to the functionality and performance of MATLAB, today's undisputed standard and the leader in the field [7].

Having mentioned that, the question that many lecturers, scientists, and professionals raise is what to choose. Is it better to stick to the proven though expensive commercial tool with its standard syntax and superior performance or to use Scilab and other similar products with possibly different syntax and less than superior performance — just for the sake of their unrestricted use and financial savings?

Without pretensions to answer this ubiquitous question unanimously, this article aims to provide an objective performance comparison of the commercial MATLAB and the free, open-source Scilab, and thus help the reader to answer the question herself.

After presenting the motivation for this paper, here we briefly describe its further contents. The second section begins with some basic information about the two compared tools: MATLAB as the educational and industry standard,

and Scilab as its alternative. The section continues with a description of the benchmarking test environment. In the third section, we outline the most important features of MATLAB and Scilab programming languages and a few useful recommendations for translating the source code from one language to another. We start the fourth section with a short survey of the previous research on MATLAB and Scilab performances and then carry on with the description of our benchmarking methodology and the original benchmarks introduced in this work. Section five presents and comments on the obtained execution times of the implemented benchmarks and compares the performances of both calculation tools. The last, sixth section concludes the paper with the final thoughts and remarks about possible future work.

2 MATLAB, SCILAB, AND THE TESTING PLATFORM

In this section, we present basic information about MATLAB and Scilab and describe the test platform that we used for their performance measurements.

2.1 MATLAB and Scilab

MATLAB—as it was already pointed out in the introduction—is almost consensually described as the leading numerical computing and simulation software nowadays. Since its appearance in the 1980s, it has been used and approved in many high-scale scientific and professional projects and has become the de-facto standard in many scientific and industrial areas, starting from mathematics, automatic control, digital signal processing, machine learning, numerical simulations of all kinds, system modeling, and many more. It is often the first choice of professionals and because of that, it is also used in higher-level educational institutions all over the world.

MATLAB releases are commercially available with different licensing options, including academic and student licensing [8]. Although these licenses are considered the low-cost versions, they are still not free, and this can present a

serious obstacle to recommending MATLAB to students. Furthermore, the "low-cost" versions, either for educational or professional use, may still be quite pricy for the buyers, especially in the underdeveloped parts of the world.

If the user's finances are critical or highly limited, one should consider the MATLAB alternatives as the way to equip more educational or professional workplaces, including those at home. The free open-source software alternative to MATLAB discussed in this paper is **Scilab**, available under GNU v3 license [9].

Scilab was initiated by two French institutions: INRIA (*Institut National de Recherche en Informatique et en Automatique*) and ENPC (*École Nationale des Ponts et Chaussées*), and was first released in 1994. Since 2003, it is governed by the *Scilab Consortium* and is now part of the ESI Group [10, 11]. Although it originated from a scientific and academic background, Scilab soon proved its usage in industry applications, too, and has evolved into a reputable numerically oriented modeling platform, with a good user interface and a large number of versatile simulation modules [12].

For several years, Scilab has been used at the University North in Varaždin in the courses Signals and Systems, Automatic Control, and Digital Signal Processing. Even after the purchase of MATLAB licenses for classrooms, it serves as a versatile numerical software that can be freely distributed to students.

2.2 Testing Platform and the Software Versions

All our benchmark tests were performed on the same Windows 10 64-bit computer with an Intel Core i3 5005U CPU, 8 GB of RAM memory, and a 1TB standard hard disk drive, which was running the following:

1. MATLAB Version 9.5 (R2018b), 64-bit release (2018);
2. Scilab Version 6.0.2, 64-bit release (2019).

At the time of performing the tests, which was in the first half of 2019, MATLAB 9.6 was released, but we had no official commercial license and kept using the 9.5 version. Both tools were installed with the standard setup and settings.

3 UNDERLYING PROGRAMMING LANGUAGES

MATLAB and Scilab software packages contain a development environment and a numerically oriented high-level programming language suitable for matrix calculations. The first part of this section describes the most important features of these programming languages. Since there are certain differences in the syntax of the MATLAB and Scilab programming languages, the second part of this section ends with a few useful recommendations for translating the source code from one language to another.

3.1 High-Level, Matrix-Based Programming Languages

Besides the usual programming forms and structures, like the data types, arithmetic and other operations, program flow (conditional statements and loops), functions, and all other aspects of the standard programming languages, the

MATLAB and Scilab programming languages have many standard and specialized functions and features for numerical and, in general, mathematical solving of scientific and technical problems.

The basic and essentially only inherent data structure of the MATLAB and Scilab languages is the *multidimensional matrix*, implemented as a multidimensional array. The standard specializations of such an array are as follows:

1. *2-dimensional $N \times M$ matrix*, which is equivalent to the common $N \times M$ matrix;
2. *1-dimensional matrix*, equivalent to a 1-row matrix or vector with N elements, or the standard 1-dim array;
3. *0-dimensional matrix*, or 1-component vector, equivalent to a scalar, i.e. a single numerical value.

The elements of the multidimensional arrays are weakly typed. The numbers are implicitly of the double-floating point type but can be simply declared as different integer or other standard data types.

This was designed with one aim in mind: to provide a software environment for (heavy) numerical matrix and vector computation. Thus, these languages provide the standard matrix and vector (array) operations, with elements that are both real and complex numbers, right "out of the box". There is no need for any additional programming or inclusion of any extra libraries and packages. Also, the MATLAB and Scilab languages offer many advanced and specialized functions and features for dealing with these structures—either as with the standard mathematical objects or merely as with data sets organized in that way. For example, an n -component 1-dimensional matrix, or a vector, can be used to represent a data series.

The MATLAB and Scilab languages were originally designed as interpreters. Later on, MATLAB introduced the possibility of the code compilation, as well as the concept of *just-in-time* (JIT) compilation [13]. Thus, although MATLAB documentation emphasizes the improved JIT compilation and its many benefits in the recent software versions, it also advises the users to write the "normal programming code", without forcing the compilation by any side tools, and to obey the standard rules for writing "good interpreted computer programs" [14]. Regarding that, Scilab programming language has been always clearly defined as being interpreted-only [15], which makes it especially sensitive to the used programming practice (cf. §4.2).

3.2 MATLAB and Scilab Source Code Conversion

The MATLAB programming language was copy-right-protected from its very beginning. Nevertheless, the Scilab programming language was—as it seems—designed with the aim to be as alike the MATLAB language as possible but without being a direct copy. Scilab provided M2SCI tools to convert the MATLAB code to Scilab code [16]. The list of equivalent functions in MATLAB and Scilab is also available, together with the conversion tips [17].

The M2SCI tools were not used in this paper. Namely, in the case of relatively simple programming, as was done here, the conversions from one language to the other could be done

simply in the code editor of the targeted software by using its find-replace function. Some of the replacements that were used in this work are listed in Tab. 1. By this method, our benchmark source codes were easily and effectively translated from one language to the other. Of course, the programmer who performs such translations should be acquainted with the syntax of both languages, because if some problems occur, she should be able to correct the code and effectively use the available user manuals.

4 BENCHMARKING METHODOLOGY

Because numerical computing and modeling software packages often solve difficult computational tasks, the measurement of their performance is of crucial interest to both their designers and users. After a survey of the existing MATLAB and Scilab benchmarking in the works of others, this section describes and justifies the concept of the benchmark programs designed by the authors and used for benchmarking purposes in this paper.

4.1 Benchmarking in the Works of Others

For measuring the performance of a particular software version installed on a certain computer platform, MATLAB offers the `bench` function [18]. It performs the required number of runs of the following six tasks: standard matrix calculations, solving the system of linear equations and the nonlinear differential equations, finding Fast Fourier Transform (FFT), and performing one 2-D and one 3-D animated graphics. The common result is returned in the form of the execution speed, which is inversely proportional to the execution time. Scilab uses the `bench_run()` function [19], which performs a large series of predefined tests and returns their execution time in milliseconds, along with the number of the test repetitions. For the performance comparison of two or more software packages, one must choose the same benchmark tests and run them in the same environment.

The first comparison of features and performances of numerical and mathematical software packages known to the authors of this paper is given in [20]. The report compared some commercial software, including MATLAB, and some free software tools, including Scilab.

In [21], MATLAB, Scilab, GNU Octave, and NSP (a descendant of the early, pre-Java version of Scilab) were compared by testing their performance on a set of originally proposed benchmarks given in the form of closed functions. The author documented the descriptions and source codes of these functions in Scilab, which makes this work a good ground for vendor-independent benchmarking. The obtained summarized results of all those tests showed that MATLAB was a winner. Scilab was better than or equal to MATLAB in 6 of the 28 tests (21.4%), with the best performances in the calculation of Fibonacci numbers and summation of harmonic series, where it was 1.58, i.e. 1.48 times faster than MATLAB. Of the remaining 22 tests (78.6%), Scilab had the worst behavior in the for-loop tests, where it was more than 124 times slower! The obvious reason for that was its lack of the JIT compilation (see also §5.1).

Table 1 Some basic differences between MATLAB and Scilab syntax

Syntax form	MATLAB	Scilab
Inline comments.	%	//
π , the Ludolf's number In MATLAB: a keyword, In Scilab: a constant.	pi	%pi
In Scilab, the keyword <code>then</code> appears after the condition in parenthesis. There is no <code>then</code> in MATLAB.	if (cond.) statement1 else statement2 end;	if (cond.) then statement1 else statement2 end
The end of function.	end	endfunction
Standard writing function with several target and format options.	fprintf	mprintf
Quicksort function.	sort	gsort

MATLAB, Scilab, and GNU Octave were further compared in [22], where the author examined how their performance is influenced by the use of different versions of BLAS (Basic Linear Algebra Subprograms) library. Scilab and Octave were tested with four different versions: RefBLAS, Atlas, OpenBLAS, and Intel's MKL (*Math Kernel Library*, a versatile library, free for non-commercial use). MATLAB was tested with the MKL only. Besides the afore-mentioned benchmark set from [21], there were two more sets composed by the other authors: `poisson` [23] and `ncrunch` [20]. That work showed the expected overall supremacy of MATLAB over Scilab and GNU Octave, especially in some of the tests.

In [24], the authors give a comparative analysis of several numerical computationally-demanding tests that were performed in MATLAB, GNU Octave, Scilab, FreeMat, R, and IDL, and run on a high-performance computing facility. Their execution time measurements (some lasting even up to 20 hours!) showed dramatically decreased performance of those 2012-Scilab with the increased problem sizes. Various special aspects of MATLAB versus its alternatives were also investigated in [25, 26].

4.2 Benchmark Methodology Used in This Work

In this paper, our primary goal is to check and compare the performance of the observed computational tools on several common vector and matrix calculations, which can be found in standard scientific professional and educational use. To that aim, we have created several original benchmark functions with the following types of calculations and operations:

- basic algebra on the number series,
- basic matrix operations,
- simple matrix equations,
- signal generation and analysis,
- disk access operations, for storing and retrieving data.

Both MATLAB and Scilab, as well as other similar computing environments, are famous for using *vectorization* [27, ch. 29], in which the standard loop-based implementation of an iterative operation on a single matrix (vector)

element is substituted by the inherently optimized single operation on multiple elements of the same type. This is the core concept of *array programming* [28]. In recent decades, it can be accomplished also without the explicit parallelism (which requires multiple processors), thanks to the CPU's *vector-based* instructions, implemented on the modern general-purpose processors [29]. In this way, the *implicit parallelism* is achieved.

For example, in scalar-based languages (like C/C++), creating a vector of N randomly generated values requires a loop to generate N single random values. In contrast to that, the `rand` function, available in both tools, does it without programming on the programmer's side. Similarly, instead of finding the minimum or maximum value in a vector by using a loop, one uses the optimized `min/max` functions. Instead of multiplying the matrices A and B by multiplying their elements within the nested `for`-loops, one simply writes the product as $A*B$, leaving the evaluation of the code and its optimization to the underlying language. This approach not only simplifies the task for programmers but also replaces either interpreted or JIT-compiled execution of the operation on a single value with highly optimized operations on multiple values. In general, as already hinted in §3.1, using vectorization by applying the math-like syntax of vector and matrix operations is the standard good-programming practice in array programming [27, ch. 29]. The advantages of using it instead of writing the standard code will be especially important for the interpreted-only languages, as is Scilab. We shall elaborate on this in sec. 5 (cf. also [30]).

MATLAB and Scilab include many functions to visualize the numerical results, but such functions, as well as all GUI-related operations, were all turned off during the benchmark test computations. No special and advanced computing techniques, like multi-threading, multi-core, or GPU-based computing were applied to any of the implemented functions. Both compared tools support several of these techniques, but in this work, we have restricted the investigation to their basic performance.

4.3 Benchmarks Designed for This Work

For the basic calculation benchmarks, we have chosen nine different common calculation tasks. Most of them have a few internal implementation variants, which will be discussed soon. These tasks were then implemented as separate functions written in the corresponding MATLAB and Scilab programming languages. The benchmarks are listed in Tab. 2. They follow the ideas of the tasks performed by the internal MATLAB and Scilab benchmark functions mentioned in §4.1, as well as implement the tasks that the authors of this paper have encountered in solving problems in the areas mentioned at the end of §2.1, in their educational, professional, and scientific practice.

As can be seen from the table, all the implemented benchmarks have the `Nrp` parameter, which defines the number of repetitions of the whole function by an additional `for`-loop. Regarding the other parameters, the most common is N . It is the vector size, i.e. the number of elements in the corresponding 1-dim array, which governs both the needed

Table 2 List of the implemented benchmarks

Abbr.	Function name / Description
b_1	<code>sumIntSquared(N, Nrp)</code> Sum of the squares of integers $1, 2, \dots, N$; implemented by using the vectorization concept, and for comparison also with a <code>for</code> -loop and <code>while</code> -loop.
b_2	<code>mtrGenMN(N, M, r, Nrp)</code> Generation of a $N \times M$ matrix populated by the floating-point elements $x_{i,j} = ij/r$, where i (j) is the row (column) index: $i = 1, 2, \dots, N$, $j = 1, 2, \dots, M$, and $r \in \mathbb{R}$ is a decimal num.
b_3	<code>mtrGenNNPow(N, k, r, Nrp)</code> Generation and calculation of the k -th power of the (square) $N \times N$ matrix, initially populated by the elements $x_{i,j} = i^2/r$, where i and j are the row and column indices, respectively: $i, j = 1, 2, \dots, N$, and $r \in \mathbb{R}$ is a decimal number.
b_4	<code>rndMtrxMulti(N, M, Nrp)</code> Multiplication of the two, $N \times M$ and $M \times N$ matrices, with randomly generated elements.
$b_{5,i}$ $b_{5,b}$	<code>rndMtrxDivInv(N, Nrp)</code> <code>rndMtrxDivBack(N, Nrp)</code> Solution of the matrix equation $Ax = b$ for the $N \times 1$ vector x , where A is a randomly generated $N \times N$ matrix, and b is a $N \times 1$ randomly generated vector. $b_{5,i}$ finds the solution by the standard method, i.e. by using the A^{-1} inverse matrix, $x = A^{-1}b$. $b_{5,b}$ finds the solution by the back-divide method (matrix operator) written as $x = A \backslash b$.
b_6 $b_{6,0c}$	<code>rndMtrxDet(N, Nrp)</code> <code>rndMtrxDet0c(N, Nrp)</code> b_6 calculates the determinant of a randomly generated square matrix of dimension $N \times N$. $b_{6,0c}$ does the same, but for a matrix with a randomly chosen zero-column (Fig. 1a and 1b).
b_7	<code>rndSort(N, Nrp)</code> Sorting the N randomly generated floating point numbers, or, equivalently, sorting the vector's N elements, using the internal implementation of the quicksort algorithm.
b_8	<code>sigAMSinFFT(Fc, Fm, Nrp)</code> Generation of an AM signal with the sinus carrier and sinus modulation signal of frequencies F_c and F_m , respectively. After that, the spectrum is calculated by using FFT.
b_9	<code>rndMtrxSaveLoad(N, Nrp)</code> Disk read/write test; a randomly generated $N \times N$ matrix is first saved to a file, and then the file is reloaded to another matrix.

memory space and the overall size of the problem that is being solved. For the square matrices, this size is N^2 . For the non-square matrices, the parameter M is added, defining the problem size of $N \times M$. The meaning of other parameters is explained in Tab. 2.

The calculation results of the benchmark functions are returned via two output vectors: D , of size `Nrp`, for the function results, and T , of size 2 for the time measurements. With them, the functions with p parameters (and with $p_1 = p - 1$) are declared in MATLAB and Scilab as:

function [D, T] = *bfuncName*($N, \text{Par2}, \dots, \text{Parp1}, \text{Nrp}$) .

The return parameters are defined in the square brackets after the keyword `function`. The function name is followed by the list of input parameters in parenthesis. For the i -th repetition of the function, $i = 1, 2, \dots, N_{rp}$, the function result is returned via the i -th element of vector D . Via the two components

of vector T , the overall T_{tot} execution time for the N_{rp} repetitions and the average T_{avg} time per one repetition of the benchmark routine are returned.

For some benchmarks that operate on the N -size vectors or on the $N \times M$ (N^2) matrices, besides the “default,” vectorized form, we have provided the alternative, loop-form. Additionally, the same benchmarks were implemented with different data types. All this was done to investigate the influence of different factors on benchmark performance.

In the following sub-subsections, we shall first briefly discuss the implemented benchmarks and then present the source code of a selected exemplary benchmark function.

4.3.1 Description and Discussion of the Benchmarks

Benchmark b_1 sums the squares of the successive integers. It is one of the benchmarks realized in the vectorized version and in the version with the standard for- and while-loops. Furthermore, it has versions with 64-bit integer and 64-bit floating point types. Benchmark b_2 creates an $N \times M$ matrix and populates it with the elements whose values are calculated from their indices and a given constant real number as $x_{i,j} = ij/r$. Again, besides the version with vectorization, it has the loop-based version, too. Benchmark b_3 creates a square, $N \times N$ matrix, populates it with the elements whose values are calculated in a similar way as in b_2 , and then, additionally, calculates the k -th power of the matrix.

Benchmarks from b_4 to b_7 create the matrices initially populated with randomly generated elements and — on top of that — perform some additional calculations. Benchmark b_4 multiplies the two generated matrices. There are two versions of benchmark b_5 , which solve the linear system of equations, $Ax = b$, in two ways: $b_{5,i}$ by finding the inverse matrix of A , and $b_{5,b}$ by using the back-divide method (operator) for the matrix division (cf. [31, 32]). Benchmark b_6 calculates the determinant of a matrix by applying the `det` function (valid in both tools). The standard version of the benchmark forms the matrix with the randomly generated elements. In addition to that, the variant $b_{6,0c}$ inserts a randomly chosen zero-column into the matrix, making the determinant equal to zero. Thus, one can test if the `det` function checks the existence of zero-columns before performing the standard calculation of the determinant.¹ Benchmark b_7 generates an N -size vector populated with random values and then sorts its elements by using the function `sort` (`gsort`) in MATLAB (Scilab), which performs the standard quicksort algorithm.

Benchmark b_8 illustrates a standard educational example in the field of signals and digital signal processing: the generation of an amplitude-modulated (AM) signal with the sinus carrier signal of frequency F_c , modulated (multiplied) by the sinusoidal signal of frequency F_m [33]. The Fast Fourier Transform (FFT) is applied to the obtained AM signal by using the original MATLAB and Scilab FFT

functions. After that, the benchmark draws the spectra of the individual signals and the modulated signal.

Finally, benchmark b_9 checks the efficiency of the operations with the external memory, which was in our case the standard hard disk (cf. §2.2). After randomly generating an $N \times N$ matrix, it saves the matrix to the disk and then reloads it into another matrix. The original save and load functions from both tools were used.

4.3.2 Source Code of an Exemplary Benchmark

Because the presentation of the source code for all our benchmarks would be too voluminous for this paper format, here we exemplify the source code of the $b_{6,0c}$ benchmark function. It is a bit more elaborate version of b_6 , with a few specifics. Its MATLAB and Scilab versions are shown in Fig. 1a and 1b. The similarity between the two codes is striking.

In lines number 1, `rndMtrxDet0c` benchmark function is declared. In lines number 2, the return parameters `D` and `T` are declared as the vectors of size `Nrp` and 2, respectively, and their elements are set to zero. In lines number 3, the specialized stopwatch timer function `tic` starts the measurement of the execution time. Both MATLAB and Scilab include the stopwatch time functions `tic` and `toc`, which measure the elapsed time with millisecond precision [34, 35]. In our case, they will measure the time of the `Nrp` repetitions of the same task, governed by the for-loops which start at lines number 4. The rationale for measuring the total (T_{tot}) and not the particular time of each repetition was to avoid the possible errors when the measured time is shorter than the not-so-great precision of the provided stopwatch functions.

The body of the for-loops, in lines 5 to 19, is the actual task routine of the benchmark. Each routine starts with the declaration and initialization of the single-row and single-column vectors, `zrow` and `zcol` (lines 5 and 6). After that, the benchmark generates the random $N \times N$ square matrix `A` (lines 7). Generally, the declarations in the array programming are demanding and important benchmark parts.

The code part from lines 8 to 10 is used for zeroing the randomly selected column of the matrix `A`, assuring that (at least) one such column exists in the matrix.

The lines from 11 to 19 contain the parts of the benchmarks that calculate the matrix-`A` determinant. First, there is a check if there is at least one zero-column or one zero-row in the matrix. Although in this benchmark we did not provide the insertion of (at least) one zero-row, we still provide the check for such a row for the sake of completeness. These checks are done in lines 11 to 16. The variable `is0` serves as a logic flag that is first set to 0 (false). An additional for loop is started that runs the column or row indices from 1 to `N`, and checks if some of the columns, `A(:, r)`, or rows, `A(r, :)`, are equal to the previously defined zero-valued `zcol` and `zrow`. If either of these is true, the `is0` flag is set to 1 (true) and the for-loop is exited by the `break` in

¹ For the matrices with two or more rows or columns being equal — and because of that having the zero determinant — we have encountered a precision problem worth noticing. Namely, in some cases, especially when such matrices are very large, both MATLAB and Scilab do not return the zero result but a value of the order of magnitude of 10^{-20} . This problem

probably occurs because of the accumulation of errors after many additions and subtractions of numbers with the absolute values differing less than the calculation precision. Although small, the occurrence of such results require additional checking and zeroing if they are lower than a certain threshold.

```

1 function [D, T] = rndMtrxDet0c(N, Nrp)
2   D = zeros(Nrp); T = zeros(2);
3   tic;
4   for j = 1 : Nrp
5     zrow = zeros(1, N);
6     zcol = zeros(N, 1);
7     A = rand(N, N);
8     c0 = round(N*rand());
9     if (c0 == 0) c0 = 1; end
10    A(:, c0) = 0;
11    is0 = 0;
12    for r = 1 : N
13      if (A(:, r) == zcol) is0 = 1; end
14      if (A(r, :) == zrow) is0 = 1; end
15      if (is0 == 1) break; end
16    end
17    if (is0 == 1) d = 0;
18    else d = det(A); end
19    D(j) = d;
20  end
21  Ttot = toc;
22  Tavg = Ttot/Nrp;
23  T(1) = Ttot;
24  T(2) = Tavg;
25 end

```

Figure 1a Source code of the benchmark $b_{6,0c}$ function in MATLAB

```

1 function [D, T] = rndMtrxDet0c(N, Nrp)
2   D = zeros(Nrp); T = zeros(2);
3   tic();
4   for j = 1 : Nrp
5     zrow = zeros(1, N);
6     zcol = zeros(N, 1);
7     A = rand(N, N);
8     c0 = round(N*rand());
9     if (c0 == 0) then c0 = 1; end
10    A(:, c0) = 0;
11    is0 = 0;
12    for r = 1 : N
13      if (A(:, r) == zcol) then is0 = 1; end
14      if (A(r, :) == zrow) then is0 = 1; end
15      if (is0 == 1) then break; end
16    end
17    if (is0 == 1) then d = 0
18    else d = det(A); end
19    D(j) = d;
20  end
21  Ttot = toc();
22  Tavg = Ttot/Nrp;
23  T(1) = Ttot;
24  T(2) = Tavg;
25 endfunction

```

Figure 1b Source code of the benchmark $b_{6,0c}$ function in Scilab

lines number 15. In this part of the benchmark, we combine vectorization (for the comparison of rows and columns) with the standard implementation of the iterative programming structure. A careful reader will note that by adding one more conditional break after line 13, an unnecessary checking for the zero-row could have been avoided if a zero-column had been found. If $is0 = 1$, the determinant d is set to zero (lines 17). In our case, matrix A will always have at least one zero-column, so this will always be the case. Still, the other possibility, in which the determinant is calculated by using the `det` function (lines 18), is also accounted for.

Lines 19 finish the benchmark routine by assigning the determinant value (here $d = 0$) found in the j -th iteration of the outer for-loop to the j -th element of vector D .

The last part of the function starts by assigning the value returned by the end of the time-measurement function `toc` to the variable T_{tot} (lines 21). After that, the average $\bar{T}_B$ time per one repetition of the benchmark routine is calculated and stored in the variable T_{avg} (cf. tables in sec. 5). Finally, both of these times are returned via vector T (lines 23 and 24).

The source code of the other benchmark functions developed in this work is available in [36].

4.3.3 Benchmarking Scripts

The benchmarking scripts for each of the two observed calculation tools include the following parts:

1. Definitions of all constants needed to run the benchmark routines and their capturing function:
 - the number N_{set} of the benchmark *measurement sets*, i.e. of the repetition of each benchmark function, in our case $N_{set} = 5$;

- the number N_{rp} of the benchmark repetitions in each measurement, i.e. in each benchmark function, which is most often $N_{rp} = 100$;
 - the problem size $(N, N \times M)$, which is taken as a value of an element from the vector Nps of size 5.
2. Procedure for executing the benchmark functions.
 3. Procedure for storing the obtained results in a text file, suitable for additional analysis in spreadsheets or other types of software.

Fig. 2 depicts a piece of Scilab script source code for measuring the $b_{6,0c}$ benchmark performance. It results with $N_{set} = 5$ benchmark measurement sets, one for each predefined problem size, $Nps(i)$, in the Nps vector. As explained in the previous sub-subsection (§4.3.2), each benchmark function executes the benchmark routine N_{rp} times and returns the total and average execution times. The shown script outputs these results to the standard output (the Scilab window). To store the results to a text file, instead of `mprintf`, one should use the `mfprintf` function.

```

1 Nset = 5;
2 Nrp = 100;
3 Nps = [100, 500, 1000, 2000, 4000];
4 mprintf('Repetitions=%d\n', Nrp);
5 mprintf('Function(size)\tTotal\tAvg\n');
6 for i = 1 : Nset
7   [r,t] = rndMtrxDet0c(Nps(i), Nrp);
8   mprintf('rndMtrxDet0c(%d)\t%f\t%f\n', Nps(i), t(1), t(2));
9 end

```

Figure 2 Part of Scilab script, which calls the benchmark function `rndMtrxDet0c` ($b_{6,0c}$) and displays the obtained results for N_{set} benchmark measurement sets with the problem size $Nps(i)$, $i = 1, 2, \dots, N_{set}$.

5 MATLAB AND SCILAB BENCHMARK PERFORMANCE

In this section, we present the results of the benchmark execution times measurements performed in MATLAB and Scilab and discuss them.

In modern, multitasking operation systems, there are a lot of background processes that cannot be controlled by the user and whose running may interfere with the execution of the benchmark functions. Indeed, when analyzing the obtained results, there were sporadic cases of enlargements of the execution times of some benchmarks. To exclude that kind of result deviation, we were selecting the three best out of the five performed time measurements and took their mean value as the indicator of the benchmark performance. The benchmarks described in the previous section were grouped according to the division in §4.2, and the results of their benchmark execution times are presented in the following subsections.

5.1 Basic Calculation Benchmarks ($b_1 - b_3$)

Tab. 3 presents the average execution times of the basic vector and matrix calculations in benchmarks b_1 and b_2 , and slightly more complex operations in b_3 (cf. Tab. 2). For these benchmarks, there are versions with and without the use of vectorization. Additionally, for b_3 , the execution times for the k -th matrix power are given with $k = 3$ and $k = 11$.

By observing the obtained results, we can easily note that the execution times for both of the calculation tools are proportional to the total problem size, i.e. to the number of elements in the vectors or matrices. Next, if we focus on comparing the benchmark versions with and without vectorization, we can note that—for benchmark b_1 —MATLAB gives surprisingly better results for the version without vectorization! The respective performance, for the problem size $N = 1 \times 10^6$ ($N = 4 \times 10^6$), is even 7.95 (8.11) times better than for the corresponding vectorized versions. It must be that the simplicity of this benchmark operation performed on the number series in a vector, combined with the effective MATLAB JIT compiler, results in a very efficient code that largely surpasses vectorization. Anyhow, this peculiarity does deserve additional investigation.

For the remaining two benchmarks, MATLAB versions with vectorization outperform those without it, but still not by much. For b_2 with $N^2 = 1 \times 10^6$ ($N^2 = 4 \times 10^6$), the vectorized benchmarks are 49.3% (82.2%) faster than the non-vectorized. However, for b_3 , this superiority is severely diminished. With $k = 3$ and $N^2 = 1 \times 10^6$ ($N^2 = 4 \times 10^6$), vectorization improves performance for only 3.1% (2.0%). For $k = 11$ and the same problem size(s), the improvement is only 1.4% (1.1%). Again, we may attribute this to the effective JIT compilation.

On the other hand, in Scilab, the benchmarks with vectorization consistently and hugely outperform those without it. The cause of that is the strictly interpreted programming code in Scilab, without any sort of compilation, which leads to the very bad performance of the code with programmed loops (§4.2). Thus, the vectorized versions of

Table 3 Average $\bar{T}_B$ execution times of the basic calculation benchmarks, $b_1 - b_3$. The averages are calculated from the 3 best out of the total of 5 sets of measurements, each with $N_{rp} = 100$ repetitions of every benchmark.

Benchmark		Problem size	$\bar{T}_B$ /ms		$\bar{T}_{B,Sci} / \bar{T}_{B,MAT}$
Abbr.	Variant	N or $N \times M$	MATLAB	Scilab	
b_1	Vectoriz.	1×10^6	12.25	22.46	1.83
		4×10^6	49.15	92.35	1.88
	No vectoriz. (for-loop)	1×10^6	1.54	1239.72	805.01
		4×10^6	6.06	4956.01	818.82
b_2	Vectoriz.	$1 \times 10^3 \times 1 \times 10^3$	7.93	25.04	3.16
		$2 \times 10^3 \times 2 \times 10^3$	31.18	101.00	3.24
	No vectoriz. (for-loop)	$1 \times 10^3 \times 1 \times 10^3$	11.84	1910.73	161.38
		$2 \times 10^3 \times 2 \times 10^3$	56.81	7668.45	134.98
b_3	Vectoriz. ($k = 3$)	$1 \times 10^3 \times 1 \times 10^3$	109.64	191.86	1.75
		$2 \times 10^3 \times 2 \times 10^3$	748.54	1185.42	1.58
	Vectoriz. ($k = 11$)	$1 \times 10^3 \times 1 \times 10^3$	263.50	325.96	1.24
		$2 \times 10^3 \times 2 \times 10^3$	1766.77	2151.14	1.22
	No vectoriz. ($k = 3$)	$1 \times 10^3 \times 1 \times 10^3$	113.06	2180.06	19.28
		$2 \times 10^3 \times 2 \times 10^3$	763.21	8789.01	11.52
	No vectoriz. ($k = 11$)	$1 \times 10^3 \times 1 \times 10^3$	267.30	2263.58	8.47
		$2 \times 10^3 \times 2 \times 10^3$	1785.77	9736.38	5.45
$\Sigma(\text{vectr. only}):$		20×10^6	2988.96	4095.23	1.37
$\Sigma(\text{non-vectr. only}):$		20×10^6	3005.59	38743.94	12.89
$\Sigma(\text{all}):$		40×10^6	5994.55	42839.17	7.15

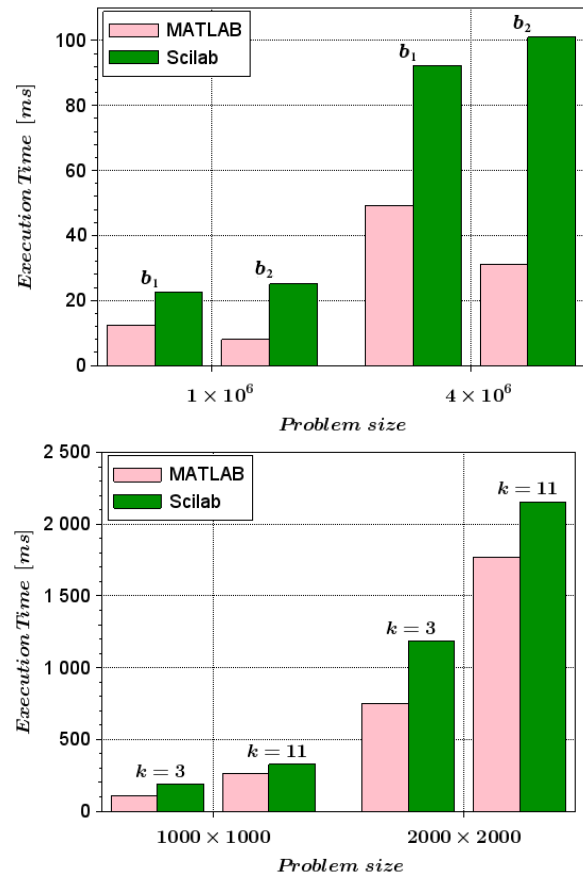


Figure 3 Graphical presentation of the benchmark execution times in MATLAB and Scilab: b_1 and b_2 (up) and b_3 (down). Shorter is better.

the benchmarks with $N^2 = 1 \times 10^6$ ($N^2 = 4 \times 10^6$) are faster than the non-vectorized ones as follows: for b_1 , 55.2 (53.7) times; for b_2 , 76.3 (75.9) times; for b_3 , with $k = 3$, 11.4 (7.41) times, and with $k = 11$, 6.9 (4.5) times. It turns out the efficiency of vectorization highly depends on the character of the benchmark calculations.

As for the comparison of the MATLAB and Scilab benchmark performances, the last column in Tab. 3 shows that MATLAB is better in all of them and—as discussed in the previous paragraph—greatly superior in all cases without vectorization. Furthermore, it is interesting to note that Scilab vectorized benchmark versions perform worse than the corresponding MATLAB vectorized or non-vectorized versions of the same problem size.

The last three rows in Tab. 3 show the cumulative execution times of all the benchmarks and their ratio in the last column. Overall, MATLAB is faster than Scilab by more than 7 times. Even more, it is faster by almost 13 times when observing the non-vectorized benchmark routines only.

On the other hand, for the benchmark routines with vectorization—which is, after all, the recommended form of programming code in this sort of languages—Scilab shows very good performance. It is slower than MATLAB from only 1.22 for b_3 up to 3.24 for b_2 . Overall, it is 37% slower than MATLAB, which can be considered an excellent performance. Fig. 3 shows these relations graphically.

5.2 Matrix Calculation and Sorting Benchmarks ($b_4 - b_7$)

Tab. 4 shows the performance of the benchmarks b_4 to b_7 , all with vectorized versions only. An interesting thing to note is that the execution times are only nearly linearly proportional to the problem size, and not dependent only on it. In b_4 , the four total problem sizes are 1×10^6 , 2×10^6 , 2×10^6 , and 4×10^6 , and the corresponding execution times in MATLAB (Scilab) expressed relative to the first problem size go as: 1.00 (1.00), 1.91 (1.91), 2.88 (2.73), and 5.49 (5.11). The matrices in the second and third cases have the same number of elements, but in the latter case, there is a slowdown of 51% for MATLAB and 43% for Scilab. The possible cause of this is the double number of rows ($N = 2M$) in the matrix in the third case (the first of the two matrices in the multiplication), which are half as long as the rows in the second case. Recalling the matrix multiplication rules, this means that there are twice as many element-multiplication "chains" that are half as long as in the previous case, possibly causing the less efficient multiplication. Still, the proportion of this slowdown is quite surprising.

Comparing the performance of the two tools, MATLAB is again faster in all proposed benchmarks, but Scilab performs very well, especially in benchmarks b_6 , $b_{5,b}$, and b_4 (Fig. 4). In them, Scilab lags by a factor of only 1.14–1.29. It performs much worse in the quicksort (b_7), especially when the numbers are not rounded to integers (Fig. 5).

Both tools solve the matrix equation (b_5) faster by using the back-divide algorithm ($b_{5,b}$) than by their internal function ($b_{5,i}$), suggesting its possible improvement (Fig. 4).

Table 4 Average $\bar{T}_B$ execution times of the benchmarks $b_4 - b_7$ (cf. Tab. 3)

Benchmark		Problem size	$\bar{T}_B$ /ms		$\frac{\bar{T}_{B,Sci}}{\bar{T}_{B,MAT}}$
Abbr.	Variant	$N \times M$ or N	MATLAB	Scilab	
b_4	Vectoriz.	$1 \times 10^3 \times 1 \times 10^3$	94.06	121.42	1.29
		$1 \times 10^3 \times 2 \times 10^3$	179.53	232.20	1.29
		$2 \times 10^3 \times 1 \times 10^3$	270.68	331.25	1.22
		$2 \times 10^3 \times 2 \times 10^3$	516.78	620.04	1.20
$b_{5,i}$	Vectoriz., invs. mtrx.	$1 \times 10^3 \times 1 \times 10^3$	106.68	193.50	1.81
		$2 \times 10^3 \times 2 \times 10^3$	580.96	1390.01	2.39
$b_{5,b}$	Vectoriz., back-divide	$1 \times 10^3 \times 1 \times 10^3$	66.31	81.81	1.23
		$2 \times 10^3 \times 2 \times 10^3$	307.17	373.84	1.22
b_6	Vectoriz.	$1 \times 10^3 \times 1 \times 10^3$	55.45	63.42	1.14
		$2 \times 10^3 \times 2 \times 10^3$	267.98	316.94	1.18
$b_{6,0c}$	Vectoriz.	$1 \times 10^3 \times 1 \times 10^3$	23.43	48.08	2.05
		$2 \times 10^3 \times 2 \times 10^3$	96.70	185.37	1.92
b_7	(Quicksort)	1×10^6	101.61	533.04	5.25
		2×10^6	212.34	1063.18	5.01
	With.round.	2×10^6	173.94	601.06	3.46
$\Sigma(all, \text{ except } b_7)$		20×10^6	2565.73	3957.88	1.54
$\Sigma(all)$		40×10^6	3053.62	6155.16	2.02

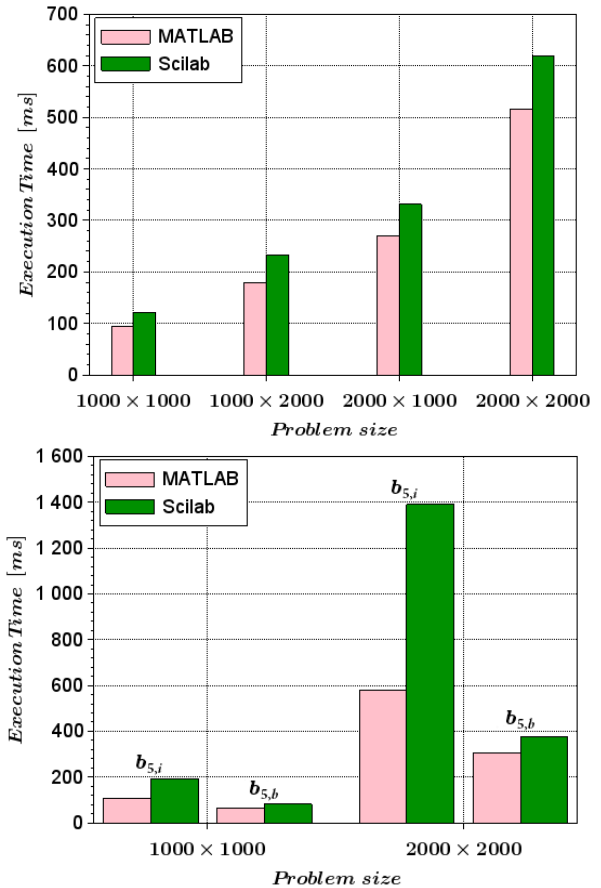


Figure 4 Performance of the benchmarks b_4 (up) and b_5 (down)

Overall, for these benchmarks, Scilab is two (2) times slower, but if the sorting is excluded, it is only about one and a half times (1.5) slower than MATLAB. For the relatively short absolute times, this is again a very good performance.

5.3 Signal Generation and DSP, Disk File Manipulation (b_8, b_9) and the Benchmarks Overall

Tab. 5 shows the performance of the benchmark b_8 . The changes of the parameters influence the execution times rather unexpectedly. Furthermore, the enlargement of N_{rp} for ten times should not influence $\bar{T}_B$. Of course, this illustrates how too small N_{rp} values may result in too short times, which are below the time measurement accuracy, and thus produce wrong measurements (grayed out in the table). In other words, we ought to observe the times obtained for $N_{rp} = 50$ only. Here, MATLAB is superior again, but Scilab performs satisfactory, especially if one considers that the whole job was done in a few milliseconds.

The performance of our last benchmarks, for the disk file manipulation (b_9), is visible in Tab. 6 (cf. also §2.2 and Tab. 2). To reduce the number of disk operations, the number of repetitions was decreased to 10. Surprisingly, in these operations, Scilab is faster than MATLAB! The only and very slight exception is for the load of 100×100 matrices. Moreover, this advantage is better for larger files, where the Scilab save and load operations are roughly 6 up to 15 times faster than those of MATLAB.

Tab 7 contains the cumulative benchmark execution times for the vectorized benchmarks from Tab. 3, 4, and 5 & 6. To reduce the influence of the disk operations, the sums for b_8 and b_9 were multiplied with the weight factor of 0.1. By that, Scilab is overall only 57% slower than MATLAB, and without that, it would be equal (faster by 0.005%)!

6 CONCLUSION

In this research, the primary goal was to compare the performance of the two well-known specialized computing software environments, MATLAB and Scilab, on a set of purposefully created benchmarks that resemble the often-used vector and matrix calculations in the professional, scientific and educational domain (sec. 4).

The benchmark execution times, presented in sec. 5, generally show the supremacy of MATLAB that was known from the previous works of others (cf. §4.1). However, after excluding the bad Scilab performance in the non-vectorized benchmarks due to its interpreted-only nature, we showed that this popular freeware can compete well with its highly-commercial rival in several tasks. For the benchmarks b_3 (with $k = 11$), b_4 , $b_{5,b}$, and b_6 (the basic version), it lags behind MATLAB for as little as 10%–30%; for the benchmarks b_1 , b_3 (with $k = 3$), $b_{5,i}$, and $b_{6,0c}$ it is slower for a factor from approximately 1.5 to 2.4. The cumulative results in Tab 7—our version of bench function—show this clearly, and qualify Scilab as a serious choice for several needs.

Even for the benchmarks in which Scilab performed significantly worse than MATLAB, like those in b_2 (the k -th matrix power) and b_7 (quicksort), the fact that it was slower than MATLAB for the factor of approximately 3 up to 5, should be considered in the context of the overall very short absolute times of these calculations. The benchmarks in b_2 perform calculations on 1×10^6 (4×10^6) array elements in around 25 ms (100 ms), and the benchmarks in b_7

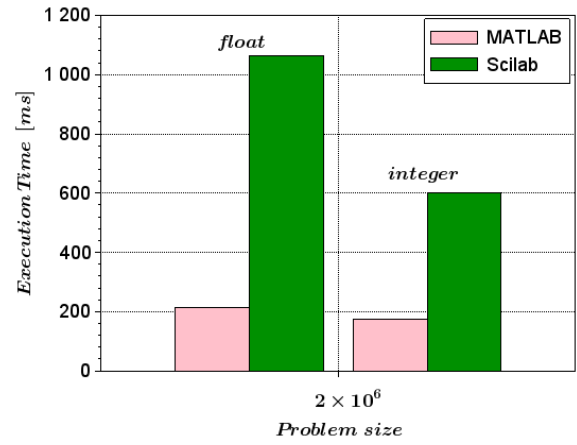


Figure 5 Performance of the benchmark b_7 (quicksort)

Table 5 Average $\bar{T}_B$ execution times of the benchmark b_8 , for the generation and FFT analysis of the AM signal with different F_c and F_m parameters and the variable N_{rp} parameter. The rest of the markings are as in Tab. 3.

Benchmark		(Probl. size)	$\bar{T}_B$ /ms		$\bar{T}_{B,Sci}$ / $\bar{T}_{B,MAT}$
Abbr.	Par. (F_c, F_m)/Hz	N_{rp}	MATLAB	Scilab	
b_8	1000, 100	5	0.12	1.43	11.92
		50	0.75	2.53	3.37
	4000, 100	5	0.20	0.59	2.95
		50	2.56	7.48	2.92
Σ ($N_{rp} = 50$ only):			3.31	10.01	3.02

Table 6 Average $\bar{T}_B$ execution times of the benchmark b_9 , for the file-handling save and load disk operations, with $N_{rp} = 10$.

Benchmark		Problem size	$\bar{T}_B$ /ms		$\frac{\bar{T}_{B,Sci}}{\bar{T}_{B,MAT}}$
Abbr.	Operation	$N \times M$	MATLAB	Scilab	
b_9	Save	$1 \times 10^2 \times 1 \times 10^2$	28.55	20.72	0.73
	Load	- -	1.80	1.85	1.03
	Save	$1 \times 10^3 \times 1 \times 10^3$	387.80	24.18	0.06
	Load	- -	99.60	14.56	0.15
	Save	$3 \times 10^3 \times 3 \times 10^3$	3406.63	445.87	0.13
	Load	- -	888.86	91.02	0.10
Σ (all):		20.02×10^6	4813.24	598.20	0.12

Table 7 Cumulative execution times for benchmarks with vectorization

Benchmrks. (with vectrz.)	Wght. w_{b_i}	Tot. probl. size $\Sigma(N, N \times M)$	$\Sigma \bar{T}_B$ /ms		$\bar{T}_{B,Sci} / \bar{T}_{B,MAT}$
			MATLAB	Scilab	
$b_1 - b_3$	1.0	20.0×10^6	2988.96	4095.23	1.37
$b_4 - b_7$	1.0	34.0×10^6	3053.62	6155.16	2.02
b_8, b_9	0.1	2.0×10^6	481.66	60.82	0.13
Σ :			6524.24	10311.21	1.58

sort 1×10^6 (2×10^6) elements in 0.53 s (1.1 s, i.e. 0.6 s for the numbers rounded to integers). Obviously, such performance, though worse than that of MATLAB, will satisfy not only educational but also many demanding professional and scientific requirements.

Of the few surprises from this research, the first one concerns MATLAB: benchmark b_1 with vectorization performed significantly slower than the one without it, as

commented in §5.1. The second is the rather large benchmark b_4 slowdown ($\approx 50\%$) in both tools when the first, non-square matrix in multiplication has the same number of elements but twice as many rows as columns (§5.2). Lastly, the third is that Scilab outperforms MATLAB in the disk file operations several times, especially for the (very) large files (§5.3). These, and a few other observations deserve additional investigation. Furthermore, the inclusion of multiprocessing and distributed computing in these software environments present challenging topics for future work.

Finally, what can we suggest to the reader? The winner of this race is obvious and could have been predicted even before this research (cf. §4.1). If the ultimate speed and the compatibility of the programming code with the industry-standard language are a must, and if the finances for the initial and yearly renewal of the licenses are not an issue, MATLAB is a sure pick. However, if the developers are either writing their code from scratch or are willing to adapt it by performing straightforward changes, this paper shows that Scilab is an excellent calculation environment that will in most cases perform either almost as well as MATLAB or rather close. It will even manipulate the disk files faster than its big rival, and — of course — it will cost you nothing!

7 REFERENCES

- [1] MathWorks Inc., MATLAB: <https://www.mathworks.com/products/matlab.html>
- [2] Scilab Home Page, <https://www.scilab.org>
- [3] GNU Octave, <https://www.gnu.org/software/octave/index>
- [4] SageMath, <https://www.sagemath.org>
- [5] FreeMat, <http://freemat.sourceforge.net>
- [6] Maxima, <https://maxima.sourceforge.io>
- [7] Almeida, E. S., Medeiros, A. C., & Frery, A. C. (2012). How good are MATLAB, Octave and Scilab for computational modelling? *Computational & Applied Mathematics*, 31(3), 523-538. <https://doi.org/10.1590/S1807-03022012000300005>
- [8] MathWorks, Pricing and Licensing — MATLAB & Simulink, <https://www.mathworks.com/pricing-licensing.html>
- [9] Scilab Open Source, <https://www.scilab.org/about/scilab-open-source-software>
- [10] Scilab History, <https://www.scilab.org/about/company/history>
- [11] ESI Group, <https://www.esi-group.com>
- [12] Scilab, Use cases, <https://www.scilab.org/use-cases>
- [13] MathWorks, MATLAB Execution Engine, <https://www.mathworks.com/products/matlab/matlab-execution-engine.html>
- [14] MathWorks, Techniques to Improve Performance, https://www.mathworks.com/help/matlab/matlab_prog/techniques-for-improving-performance.html
- [15] Scilab Wiki, Introduction, <https://wiki.scilab.org/Introduction>
- [16] Scilab Help, About M2SCI tools, https://help.scilab.org/docs/5.3.3/en_US/About_M2SCI_tools.html
- [17] Scilab Help, MATLAB to Scilab Conversion Tips, https://help.scilab.org/docs/5.3.3/en_US/section_c592a9ecd0ed2b4d08f8a4de718ee9aa.html
- [18] MathWorks Help, bench – MATLAB benchmark, <https://mathworks.com/help/matlab/ref/bench.html>
- [19] Scilab Help, bench_run, https://help.scilab.org/docs/6.1.1/en_US/bench_run.html
- [20] Steinhilber, S. (2008). Comparison of Mathematical Programs for Data Analysis, Munich. Retr.: <https://www.additive-net.de/images/software/wolfram/publicon/downloads/numbercrunch5.pdf>
- [21] Pinçon, B. (2022). Quelques tests de rapidité entre différents logiciels matriciels. University of Lorraine, Retrieved from <https://cermics.enpc.fr/~jpc/scilab-gtk-tiddly/bench.pdf>
- [22] Baudin, R. (2016). Run time comparison of MATLAB, Scilab and GNU Octave on various benchmark programs, Retrieved from <http://roland65.free.fr/benchmarks/benchmarks-0.2.pdf>
- [23] Sharma, N. & Gobbert, M. K. (2010). A comparative evaluation of MATLAB, Octave, FreeMat, and Scilab for Research and Teaching. Department of Mathematics and Statistics, University of Maryland, Retr. from <https://userpages.umbc.edu/~gobbert/papers/SharmaGobbertTR2010.pdf>
- [24] Coman E., Brewster, M. W., Popuri, S. K., Raim, A. M., & Gobbert, M. K. (2012). A Comparative Evaluation of Matlab, Octave, FreeMat, Scilab and R on Tara, Retrieved from <http://profs.scienze.univr.it/~caliari/pdf/octave.pdf>
- [25] Shaukat, K., Tahir, F., Iqbal, U., & Ajmad S. (2018), A Comparative Study of Numerical Analysis Packages. *International Journal of Computer Theory and Engineering*, 10 (3), 67-72. <https://doi.org/10.7763/IJCTE.2018.V10.1201>
- [26] Leros, A., Andreatos, A., & Zagoranos A. (2010). Matlab — Octave science and engineering benchmarking and comparison. Proc. of the 14th WSEAS inter. conf. on Computers (II) 746-754. <https://dl.acm.org/doi/10.5555/1984366.1984421>
- [27] MathWorks (1984-2022). *MATLAB Programming Fundamentals (R2022a)*. Natic, MA: MathWorks Inc. Retr. from: https://www.mathworks.com/help/pdf_doc/matlab/matlab_prog.pdf
- [28] Wkp. article: Array Programming; <https://en.wikipedia.org>
- [29] Intel, Vectorization Basics for Intel® Architecture Processors.
- [30] Mikac, M., Horvatić, M., & Mikac, V. (2020). Using vectorized calculations in Scilab to improve performances of interpreted environment. *INTED2020 Proceedings – the 14th International Technology, Education and Develop. Conf.*, Valencia, Spain, 2127-2136. <https://doi.org/10.21125/inted.2020.0664>
- [31] MathWorks Help, mldivide, \, MATLAB, <https://www.mathworks.com/help/matlab/ref/mldivide.html>
- [32] Scilab Help, backslash, https://help.scilab.org/docs/6.0.2/en_US/backslash.html
- [33] Mikac, M. & Horvatić, M. (2020). Using Open-Source Numerical Computation Software in Education — Basic Performance Comparison and Lab Examples. *EDULEARN20 Proc. – The 12th International Conf. on Education and New Learning Technologies*, 2319-2327. <https://doi.org/10.21125/edulearn.2020.0714>
- [34] MathWorks Help, Start stopwatch time, MATLAB, <https://www.mathworks.com/help/matlab/ref/tic.html>
- [35] Scilab Help, tic – Start a stopwatch time, https://help.scilab.org/doc/6.0.2/en_US/tic.html
- [36] Mikac, M., Logožar, R., & Horvatić M. (2022). Scilab vs. MATLAB, The benchmark functions source-code repository: <https://papers.unin.com.hr/matscilab2022.html>

Authors' contacts:

Matija Mikac, M.Sc.E.E.

(Corresponding author)

University North,
Jurja Križanića 31b, 42000 Varaždin, Croatia
matija.mikac@unin.hr

Robert Logožar, Ph.D. C.S.

University North,
Jurja Križanića 31b, 42000 Varaždin, Croatia
robert.logozar@unin.hr

Miroslav Horvatić, M.E.E.

University North,
Jurja Križanića 31b, 42000 Varaždin, Croatia
miroslav.horvatic@unin.hr

Free Vibration of a Cantilever Euler-Bernoulli Beam Carrying a Point Mass by Using SEM

Saida Hamioud

Abstract: The objective of this research is to study the free vibration of a cantilever Euler-Bernoulli beam carrying a point mass with moment of inertia at the free end using the spectral element method (SEM). Typically, the shape (or interpolation) functions used in the Spectral element method are derived from exact solutions of the governing differential equations of motion in the frequency domain. The beam was discretized by a single spectral element which was connected by a point mass at the free end. The dynamic stiffness matrix of the beam is formulated in frequency domain by considering compatibility conditions at the additional mass position. Then, the first three natural frequencies of the cantilever beam are determined. After the validation of the spectral element method, the influence of the non-dimensional mass parameter and the non-dimensional mass moment of inertia on the first three natural frequencies and shape mode are examined.

Keywords: cantilever beam; free vibration; point mass; spectral element method

1 INTRODUCTION

Beams structures are used as one of the fundamental structural components in civil and mechanical engineering. Free vibration analysis of beams is of great importance in design and fabrication of structures and machines such as highway bridges, railroads, tall building, wind turbines and huge cranes.

There have been extensive research published in the area about the vibration characteristics of a uniform beam with different boundary conditions and with/without various concentrated elements (such as intermediate point masses, rotary inertias, linear springs, rotational springs, spring-mass systems, etc.). Analytical and numerical methods were used to obtain the natural frequencies and different aspects have been considered [1-8]. For the vibration analysis of beams carrying concentrated masses at arbitrary locations, a lot of studies have been published. Hong et.al [9] investigated the transverse vibration of clamped-pinned-free Euler-Bernoulli beam with mass at free end; obtained analytical eigenvalues of the system were compared to experimental data. Natural frequencies and model shapes of a clumped beam with mass at free end have been determined [10]. Wang et al. [11] examined the transverse vibration of a cantilever Euler-Bernoulli beam that has a mass with moment of inertia on its free end. Wu and Lin [12] analysed the frequency equation of flexural vibrating cantilever beam with masses attached at multiple points by using an analytical-numerical combined method. They derived the eigenvalue equation analytically by using an expansion theorem and frequencies and mode shapes were calculated numerically. Chang [13] performed the free vibration analysis of a simply supported beam carrying a rigid mass at the middle. Low [14] used both the methods of frequency determinant and the method of Laplace transform to determine the Eigenvalues of a beam with any number of point masses. Naguleswaran [15, 16] studied free vibration of an Euler-Bernoulli beam with point masses and negligible inertia moment of the point mass. Gürgöze [17-19] has carried out several studies on the frequency equation of flexural vibrating beam carrying a rigid mass. Reference [20] dealt with the determination of the eigenvalues of an Euler-

Bernoulli beam with one end spring-mass system attached and the other fixed.

The purpose of this study is to use the spectral element method to establish the dynamic equation of a cantilever Euler-Bernoulli beam with tip mass at the free end. The spectral formulation requires that the equation of motion is solved in the frequency domain and the fast Fourier transform (FFT) is utilized to convert the time domain responses to the wave domain and back. The accuracy and the validation of the proposed approach are confirmed by the comparison with existing results in literature. Finally, the effects of non-dimensional mass parameter and the non-dimensional mass moment of inertia on the natural frequencies and shape mode are investigated.

2 MATHEMATICAL MODEL AND FORMULATIONS

Consider a cantilever beam with an additional mass M_0 having a moment of inertia J_0 at the free end of length L as shown in Fig.1. The beam has flexural stiffness EI and mass per unit length M . Based on the Euler-Bernoulli beam theory, the governing equation of motion can be written:



Figure 1 Schematic of a cantilever beam with an additional mass at free end

$$\frac{\partial^4 w(x,t)}{\partial x^4} + \frac{M}{EI} \frac{\partial^2 w(x,t)}{\partial t^2} = 0, \quad (1)$$

The solution of Eq. (1) can be assumed in the spectral form as:

$$w(x, t) = \frac{1}{N} \sum_{n=0}^{N-1} W_n(x) \cdot e^{i\omega_n t}, \quad (2)$$

Where $w(x, t)$ is the transverse displacement, E is Young's modulus, I is the area moment of inertia about the neutral axis.

Inserting Eq. (2) into Eq. (1) and after rearrangement, the following equation is derived:

$$\frac{d^4 W(x)}{dx^4} - \alpha^4 W(x) = 0, \quad (3)$$

In which $\alpha^4 = \frac{M\omega^2}{EI}$, and ω is the circular frequency.

Then the solution of Eq. (3) can be written as:

$$W(x, \omega) = \bar{A}_1 \cos \alpha x + \bar{A}_2 \sin \alpha x + \bar{A}_3 \cosh \alpha x + \bar{A}_4 \sinh \alpha x, \quad (4)$$

Writing displacements and rotations at nodes:

$$w_1 = w(0, t), \quad \vartheta_1 = \frac{\partial w(0, t)}{\partial x}, \quad w_2 = w(L, t), \quad \vartheta_2 = \frac{\partial w(L, t)}{\partial x}.$$

The nodal displacement and slope at both ends can be expressed as:

$$\{q_e\} = [D] e^{i\omega t} \{\bar{A}\}. \quad (5)$$

Where $\{\bar{A}\} = \{\bar{A}_1 \quad \bar{A}_2 \quad \bar{A}_3 \quad \bar{A}_4\}^T$ is the constant

Vector and the matrix $[D]$ has the form

$$[D] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \alpha & 0 & \alpha \\ \cos \alpha L & \sin \alpha L & \cosh \alpha L & \sinh \alpha L \\ -\alpha \cdot \sin \alpha L & \alpha \cdot \cos \alpha L & \alpha \cdot \sinh \alpha L & \alpha \cdot \cosh \alpha L \end{bmatrix}. \quad (6)$$

Corresponding to the general problem Fig. 1, the vector of forces at the ends of the beam can be established by using boundary conditions of the structural system, which are:

$$\text{At } x = 0 \begin{cases} V_1 = EI \frac{\partial^3 w(0, t)}{\partial x^3} \\ M_1 = -EI \frac{\partial^2 w(0, t)}{\partial x^2} \end{cases}, \quad (7)$$

$$\text{At } x = L \begin{cases} V_2 = EI \frac{\partial^3 w(L, t)}{\partial x^3} + M_0 \omega^2 \cdot w(L, t) \\ M_2 = EI \frac{\partial^2 w(L, t)}{\partial x^2} + J_0 \omega^2 \frac{\partial w(L, t)}{\partial x} \end{cases}, \quad (8)$$

In matrix form, the forces at the ends of beam are

$$\{F_e\} = \begin{Bmatrix} V_1 \\ M_1 \\ V_2 \\ M_2 \end{Bmatrix} = [F] \cdot e^{i\omega t} \{\bar{A}\}. \quad (9)$$

With,

$$[F] = \begin{bmatrix} 0 & -EI\alpha^3 & 0 & EI\alpha^3 \\ EI\alpha^2 & 0 & -EI\alpha^2 & 0 \\ C_1 & C_2 & C_3 & C_4 \\ R_1 & R_2 & R_3 & R_4 \end{bmatrix}. \quad (10)$$

$$C_1 = EI\alpha^3 \sin \alpha L + M_0 \omega^2 \cos \alpha L,$$

$$C_2 = -EI\alpha^3 \cos \alpha L + M_0 \omega^2 \sin \alpha L,$$

$$C_3 = EI\alpha^3 \sinh \alpha L + M_0 \omega^2 \cosh \alpha L,$$

$$C_4 = EI\alpha^3 \cosh \alpha L + M_0 \omega^2 \sinh \alpha L,$$

$$R_1 = -EI\alpha^2 \cos \alpha L - J_0 \omega^2 \sin \alpha L,$$

$$R_2 = -EI\alpha^2 \sin \alpha L + J_0 \omega^2 \cos \alpha L,$$

$$R_3 = EI\alpha^2 \cosh \alpha L + J_0 \omega^2 \sinh \alpha L,$$

$$R_4 = EI\alpha^2 \sinh \alpha L + J_0 \omega^2 \cosh \alpha L.$$

The relationship between nodal force and degree of freedom vectors is expressed by

$$\{F_e\} = [F][D]^{-1} \{q_e\}. \quad (11)$$

Where $[F][D]^{-1}$ is the spectral stiffness matrix of Euler-Bernoulli beam with mass M attached at end ($x = L$).

3 NUMERIC APPLICATIONS

The applications reported in this section are provided by applying the proposed approach on a cantilever beam with a mass attached at the free end with the following non-dimensional parameters; β_M , β_J , and λ denoting the non-dimensional mass parameter, non-dimensional rotary mass moment of inertia and non-dimensional natural frequencies of the beam, respectively.

$$\beta_M = \frac{M_0}{M \cdot L} \quad (12)$$

$$\beta_J = \frac{J_0}{M \cdot L^3} \quad (13)$$

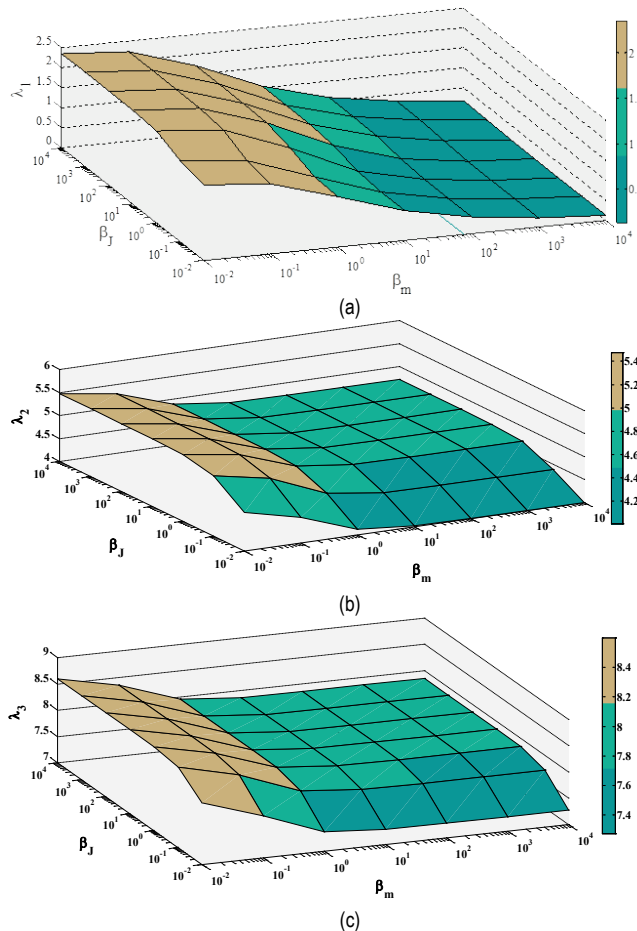
$$\lambda = \alpha L \quad (14)$$

To check the numerical model, the first three eigenvalues for the cantilever beam are determined by finding the nontrivial solutions of the determinant in Eq. (9). Computer programs based on SEM have been developed in MATLAB software to calculate numerical results.

In order to check the effectiveness of the proposed method, the first three dimensionless frequencies of a cantilever beam carrying a tip mass beam are shown in Tab. 1. Without considering the moment of inertia of the attached mass, it is clear from Tab. 1 that the obtained values are in good agreement with those obtained by Rao [1] and Gürgöze [16].

Table 1 The effect of the variation of β_M and β_J on the first three dimensionless frequencies λ of cantilever beam carrying a tip mass is presented in Fig. 2

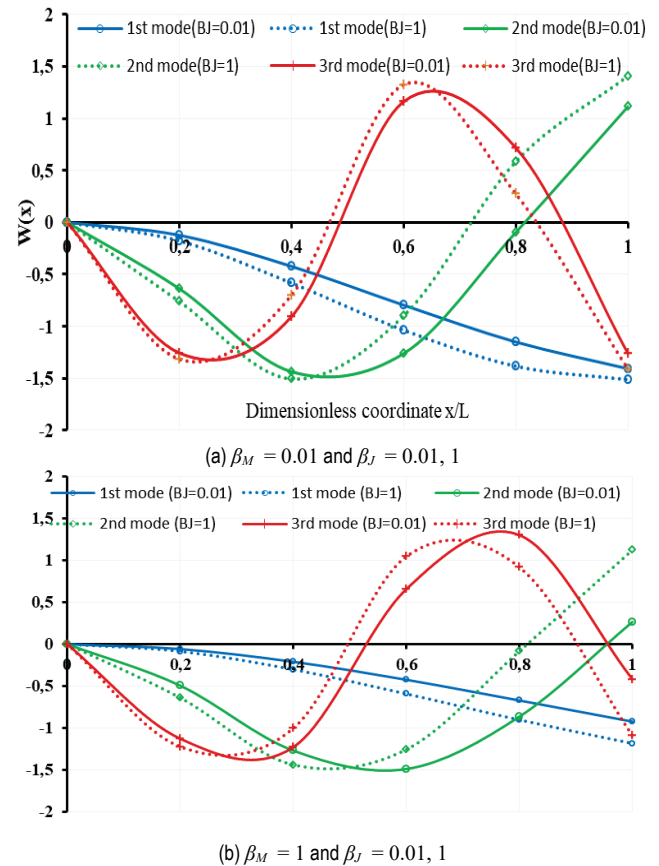
β_M, β_J	Method	λ_1	λ_2	λ_3
$\beta_M = 0.01 \beta_J = 0$	SEM	1.857	4.650	7.783
	Rao [1]	1.852	4.650	-
	Gürgöze [16]	1.857	-	-
$\beta_M = 1 \beta_J = 0$	SEM	1.248	4.031	7.134
	Rao [1]	1.248	4.031	-
	Gürgöze [16]	1.248	-	-
$\beta_M = 100 \beta_J = 0$	SEM	0.416	3.928	7.069
	Rao [1]	0.416	3.928	-
	Gürgöze [16]	0.416	-	-
$\beta_M = 0.01 \beta_J = 0.01$	SEM	1.875	4.843	8.168
$\beta_M = 0.01 \beta_J = 0$	SEM	2.286	5.454	8.591
$\beta_M = 1 \beta_J = 0.01$	SEM	1.252	4.128	7.367
$\beta_M = 1 \beta_J = 1$	SEM	1.540	4.854	7.949

**Figure 2** Variation of the natural frequency parameter (λ_1) as β_M and β_J varies from 10^{-2} to 10^4 . (a) λ_1 , (b) λ_2 , (c) λ_3 .

It observed from Fig. 2a that surface has a region of lowest dimensionless frequencies λ_1 corresponds to both β_M and β_J ($\beta_M \geq 100, \beta_J \geq 100$) being very extremely large values which represents the clumped-simply supported system. For negligible effect of both β_M and β_J ($\beta_M \leq 0.01, \beta_J \leq 0.01$), the values of dimensionless frequencies λ_1 represents the clumped-free beam. The Fig. 2 shows that the vibration dimensionless frequencies are significantly influenced by the increases in the non-dimensional mass parameter and non-dimensional rotary mass moment of inertia. The dimensionless frequencies will decrease by any increase of

non-dimensional mass parameter (β_M) and also increase by any increase of non-dimensional rotary mass moment of inertia (β_J).

Moreover, the first three transverse mode shapes of the cantilever beam carrying a point mass with moment of inertia are illustrated in Fig. 3 for different values of β_M and β_J . As seen in this Fig. 3, β_M and β_J effects on fundamental mode shapes are significantly observed. Note that β_M have great effects on higher mode shapes.

**Figure 3** First three mode shapes for a cantilever Euler-Bernoulli beam considering the Rotary Inertial Moment of an Attached Mass at right end with (a) $\beta_M = 0.01$ and $\beta_J = 0.01, 1$ (b) $\beta_M = 1$ and $\beta_J = 0.01, 1$

4 CONCLUSION

This paper presents the free vibration analysis of a cantilever beam with a rigid body exciting flexural vibration using SEM. The spectral stiffness matrix of the problem was derived and the first three eigenvalues are determined. For special cases results compared with existing results in literature and very good agreement was achieved. The proposed approach yields high accuracy and rapid convergence. Also, the effect of non-dimensional mass parameter and non-dimensional mass moment of inertia on the dimensionless frequency parameter and mode shapes of the system was investigated. The results show that the values of non-dimensional mass parameter and non-dimensional rotary mass moment of inertia had significant effects on the on the dimensionless frequency parameter and mode shapes of the cantilever beam with attached mass at the free end.

5 REFERENCES

- [1] Rao, S. S. (2004). *Mechanical vibrations*. Pearson Education, Upper Saddle River, NY, USA, 4th edition.
- [2] Kim, H. K. & Kim, M. S. (2001). Vibration of beams with generally restrained boundary conditions using Fourier series. *Journal of Sound and vibration*, 245(5), 771-784. <https://doi.org/10.1006/jsvi.2001.3615>
- [3] Chun, K. R. (1972). Free vibration of a beam with one end spring-hinged and the other free. *Journal of Applied Mechanics*, 39, 1154-1155. <https://doi.org/10.1115/1.3422854>
- [4] Avcar, M. (2014). Free vibration analysis of beams considering different geometric characteristics and boundary conditions. *Journal of mechanics and applications*, 4(3), 94-100. <https://doi.org/10.5923/j.mechanics.20140403.03>
- [5] Banerjee, J. R. (2012). Free vibration of beams carrying spring-mass systems - A dynamic stiffness approach. *Computers and Structures*, 104-105, 21-26. <https://doi.org/10.1016/j.compstruc.2012.02.020>
- [6] Cha, P. & Wong, W. (1999). A novel approach to determine the frequency equations of combined dynamical systems. *Journal of Sound and Vibration*, 219, 689-706. <https://doi.org/10.1006/jsvi.1998.1916>
- [7] Gonçalves, P. J. P., Brenner, M. J., & Elliott, S. J. (2007). Numerical evaluation of high-order modes of vibration in uniform Euler-Bernoulli beams. *Journal of sound and vibration*, 301, 1035-1039. <https://doi.org/10.1016/j.jsv.2006.10.012>
- [8] Hamioud, S. & Khalfallah, S. (2016). Free-vibration of Bernoulli-Euler beam using the spectral element method. *Tehnički glasnik*, 10(3-4), 106-112. Retrieved from <https://hrcak.srce.hr/171830>
- [9] Hong, J., Dodson, J., Laflamme, S., & Downey, A. (2019). Transverse Vibration of Clamped-Pinned-Free Beam with Mass at Free End. *Journal applied sciences*, 9(15), 2996. <https://doi.org/10.3390/app9152996>
- [10] Laura, P. A. A., Pombo, J. L., & Susemihl, E. A. (1974). A note on the vibration of a clamped-free beam with a mass at the free end. *Journal of sound and vibration*, 37, 161-168. [https://doi.org/10.1016/S0022-460X\(74\)80325-1](https://doi.org/10.1016/S0022-460X(74)80325-1)
- [11] Binghui, W., Zhihua, W., & Xi, Z. (2017). Frequency equation of flexural vibrating cantilever beam considering the rotary inertial moment of an attached mass. *Mathematical Problems in Engineering Volume 2017*, 5 pages. <https://doi.org/10.1155/2017/1568019>
- [12] Wu, J. S. & Lin, T. L. (1990). Free vibration analysis of a uniform cantilever beam with point masses by an analytical and numerical combined method. *Journal of Sound and Vibration*, 136, 201-213. [https://doi.org/10.1016/0022-460X\(90\)90851-P](https://doi.org/10.1016/0022-460X(90)90851-P)
- [13] Chang, C. H. (2000). Free vibration of a simply supported beam carrying a rigid mass at the middle. *Journal of Sound and Vibration*, 237(4), 733-744. <https://doi.org/10.1006/jsvi.2000.2990>
- [14] Low, K. H. (2001). On the methods to derive frequency equations of beams carrying multiple masses. *Int J Mech Sci*, 43(3), 871-881. [https://doi.org/10.1016/S0020-7403\(00\)00052-7](https://doi.org/10.1016/S0020-7403(00)00052-7)
- [15] Naguleswaran, S. (2002). Transverse vibrations of an Euler-Bernoulli uniform beam carrying several particles. *Int J Mech Sci*, 44(12), 2463-2478. [https://doi.org/10.1016/S0020-7403\(02\)00182-0](https://doi.org/10.1016/S0020-7403(02)00182-0)
- [16] Naguleswaran, S. (2001). Transverse vibrations of an Euler-Bernoulli uniform beam carrying two particles in-span. *International Journal of Mechanical Sciences*, 43, 2737-2752. [https://doi.org/10.1016/S0020-7403\(01\)00072-8](https://doi.org/10.1016/S0020-7403(01)00072-8)
- [17] Gürgöze, M. (2005). On the representation of a cantilevered beam carrying a tip mass by an equivalent spring-mass system. *Journal of Sound and Vibration*, 282, 538-542. <https://doi.org/10.1016/j.jsv.2004.04.006>
- [18] Gürgöze, M. (1996). On the eigenfrequencies of a cantilever beam with attached tip mass and a spring-mass system. *Journal of Sound and Vibration*, 190(2), 149-162. <https://doi.org/10.1006/jsvi.1996.0053>
- [19] Gürgöze, M. (1998). On the alternative formulation of the frequency equation of a Bernoulli-Euler beam to which several spring-mass systems are attached in-span. *Journal of Sound and Vibration*, 217, 585-595. <https://doi.org/10.1006/jsvi.1998.1796>
- [20] Rossit, C. A. & Laura, P. A. A. (2001). Free vibrations of a cantilever beam with a spring-mass system attached to the free end. *Ocean Engineering*, 28(7), 933-939. [https://doi.org/10.1016/S0029-8018\(00\)00055-X](https://doi.org/10.1016/S0029-8018(00)00055-X)

Author's contacts:

Saida Hamioud, Doctor in Civil Engineering
Jijel, Algeria
hamioud.saida@yahoo.com

The Importance of Developing Preventive Techniques for SQL Injection Attacks

Nenad Bedeković, Ladislav Havaš, Tomislav Horvat*, Dražen Crčić

Abstract: Many intentionally vulnerable web applications are circulating on the Internet that serve as a legal test ground for practicing SQL injection attacks. For demonstration purposes the attacks will target an Acunetix test web application created using PHP programming language and MySQL relational database. In the practical part, the execution of the attack itself largely depends on the database management system, so the displayed syntax is intended only for the MySQL database management system. Example of an automated attack will be executed on SQLmap in a Linux Kali virtualized environment. Security guidelines with a purpose of protecting databases are also discussed.

Keywords: database; MySQL; relational database; SQL; SQL injection attack; SQLmap

1 INTRODUCTION

Nowadays, databases are used for almost all systems and websites because data must be stored somewhere. Collecting massive amounts of structured and unstructured data, known as Big data, has revolutionized today's way of life. Large data amounts are stored in a database so there is an increasing risk involved in the operation of security systems. One of possible security attacks are SQL injection attacks. An attacker sends a malicious SQL query to a vulnerable field that exploits a vulnerability in the SQL implementation of the application itself. Although this vector can be used to attack any database and the most common targets are still web applications. Attack vector is a method used by a malicious person to penetrate or unauthorized access into the vulnerable system.

Providing a unique database entry, most web applications allow visitors to enter and retrieve data. Improperly programmed applications will allow attackers to use vulnerabilities and access the database, after which an attacker can manipulate the content of a database at will. Discovering which database management system (DBMS) is being used is a key to further exploiting the vulnerability. Without discovering DBMS it is not possible to determine which tables to attack, which built-in system functions are involved and which detection methods to avoid. SQL injection is nowadays one of the most popular attack vectors because of its simplicity and does not require a lot of resources and expertise. The attack vector defines the path that is used to take advantage of the security breach.

Since 1998 until today SQL injection attacks still hold the top place according to the Open Web Application Security Project (OWASP) organization, aimed at improving software security. Interestingly, through its lifetime SQL injection attacks have not changed much. A common vulnerability, which is relatively easy to fix, continues to create problems and if not detected before implementation provides a window of opportunity for potential attackers. This paper will demonstrate a direct and automated approach for finding vulnerabilities and performing attacks using SQL injection attacks. A deliberately vulnerable free and legal PHP application on "Acunetix.com" will be used with a purpose of better understand programming errors.

Acunetix was founded to combat the alarming rise in web attacks. It delivers a comprehensive set of options for

automated and manual penetration testing tools, allowing security professionals to perform vulnerability assessments and repair detected threats with just one product. The web application uses MySQL as a DBMS, so we will be using attacks that are specific to that system.

Section 2 gives an insight into SQL injection attacks related scientific papers and newspaper articles. Section 3 defines the basic SQL injection attacks concept, while Section 4 depicts about automated SQL injection attacks. Section 5 provides an analysis of the obtained results and possible prevention techniques. Finally, Section 6 concludes the paper.

2 SQL INJECTION ATTACKS

Despite the fact that more than 20 years have passed since the first SQL injection attacks and the exponential development of technology, we are still witnessing many attacks on vulnerable web-based applications. As the number of available web applications increases, so does the number of potential omissions in the systems that malicious hackers will try to use for their attacks. Numerous techniques for SQL injection attacks have been developed, as well as techniques for preventing such attacks. More about that will be said later in this paper.

2.1 A Brief History of SQL Injection Attacks

In 1998, security researcher Jeff Forristal under the pseudonym "rain.forest.puppy", published an article in Phrack Magazine, describing problems with the Microsoft SQL Server and described how a person with basic programming skills can obtain sensitive information from a database with unauthorized SQL commands on an unsecured website [1]. Author informed Microsoft about the problem but they did not perceive it as a problem. Four years later, more precisely in 2002, SQL injection attacks attracted some attention in computer security community. Daily increase in a number of computer viruses and worms alarmed security experts to start developing software tools for defend against various cyberattack methods. One of the informers was Bill Gates who sent a letter to his employees and warned them about the security their software [2]. This initiative has

triggered the development of secure tools and the gaps created by poor programming are beginning to be taken more seriously. Since then, countless articles have appeared about SQL injection attacks. The first security tool, Analysis and Monitoring for Neutralizing SQL Injection Attacks (AMNESIA), for preventing and detecting SQL injection attacks appeared in 2005. AMNESIA makes a table model of possible malicious SQL code that would be used for the injection and then analyzes each entry point where the code could be inserted [3].

OWASP, a non-profit community on the Internet, aims to improve software security by regularly publishing documentation, tools, and technologies to protect against potential security risks. Every few years, OWASP publishes Top 10 SQL injections attacks and tips for protecting against ten most critical security web application risks [4].

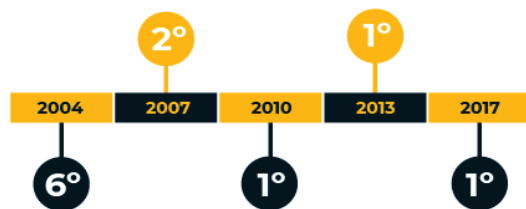


Figure 1 OWASP Top 10 SQL injection attacks by year [4]

It can be concluded that the growing popularity lies in the ease of use. On a vulnerable web application with only a few lines of code on the application layer, it is possible to send queries to the database. Unlike other attack vectors, the resources required to perform the attack are extremely minimal and reduced to a single computer. Other attacks require huge resources and time, involving thousands of computers like a botnet in carrying out attacks. Some of the SQL injection attacks are shown in Tab. 1.

Table 1 SQL injection attacks by year and description

Year	#	Description
2009	[5]	130 million stolen credit card numbers, the biggest identity theft in American history
2009	[6]	RockYou's social media web application development site suffers a leak of 32 million unencrypted user accounts and passwords stored in a text database
2011	[7]	The MySQL official DBMS page had a leak with user accounts and passwords of leading people in the company
2011	[8]	Sony Pictures, the official website of the company Sony, had a database leak with millions of unencrypted user accounts and related information stored in text format
2012	[9]	Yahoo had a database leak with 450,000 user accounts and associated passwords stored in text format
2015	[10]	Vtech, a global supplier of electrical products, suffered a database leak of five million user accounts
2016	[11]	The U.S. National Election Assistance Commission in charge of assisting in the conduct of local and state elections had a database leak of administrator accounts
2019	[12]	Georgia Institute of Technology one of America's leading technical colleges had a database leak of 1.3 million users

2.2 Web Application Architecture

Web application architecture idea is that there is a front-end interface that communicates with the back-end on the server side. Applications are built in layers, so there are

several types of architectures and the choice of architecture depends on the purpose of the application. An example is a web application created using three-tier architecture, consisting of a presentation, logical and data layer. It is often used in applications as a specific type of client-server system, providing numerous benefits for the development environment by modifying the user interface, logic, and storage layers. It is always possible to flexibly develop, upgrade or relocate any layer independently of another. The basic rule of this model is that the presentation layer communicates with the data layer through the logical layer.

The presentation layer belongs to the highest layer of application architecture. It is considered as a communication layer where the client communicates with the application via the user interface and is usually developed through HTML, CSS and JavaScript.

The purpose of the logical layer is to control the application functionalities. It is the main part of the application where the obtained data from the presentation layer is computed and performed. The development of this layer is usually done through PHP, Java, Python, C++, Ruby, .NET, etc.

The data layer consists of a database and a DBMS such as MySQL, Oracle, Microsoft SQL Server, PostgreSQL, etc. The information is here kept separate from the application and the logical layer, so it can be stored and returned on the server.

2.3 Related Work

Great interest in analyzing this security issue was aroused by Jeff Forristal, who studied the security of a growing number of web-enabled applications in the late 1990s. He published his previous research in [1], in which he studied the problem related to Web Bot programs, developed in 1997, and the security problems of MS SQL Server version 6.5.

The development of preventive methods for ASP or PHP SQL injection attacks, which are being worked on intensively, leads to rules that every application developer must incorporate into their applications. At the same time, the development of new versions of database servers and scripting languages, opens up new possibilities for attacking such systems. This is discussed by XuePing-Chen in [13]. He notes that SQL injection is often used in combination with some other hacking methods, which greatly increases the vulnerability of the system.

In addition to the analysis of standard SQL injection attack methods analysis, some authors provide an overview of advanced methods and appropriate defensive mechanisms [14].

When developing web applications, developers try to prevent SQL injection using their personal experience in coding application programs. Despite the extensive experience of these professionals, SQL injection is still the most vulnerable of typical web applications according to a 2014 report from the National Security Agency (NSA) [15].

Trying to prevent as many possible coding flaws as possible, some authors also introduce machine learning

methods in prevention. Thus Kamuto and Soomlek [16] train 4 machine learning models (Support Vector Machine, Boosted Decision Tree, Artificial Neural Networks and Decision Tree) on 1100 samples of vulnerable SQL commands.

In addition to the aforementioned scientific papers on SQL code insertion attacks, there are many review papers that address the issue of comprehensive ways to use the weaknesses of web-based applications and techniques for preventing SQL injection attacks.

Recently, the authors in paper [17] provided an extensive overview of the most commonly used methods and techniques for unauthorized intrusion. Authors also proposed a method for avoiding SQL Injection based on Blockchain technology concept for overcoming this vulnerability via nodes verification using IP.

3 THE BASIC SQL INJECTION ATTACK CONCEPT

SQL injection attacks are usually performed to get inaccessible information by sending queries to the database and manipulating the applications programming logic. Such attacks allow attackers to bypass authentication and gain access to the database, thereby copying, modifying and deleting the data. Reconnaissance and information gathering are always the first step of any successful attack plan and identifying and testing the system configuration greatly helps in choosing how the attack will be performed.

SQL injection attack is a technique where a specially defined SQL query is inserted into the data entry field. This can be any input field in the application or a URL (Uniform Resource Locator) address. A security vulnerability is exploited where the data entered is not checked by the server because the success of the attack depends on the syntax of a correctly written SQL query. If application through sent query responds with an error in the database, it suggests there is a possibility that the application is vulnerable. Using information from the generated error, the attacker is able to understand how the database works.

It is also important to visualize the application programming logic in order to modify the query that caused the original error and test further database behaviour. On the data received by the user, the application server tries to connect to the database and forwards the query. The same communication channel is used for attacking and collecting information, so it is possible to get an insight into the entire database through only one input field. After a successful attack, the attacker can read and manipulate existing data by deleting or creating new ones, thus even becoming a database administrator. If the received entry has passed the check and the data matches those in the database, the query is true, and the user successfully logs in (authentication phase) but if it does not match then it is false and access will be denied. An attack occurs when the application processes a specially written user entry in a way that allows entering the command context instead of data context. This changes the structure of the SQL statement being executed. The ultimate goal of each SQL query is to be true, made up of the resulting entry

(string) and command (statement) that control how the query will be executed. For example, putting any Boolean expression with an always true OR statement that returns true if either of the two conditions are true means that there will always be a true result. In this way, if the entry point is not sanitized and does not check the entered data, the attacker successfully logs on to the page, regardless of the password. Same logic can be used via URL by using web browser.

3.1 Error-based SQL Injection

When the average users encounters a confusing application error, they often ignore it and don't think much about it. In the right hands the information displayed can reveal a lot about how the system works. Errors and the information displayed in them, greatly help the attacker in choosing the vector of the attack and understanding the query sent to the database.

In order to find a way to execute a SQL injection attack, one starts by inserting special characters intended only for the database into the user application input field or directly through the URL. By monitoring the communication between the application and the database attackers follow how it will behave. In the event that an error occurs in the database, the application can generate the error as feedback and provide useful information about the database.

3.2 UNION-based SQL Injection

UNION operator is used to combine two or more SELECT statements results. Each SELECT statement must have the same number of columns, columns in each SELECT statement must have the same order and columns should contain similar data types.

When an application is vulnerable and the query results are displayed within the application response, the UNION operator can be used to retrieve data from other tables that are not visible in the database. The UNION operator is often used in attacks where databases are discovered and leaked.

3.3 Blind SQL Injection

Blind SQL injection derives name from the interaction with a database where an attacker manipulates the database without directly seeing the results of malicious actions. Unlike Error-based attacks, Blind SQL attack does not require any visual cues that something happened to the database.

In previous attack types, queries were made with predictable logical statements as the condition was true or false. Here a condition is used that could be true and thus reveal some information without showing feedback.

Using a Boolean logic, it is necessary to determine whether a query can be considered true or false. By sending a logical query to the vulnerable application database, a possibility of SQL injection attack can be determined. The feedback allows an attacker to assess how logical queries are treated, whether an application returns a true or false

response to a malicious query even though data from the database is not visually displayed.

3.4 Time-based SQL Injection

Another way of performing blind SQL injection attacks are time-based attacks. There are certain variations, but most DBMS support functions for time manipulation. It is based on sending SQL queries where the database will wait a certain amount of time or be late before responding with feedback. The response time will reveal whether the result of the malicious query is true or false, because depending on condition confirmation, the time delay will be executed and the response will be longer than it usually takes.

Sometimes the results of a malicious logical query do not differ much, so the attacker creates an additional difference in the response he receives from the database. For example, MySQL server supports functions like SLEEP() and BENCHMARK() that allow the time delay.

4 AUTOMATED SQL INJECTION ATTACKS

SQL injection attacks can be also performed through specialized software tools and scripts that can perform some of the key tasks required to successfully execute an attack much faster. Some systems or web applications can also be quite complicated, so manual testing requires a higher level of expertise and the ability to track large amounts of programming code. These automated tools are used not only by attackers but also by security experts because constant optimization simplifies and save time in finding vulnerabilities. It should be noted that even if automated tools save time, they are not always considered reliable because there is a possibility they will skip something or due to technical capabilities they cannot perform something that could have been done with the direct approach.

Nowadays there is a wide range of SQL injection attack tools. SQLmap searches and indexes pages on a web application and looks for input windows and forms to send malicious SQL queries with a purpose of causing a system syntax error. It also contains various customization options that allow the implementation of various attack techniques and a functionality for decrypting hashed passwords. The Acunetix web applications will be used for testing.

4.1 SQLMap

SQLmap is one of the most popular and powerful automated SQL injection attack open source tools developed in Python programming language [18]. It automates the detection process and exploits possible security vulnerabilities through SQL injection attack. If an attack is performed through an HTTP vulnerable link, SQLmap can scan, retrieve data from the database and under certain conditions input data into the database.

SQLmap does not have a graphical interface, so everything is done by typing console commands. For the purposes of this paper SQLmap 1.5.5 version will be used on a virtualized Linux Kali environment. Fig. 2 shows SQLmap

console and how the distribution called Kali was created, which contains a large number of tools for penetration testing of computer networks and systems.



```

root@kali:~# sqlmap -h
SQLmap
{1.5.5#stable}
http://sqlmap.org

Usage: python3 sqlmap [options]

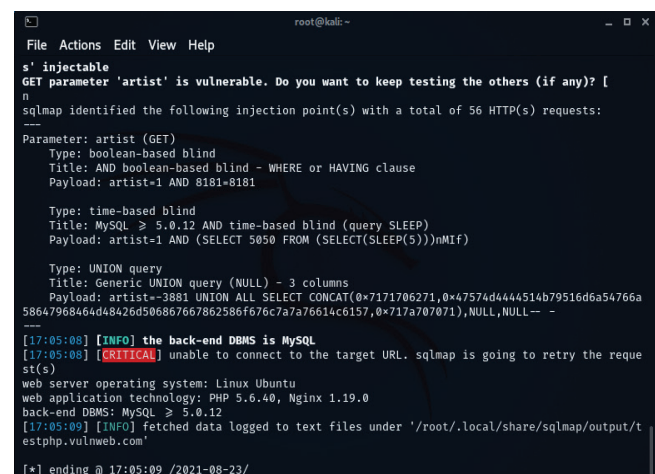
Options:
  -h, --help            Show basic help message and exit
  -hh                   Show advanced help message and exit
  --version             Show program's version number and exit
  -v VERBOSE           Verbosity level: 0-6 (default 1)

Target:
  At least one of these options has to be provided to define the
  target(s)

  -u URL, --url=URL     Target URL (e.g. "http://www.site.com/vuln.php?id=1")
  -g GOOGLEDORK         Process Google dork results as target URLs
  
```

Figure 2 SQLmap console

Through previous examples, it can be concluded the value in the artist parameter is vulnerable. Artist parameter is the identification value on the page, more precisely the string unique to the content. Changing the specified string would display other content, and the prevention of presented vulnerability can be done using specially prepared queries. It is dynamic and possible to change it with a SQL statement. In the console, it is sufficient to specify the link described as the target parameter. Through a series of questions SQLmap will scan and test different attack techniques. Because aspects of the web application are searched and depending on the results, SQLmap will ask what exactly we want to do and what attacks are possible if vulnerabilities are found.



```

root@kali:~# s' injectable
GET parameter 'artist' is vulnerable. Do you want to keep testing the others (if any)? [
n
sqlmap identified the following injection point(s) with a total of 56 HTTP(s) requests:
---
Parameter: artist (GET)
Type: boolean-based blind
Title: AND boolean-based blind - WHERE or HAVING clause
Payload: artist=1 AND 8181=8181

Type: time-based blind
Title: MySQL >= 5.0.12 AND time-based blind (query SLEEP)
Payload: artist=1 AND (SELECT 5050 FROM (SELECT(SLEEP(5)))NMIf)

Type: UNION query
Title: Generic UNION query (NULL) - 3 columns
Payload: artist=-3881 UNION ALL SELECT CONCAT(0x7171706271,0x47574d444451b79516d6a54766a
58647968464d48426d506867667862586f676c7a7a76614c6157,0x717a707071),NULL,NULL--
[17:05:08] [INFO] the back-end DBMS is MySQL
[17:05:08] [CRITICAL] unable to connect to the target URL. sqlmap is going to retry the requ
st(s)
web server operating system: Linux Ubuntu
web application technology: PHP 5.6.40, Nginx 1.19.0
back-end DBMS: MySQL >= 5.0.12
[17:05:09] [INFO] fetched data logged to text files under '/root/.local/share/sqlmap/output/t
estphp.vulnweb.com'

[*] ending @ 17:05:09 /2021-08-23/
  
```

Figure 3 SQLmap scan results

SQLmap supports attacks using the HTTP GET request. GET requests are used when a client, like a web browser, requires certain resources from a server by using the HTTP protocol. Looking at the URL links in the web browser, URL links usually start with a question mark (?), followed by the name of the variable, value and separated by an equal sign

(=). If the link contains more than one parameter then they are separated by the ampersand (&) symbol. Using a specifically designed command SQLmap can confirm that the application is vulnerable by testing different variations of malicious code. SQLmap also details which SQL query should be used to perform an attack. The results shown in Fig. 3 are automatically saved in the system log file for possible future analysis.

SQLmap can also reveal that the operating system, PHP web server and MySQL database system, along with the associated versions. Accordingly, it is possible to go deeper and to find out the names of the database and extract the information. SQLmap documentation contains many configurable parameters but for the purposes of this paper only a few of them, shown in Tab. 2, will be demonstrated.

Table 2 Parameters and their description

Parameter	Description
--dbs	enumerate DBMS databases
--tables	enumerate DBMS database tables
--dump	dump DBMS database table entries
-D	DBMS database to enumerate
-T	DBMS database table(s) to enumerate

Using a few lines of code, SQLmap was able to gather useful information from the database and provide almost direct access (Fig. 4).

```

root@kali: ~
File Actions Edit View Help
web application technology: PHP 5.6.40, Nginx 1.19.0
back-end DBMS: MySQL 8
[05:17:56] [INFO] fetching database names
available databases [2]:
[*] acuart
[*] information_schema

[05:17:56] [INFO] fetching tables for database: 'acuart'
Database: acuart
[8 tables]
+-----+
| artists |
| carts  |
| categ  |
| featured |
| guestbook |
| pictures |
| products |
| users  |
+-----+

[05:17:56] [INFO] fetching columns for table 'users' in database 'acuart'
[05:17:56] [INFO] fetching entries for table 'users' in database 'acuart'
[05:17:56] [INFO] recognized possible password hashes in column 'cart'
do you want to store hashes to a temporary file for eventual further processing
with other tools [y/N] y
[05:17:59] [INFO] writing hashes to a temporary file '/tmp/sqlmap81hy81_i2907/s
qlmaphashes-b7po0czw.txt'
do you want to crack them via a dictionary-based attack? [Y/n/q] y

```

Figure 4 SQLmap database and table names

In a real scenario the attacker would try to gain a higher level of access. To do this, attacker should try to decrypt the hashed passwords and log in as a system administrator which can allow commands directly on the server. SQLmap also supports a dictionary based attack where it immediately tries to recognize and decrypt the found hashed passwords.

5 ANALYSIS

CDNetworks, a company that operates globally by providing content network infrastructure (CDN) services and security services through cloud, published a report in November 2020 on the state of web security for the first half of 2020 [19]. The report states that attacks on web

applications have increased by more than 800% than in the same period in 2019. The main reason for the sudden jump is considered to be the current COVID pandemic, which had a serious impact on the rise of cyber-attacks and it seems logical on the one hand. This global phenomenon has shut down the world population and forced us to use the internet for most of life's socio-economic activities [20]. Comparison of blocked attacks on web application for the first half of 2019 and 2020 can be seen on Fig. 5.

Web Attacks in H1 2019 and 2020

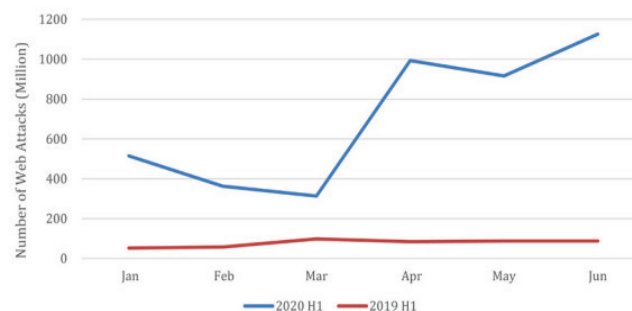


Figure 5 Comparison of blocked attacks on web applications for the first half of 2019 and 2020 [19]

In the first half of 2020, SQL injection attacks made up 16% of all web applications attacks. It indicates that even after twenty years, this vulnerability is still one of the most used vectors of attack. Fig. 6 shows types of blocked web application attacks [19].

2020 H1 WEB Application Violation Types

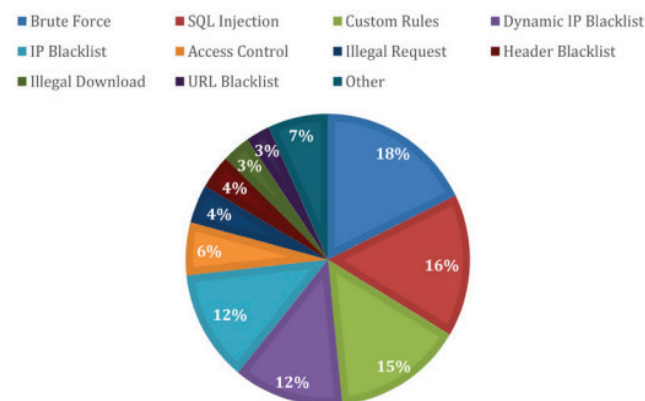


Figure 6 Types of blocked web applications attacks [19]

Akamai Technologies, a company that deals with CDN content, released similar research results in 2020. During the 24-month observation period, looking at all vertical periods SQL injection attacks made up more than 72% of all web applications attacks. The attack rate is halved to 36% if one looks at attacks targeting only the financial sector [21]. Fig. 7 shows overview of web applications attacks during the 24 months period (December 2017 - November 2019).

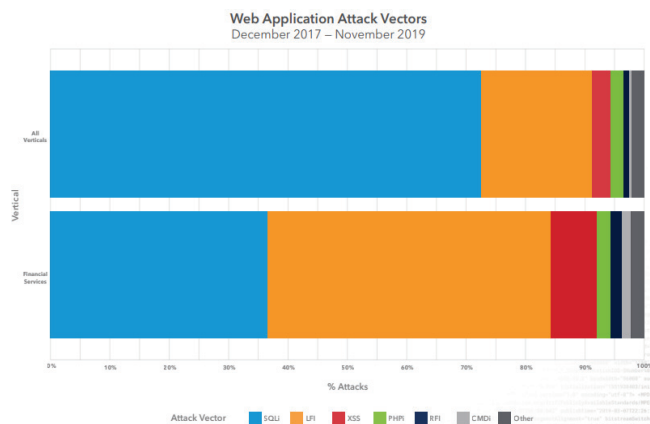


Figure 7 Overview of web applications attacks during the 24 months period, from December 2017 to November 2019 [21]

5.1 Prevention

SQL injection is an old vulnerability and usually taken into account when the application is still in a state of development. In most cases the application dragged the vulnerability out of the development phase. Developers often do not consider security aspects of the applications they work on because they place more attention on the functionalities itself. Such practices not only lead to vulnerabilities but also create additional problems that are mostly solved in the future by software patching.

Various authors suggest using more prevention techniques. Most frequent prevention techniques include using various detecting methods to validate user input as well as using type-safe SQL parameters [22]. Authors in paper [23] proposed an expectation-based prevention technique which address SQL injection attack while calculating the probability of the special characters in the attack dataset, while authors in paper [24] proposed randomization technique that modifies the user input data before the SQL query is executed at the database server.

There are also a few recommended prevention techniques. First of them is through prepared statements. When it comes to prepared statements the SQL code included in the query is predefined, which means that the database will distinguish between code and user input. If an attacker tries to enter a malicious SQL statement, the database will treat the received query as data and not as code. This approach will prevent the resulting query from turning into a malicious one. When preparing a statement, a few rules should be applied:

- prepares SQL statement with empty values are reserved places where the question mark represents a value
- associating a variable with reserved places, specifying each variable along with the data type. Allowed data types are integer, double, string and blob

Second important prevention technique are stored procedures where the principle is somewhat similar to that of prepared statements. The difference is in locating the SQL code, while the queries are prepared in the application logical layer, the stored procedures are located in the data layer called by the application. These are a set of instructions that

perform a predetermined task and their advantage is that it can centralize the data access logic in one place and change access permissions. It boils down to the least privilege principle, which means that if an attacker compromises a stored procedure that contains SQL statements, damage would be limited due to stricter access to the data. Depending on the application in cases where SQL code becomes a necessary part of user input, it is important to make a list of allowed SQL statements. Creating a list of only the most necessary expressions is used to filter invalid statements so that they do not end up in the sent query.

Other researchers results also showed that a significant amount of attacks were blocked using prevention technique which only checks that there are no special characters in an input string. As described before, special characters could potentially change the complete query logic. PHP and MySQL have built-in functions, such as *mysqli_real_escape_string*, which escapes special characters in a string for use in an SQL statement, taking into account the current charset of the connection.

When writing the code that accesses a database, a good practice is to predict possible anomalies at an early stage in order to prevent vulnerabilities instead of resolving through software patches. Nowadays there are various tools for analyzing programming code by providing guidance for best programming practices, adhering to all coding standards with a purpose of identifying potential security issues. The development process should include tests in various forms against attacks because it is important to reduce the risk of future consequences.

6 CONCLUSION

In this paper, using a vulnerable web application, several SQL injection attack techniques are explored. Despite being an older and well-known vulnerability, it is still present, emphasizing the importance of implementing preventive measures. Because of the Internet, we live in a golden age of collecting and storing vast amounts of data. Today's databases allow companies to analyze them for business improvements like profits or to better understand the preferences and needs of their customers. Their security has taken many forms because the consequences of a vulnerability can be catastrophic for any business.

Manual approach demonstrates the attack potential and how to test for vulnerabilities. It has been demonstrated that SQL injection attack is not designed to interact directly with applications, instead it is the application that, given the input received, prepares the SQL code to be sent to the database. The paper also presents the SQLmap software tool, which helps to check the vulnerability of the applications in an automated way that saves time during testing. Automated tools are not a substitute for manual testing, but they are usually sufficient to determine whether an application is vulnerable or not.

In terms of protection against SQL injection attacks, it is always about checking and applying the way how to control the input and output of information in the application. The attack is based on sending a malicious SQL query and as

shown in the paper the principle is almost always the same. The input coming from the user should always be considered potentially malicious and only then effective security measures can be implemented.

7 REFERENCES

- [1] Forristal, J. (1998). NT Web Technology Vulnerabilities. *Phrack Magazine*. Retrieved from <http://phrack.org/issues/54/8.html>
- [2] Gates, B. (2002). Global Information Assurance Certification, Microsoft toward Trustworthy Computing. Retrieved from <https://www.giac.org/paper/gsec/4243/pillars-trustworthy-computing-displayed-patch-management/106837>
- [3] Halfond, W. G. J. & Orso, A. (2006). Preventing SQL injection attacks using AMNESIA. *Proceeding of the 28th International Conference on Software Engineering – ICSE '06*. <https://doi.org/10.1145/1134285.1134416>
- [4] OWASP, (2017). OWASP Top 10. Retrieved from <https://owasp.org/www-project-top-ten/>
- [5] Trotta, D. (2009). Three indicted in largest U.S. identity theft scheme. Retrieved from <https://www.reuters.com/article/us-crime-identity-idUSTRE57G4GC20090817>
- [6] O'Dell, J. (2009). RockYou Hacker: 30% of Sites Store Plain Text Passwords, RockYou breach. Retrieved from <https://archive.nytimes.com/www.nytimes.com/external/readwriteweb/2009/12/16/16readwriteweb-rockyou-hacker-30-of-sites-store-plain-text-13200.htm>
- [7] TinKode & NeOh, (2011). MySQL breach. Retrieved from <https://pastebin.com/raw/BayvYdcP>
- [8] BBC, (2011). Sony investigating another hack. Retrieved from <https://www.bbc.com/news/business-13636704>
- [9] Gross, D. (2012). Yahoo hacked 450,000 passwords posted online. Retrieved from <https://edition.cnn.com/2012/07/12/tech/web/yahoo-users-hacked/index.html>
- [10] Vtech, (2015). Data Breach on VTech Learning Lodge (update). Retrieved from https://www.vtech.com/en/press_release/2015/data-breach-on-vtech-learning-lodge-update/
- [11] Menn, J., (2016). U.S. election agency breached by hackers after November vote. Retrieved from <https://www.reuters.com/article/us-election-hack-commission-idUSKBN1442VC>
- [12] Gorgia Tech, (2019). Unauthorized Access on Georgia Tech Network Exposes Information for 1.3 Million Individuals. Retrieved from <https://www.news.gatech.edu/news/2019/04/02/unauthorized-access-georgia-tech-network-exposes-information-13-million-individuals>
- [13] Ping-Chen, X. (2011). SQL injection attack and guard technical research. *Procedia Engineering*, 15, 4131-4135. <https://doi.org/10.1016/j.proeng.2011.08.775>
- [14] Gudipati, V. K., Venna, T., Subburaj, S., & Abuzagheh, O. (2016). Advanced automated SQL injection attacks and defensive mechanisms. *2016 Annual Connecticut Conference on Industrial Electronics, Technology & Automation (CT-IETA)*. <https://doi.org/10.1109/ct-ieta.2016.7868248>
- [15] National Security Agency, 2014. Defending against the Exploitation of SQL Vulnerabilities to compromise a Network. Retrieved from <https://www.iad.gov/iad/library/ia-guidance/tech-briefs/defending-against-the-exploitation-of-sql-vulnerabilities-to-cfm>
- [16] Kamtuo, K. & Soomlek, C. (2016). Machine Learning for SQL injection prevention on server-side scripting. *2016 International Computer Science and Engineering Conference (ICSEC)*. <https://doi.org/10.1109/ICSEC.2016.7859950>
- [17] Yunus, M. A. M., Brohan, M. Z., Nawi, N. M., & Surin, E. S. M. (2018). Review of SQL Injection: Problems and Prevention. *International Journal on Informatics Visualization*, 2 (3-2), pp. 215-219. <https://doi.org/10.30630/ijov.2.3-2.144>
- [18] GitHub, Miroslav Stampar, (2021). SQLmap History. Retrieved from <https://github-wiki-see.page/m/sqlmapproject/sqlmap/wiki/History>
- [19] CDNetworks, (2020). State of Web Security H1 2020. Retrieved from <https://www.cdnetworks.com/wp-content/uploads/2020/11/CDNetworks-State-Of-web-Security-H1-2020.pdf>
- [20] Communications of the Acm, (2021). A Year in Lockdown: How the Waves of COVID-19 Impact Internet Traffic. Retrieved from <https://cacm.acm.org/magazines/2021/7/253468-a-year-in-lockdown/fulltext>
- [21] Akamai, (2020). State of the Internet, Financial Services Hostile Takeover Attempts. Retrieved from <https://www.akamai.com/content/dam/site/en/documents/state-of-the-internet/soti-security-financial-services-hostile-takeover-attempts-report-2020.pdf>
- [22] Li Qian, Zhenyuan Zhu, Jun Hu, & Shuying Liu. (2015). Research of SQL injection attack and prevention technology. *2015 International Conference on Estimation, Detection and Information Fusion (ICEDIF)*. <https://doi.org/10.1109/icedif.2015.7280212>
- [23] Xiao, L., Matsumoto, S., Ishikawa, T., & Sakurai, K. (2016). SQL Injection Attack Detection Method Using Expectation Criterion. *2016 Fourth International Symposium on Computing and Networking (CANDAR)*. <https://doi.org/10.1109/candar.2016.0116>
- [24] Mehta, P., Sharda, J., & Das, M. L. (2015). SQLshield: Preventing SQL Injection Attacks by Modifying User Input Data. *Lecture Notes in Computer Science*, 192-206. https://doi.org/10.1007/978-3-319-26961-0_12

Authors' contacts:

Nenad Bedeković, MSc, Student
University North, University Center Varaždin,
Jurja Križanića 31b, 42 000 Varaždin, Croatia
nenad.bedekovic@unin.hr

Ladislav Havaš, Assistant Professor
University North, University Center Varaždin,
Jurja Križanića 31b, 42 000 Varaždin, Croatia
ladislav.havas@unin.hr

Tomislav Horvat, PhD
(Corresponding author)
University North, University Center Varaždin,
Jurja Križanića 31b, 42 000 Varaždin, Croatia
tomislav.horvat@unin.hr

Dražen Crčić, mag. ing. el. eng.
University North, University Center Varaždin,
Jurja Križanića 31b, 42 000 Varaždin, Croatia
drazen.crcic@unin.hr

How to Improve the Monitoring of Subjective Analytical Methods and Assessors - The Case of Wine

Ivana Alpeza*, Ivna Linke, Biserka Runje

Abstract: Due to possible various physiological errors in sensory testing, regardless of knowledge and experience, it is necessary to control the results and assessors. This paper aims the possibility of some statistical tools based on laboratory control charts to monitor assessors in wine sensory analyses done in the certification procedure. The data of control charts and reliable results were processed, and individually the results of "100 points" and "Yes/No" methods. The possibility of checking the consistency and concordance of the assessors using wine faults descriptors is also presented. It has been observed that the Chi-square test and some graphical statistics, such as scatter with correlation analysis, can be a good tool in assessors monitoring. The results of sensory analyses affect the wine declaring, and therefore, continued development of technical conditions and tools in the monitoring of assessor's performance is expected to reduce subjectivity and potential errors.

Keywords: assessors; monitoring; subjective sensory analyses; wine

1 INTRODUCTION

In addition to objectively measurable quality, wine acceptability is strongly influenced by the psychological dimension and the fact that the choice of a particular wine is a complex issue burdened with emotions and like / dislike personal aptitude. According to Jackson [1], as Maynard Amerine once said, wine quality is easier to detect than define, and it is more difficult to corroborate positive quality factors than negative ones. Negative quality, such as unpleasant taste or off-odors, is generally easier to identify and describe [2].

The EU wine production and the market are regulated by common regulations and national legislation, and some of them require pre-marketing protocols that involve objective and subjective, chemical and sensory analysis. The principle of consumer protection implies that Sensory characteristics included in the specification of Protected Designation of Origin (PDO) must be guaranteed [3]. The content of some labels is consequently linked to subjective analysis, e.g. Traditional terms, which are important as quality markers and can influence product positioning [4].

1.1 Wine Quality and Subjective Methods of Analyses

Subjective sensory analysis are a large group of methods used during production and wine maturation, bottle-aging, and official quality control, and they definitely have no alternative. Any kind of instrument cannot replace the human senses because it is limited by the key human ability to connect and recognize quality as a whole, regardless of the potential impact of an ingredient. Unfortunately, subjective methods are associated with a constant risk of error; the measuring instruments are the human senses, and, although all external conditions may be at the highest level, subjectivity may cause some of the known errors. According to Stone et al. [5], there are seven groups of psychological errors in sensory testing, and some of them are central tendency (rating of samples using the middle point of scale

and avoiding the extreme ends), error of expectation (rating of samples according to expectation based on previous knowledge of the product information), halo effect (rating the same attributes when they appear in a series of questions differently from when they were tested separately), and stimulus error (assessors rate samples according to the other stimulus and not on their perception from the samples). It is therefore important to conduct sensory analysis in isolated booths by using the proper number of assessors, by training them adequately by placing them in the proper testing environment and conditions, by proper the order of sample presentation, and by using a carefully designed protocol [6]. It is not surprising that the sensory discipline encompasses dozens of ISO standards related to the field of sensory analysis (<https://www.iso.org/ics/67.240/x/>).

Given that the new professional approach in wine sensory evaluation started in the latest fifties of the last century [7], and it should be based on extensive wine knowledge, it is expected that sensory evaluation of wine is permanently built on this knowledge, experience, and research. This is the case with the methods used in the training and selection of tasters; methods and evaluation of results are well developed and mostly standardized, either at the level of broad-spectrum methods, or for sensory analysis of specific purposes or specific products [8-11].

Quality assurance of sensory testing is a demanding discipline, with the need for serious analysis and finding appropriate support tools. Perhaps the most challenging permanent question is "How to judge the wine judges?". There are several interest groups concerning that question. Most often, papers deal with methods and analyses of the quality of sensory testing in research and food industry [12, 13]. The competencies and the results of assessors on International wine competitions are also the subject of interest [14-17]. However, evaluating the sensory compliance of food products and wine with Protected Designation of Origin (PDO) with the sensory description in official specifications is difficult because there are no standard methods for such testing [18]. Some authors

presented the solutions in panel preparing and sensory evaluation in certification of protected products in Portugal and Greece [19-21], but they cannot be used beyond the specific area and medium for which they were created. Perez-Elortondo et al. [18] analysed the current situation of official sensory control of PDO food products and wines in Spain, France, and Italy, the countries that represent almost 70% of the total PDO products registered in the EU. They found a severe problem because of the different methodological approaches and technical criteria in wine sensory evaluation. This kind of sensory testing can be organized through an internal panel with a permanent assessors who work continuously and professionally as measuring instruments. In such cases, assessor monitoring can be incorporated into the methodology of routine testing [19, 20], especially when it is computerized. The laboratory work is complex when a large number of external assessors conduct testing. Although they all need official references, their practical experiences and engagements differ. It is a potential source of an error. Furthermore, there is no repetition in such sensory testing (except control samples), which is also a potential source of error.

There is an increasing demand for platforms that facilitate and integrate panel performance measurements in the routine procedures of a sensory laboratory. Some authors found that only 20% of the studies applied some tool to control the assessors [22]. Guideline for sensory analysis of PDO food products and wine is good basic support [23], but an upgrade is needed. According to Sipos et al. [13], the ISO subcommittee for sensory analysis decided to create a working document for panel performance measurement, especially for those methods which are different from the traditional descriptive analysis.

1.2 Subjective Analytical Methods in Wine Certification in Croatia

The details of wine certification with PDO in Croatia are regulated by national legislation. The sensory testing is a part of certification protocol and should confirm that sensory properties of wine comply with the corresponding specification. With that purpose, the samples are coded, and specific information about samples, such as grape variety, and vintage are presented to assessors. The testing is performed by a 7-member commission (5 assessors) and the "100 points" method that evaluates ten different parameters. Each parameter and selected expression from excellence to inadequate is accompanied by a certain number of points; the ideal wine has 100 points. Sensory results, together with other data that precede sensory analysis, determined the rights to use PDO labels and the Traditional terms "Quality wine KZP" and "Top wine KZP". These traditional terms are therefore objective indicators of quality. The method, "Yes/No" is used in the sensory testing of wines without PDO. This is a descriptive method and five parameters are analysed; limpidity and color (visual aspect), smell, taste, and overall quality.

The assessors who participate in the sensory testing are wine professionals; with an academic level of education in

viticulture and enology, and with at least 5 years of work experience in the wine sector. They have a certificate that confirms abilities and knowledge in wine sensory testing, and are obligated to participate periodically in the aptitude tests.

In addition to the technical conditions (ISO 8589), the commission and the sample preparation and analysis are accredited under the ISO/IEC 17025. The indispensable condition in laboratory work is quality, and various external and internal tools are used to ensure the quality of the results.

The control charts are used to control the results and monitor the assessors. The assessors participate in international periodic proficiency testing with blind samples. Although all assessors are certified, the possible physiological influences are respected when combining them in different commissions. It is known that women are more sensory sensitive [24], so assessors are members of both gender. It is also found that sensitivity changes over the years [25], so assessors represent different age groups. The recommended number of assessors is between five and eight in such analyse [23], which is also provided. Assessors use software produced for this purpose to facilitate their work, to be dedicated to analysis and sample.

Based on individual results and their agreement, the reliable descriptive result (yes/no) and the median (100 points) is defined. When the result is unreliable, the second analysis is carried out in a very short time by another commission.

All the elements explained ensure the objectivity of the examination and reduce the distrust of subjective methods. However, subjective, physiological sources of error cannot be eliminated. Each assessor is an individual measuring instrument and should be monitored individually. The tools that are available and could be generally accepted in this kind of analysis are limited, as already explained [23].

The aim of this study was to examine the use of control charts to develop a new model of statistical tools in the monitoring of results and assessors, to reliable the quality of sensory instruments and low variability, and to guarantee the transparency of testing. This paper is the first one dealing with the problem of subjective methods and monitoring the assessors in wine certification in Croatia.

2 MATERIAL AND METHODS

2.1 Assessors and Sensory Analyses

The research was conducted during 2019, and the results of two assessors produced during 2017 and 2018 were analysed. Two assessors of different ages and gender were randomly selected; assessor A is a 53 old male, and assessor B is a 40 old female. They have more than ten years of experience in wine sensory testing.

They used both of explained methods. In the method "100 points", the result is a median. In the case of the "Yes/No" method, a result is defined by a minimum of three (out of five) same answer. The negative result had to be explained; the assessors could indicate one or more quality descriptors from the list in software created for this purpose. All technical requirements for sensory analyses follow

ISO/DIS 8589 standard, and the procedure is accredited to HRN EN ISO/IEC 17025.

2.2 Data analyses

The results of analyses and the data of control charts were used. Control charts are based on transformed values of results in relation to the points intervals ("100 points" method) and descriptive answers ("Yes/No" method). The intervals of points are as follows: 1-59 (negative result), 60-71 (positive result, but insufficient for PDO), 72-81 (PDO, Traditional term "Kvalitetno vino KZP", minimum median 72), and 82-100 (PDO, Traditional term "Vrhunsko vino KZP", minimum median 82). A value of 0 is recorded for each individual result that is in the same points interval as the final result, and possible deviations were in the range of ± 1 to ± 3 , according to the score ranges for the quality groups and Traditional terms. Deviations can be 1, 2, 3, -1, -2, and -3. For example, with a final result of 85 and an assessor result of 70, the relative deviation is -2, otherwise, the deviation would be 2. In the case of the "Yes/No" method with two possible answers, the deviations -1 and 1 are possible. A value 0 is recorded for each descriptive result equal to the final result. On the horizontal axis are the ordinal numbers of the samples, and on the vertical are the deviations. Fig. 1 presented the first control chart of assessor A used in this study. He tested 50 samples, and had 8 deviations.

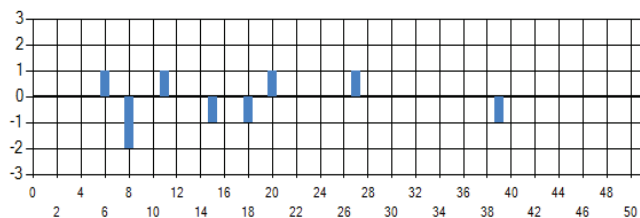


Figure 1 Control chart

Control charts contain all results, including unreliable ones. Therefore, the results were filtered and only reliable results were used in this study. The reliable result is when at least three assessors have a result in the same points interval. A reliable negative result is when at least three assessors have the same descriptors of wine faults. An unreliable result is an error caused by a measuring instrument, but it is not known with which one. To eliminate this unknown source of error and its impact on the reliability of a particular assessor in the study, we included only reliable results in analysis.

The graphical statistics, correlation analyse and Chi-square were used to test the results and the assessors. The data were analysed in Minitab Statistical Software, 2019 (Minitab, LLC, Pennsylvania, USA).

3 RESULTS AND DISCUSSION

3.1 Control Charts and Consistency of Assessor in Relation to Time; Graphical Analysis Use of Graphs

The number of testing with agreement and disagreement for assessors and per year is presented in Tab. 1. The data were calculated from the control charts, and all results were

analysed. The distribution of deviation related to intervals of points is shown in Fig. 2 and Fig. 3.

Table 1 The data of testing and results

Assessor	Year	Samples (No)	Agreement (No, %)	Deviation (No, %)
A	2017	474	409 / 86.3	65 / 13.7
	2018	433	382 / 88.2	51 / 11.8
B	2017	243	199 / 81.9	44 / 18.1
	2018	629	494 / 78.5	135 / 21.5

Most results of both assessors, in both years, were in the same points interval as results of the commissions (Tab. 1). Assessor A had a deviation in 13.7% of testing compared to the total number of testing in 2017, while in 2018, the numbers of variations were smaller. The decrease of the total number of deviations is evident in all points intervals, not only in one specific (Fig. 2).

In the case of assessor B, we found more deviations; 18.1% in 2017 and 21.5% in 2018. Variability of his results is observed in the number of variations and intervals (Fig. 3). More precisely, assessor B had a worse matching with commissions in 2018 than in 2017, and the most significant deviation was observed in intervals 1 and -1. While the share of deviations in interval -1 was significantly reduced (assessor perceives the samples as lower quality), the share of deviations in interval 1 increased in 2018 (assessor perceives the samples as better quality than the commissions).

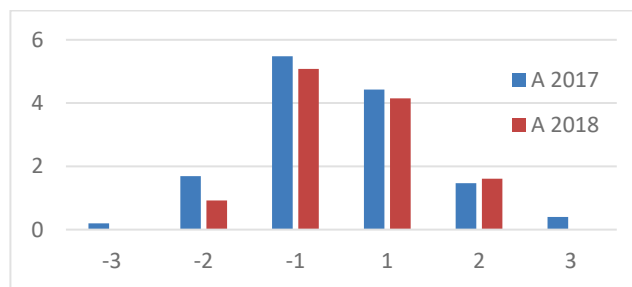


Figure 2 Assessor A; results with interval disagreement (%)

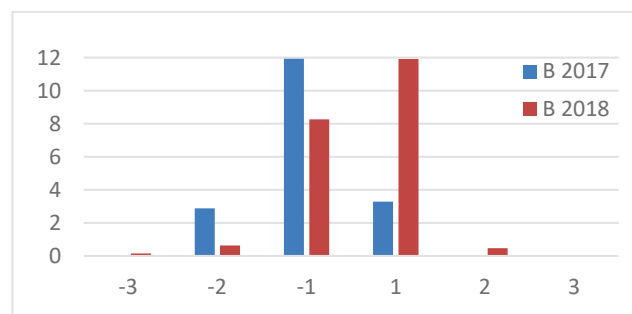


Figure 3 Assessor B; results with interval disagreement (%)

3.2 Assessor and 100-Point Method; Use of Scatter Diagram, Correlation Analysis and Chi-Square Test

A scatter plot is a graphical presentation of values in a coordinate system; the "scattering" of values determines the shape, direction, and intensity of the connection between the two variables. These diagrams are also useful for monitoring

a variable over time. In this analyse the horizontal axis presents the results of the assessor, and the vertical axis the results of the commissions for the same samples. Scatter diagrams confirm the existence of a linear relationship between the assessors and the corresponding commissions during both years (Fig. 4).

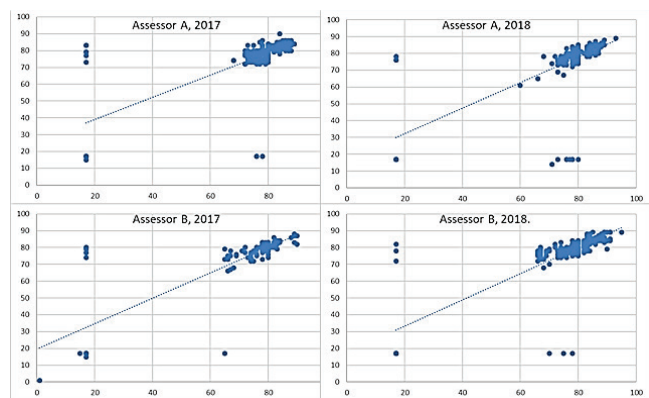


Figure 4 Scatter plot, "100 points" method

The scatter plots are used in determining the correlation and we chose the Pearson correlation coefficient as the most common measure of the direction and strength of the linear statistical correlation of the two variables. The correlation strength is defined by the range of the Pearson coefficient within the interval $[-1, 1]$. In all cases in this study, the risk of error is less than 1% given that the p-values in all cases are 0.000. In all four cases, the obtained coefficients are in the range of medium and high correlations and indicate good compliance between the assessors and the commissions, and in the appropriate direction. In the case of assessor A, Pearson's correlation coefficient was 0.68 for 2017 (257 results) and 0.66 for 2018 (264 results), which means a medium correlation. The correlation strength in the case of Assessor B was high, with a coefficient of 0.83 in 2017 (142 results) and 0.81 in 2018 (362 results). This analyse found a stronger correlation between the assessor B and the corresponding commissions, compared to assessor A, in the case of the 100-point method.

A Chi square test (χ^2) was used to test the consistency of assessors during the time. Hypotheses for the test were:

H_0 : There is no statistically significant difference between the assessor's results compared to the commissions' results during the observed period.

H_1 : There is a statistically significant difference between the assessor's results compared to the commissions' results during the observed period.

In both tests, the P-value is greater than 0.05 (Fig. 5 and 6), which means there is no evidence that there is a difference between the results of the assessor and the commissions during the observed period. Thus, it can be concluded that assessors A and B were consistent throughout the observed period in case of the "100 points" method.

It can be concluded that both assessors presented the sensitivity and consistency in case of the "100 points" method. However, a longer period than two years should be analysed to make a more significant conclusion.

These examples presented that the Chi-square test can be applicable in monitoring of assessors. It could be a good tool

for analysing the differences between assessors when they participate in multiple commissions, or when they are the permanent members of the panel in evaluating certain properties. In our study, this was not possible because assessors A and B were not members of the same commissions. Furthermore, it would be recommended to analyze a more extended period to ensure the relevance and the significance of the test.

Category	Observed	Historical Counts	Test proportion	Expected	Contribution to Chi-Sq
-3	0	2	0,007782	2	2
-2	3	4	0,015564	4	0,25
-1	22	17	0,066148	17	1,47059
0	206	215	0,836576	215	0,37674
1	18	14	0,054475	14	1,14286
2	8	5	0,019455	5	1,8
N	DF	Chi-Sq	P-Value		
257	5	7,04019	0,218		

Figure 5 χ^2 test (observed and expected counts and Chi-Square test), assessor A, 100 points method

Category	Observed	Historical Counts	Test proportion	Expected	Contribution to Chi-Sq
-2	3	4	0,011111	2	1,37178
-1	24	47	0,130556	18	1,89
0	107	287	0,797222	111	0,13126
1	5	19	0,052778	7	0,74391
2	0	3	0,008333	1	1,16
N	DF	Chi-Sq	P-Value		
257	5	7,04019	0,218		

Figure 6 χ^2 test (observed and expected counts and Chi-Square test), assessor B, "100 points" method

3.3 The Descriptors of Wine Faults as a Tool in Monitoring of Assessors

Since it is mandatory to explain the negative result and unacceptable wine quality with sensory descriptors from the available application menu, we compared the most commonly used descriptors of the commissions with the descriptors of the assessors. The wine oxidation, volatile acidity combined with the ethyl acetate, and reductive smells of hydrogen sulphide and mercaptans were tested.

It can be observed that assessor B is more sensitive in the perception of the discussed wine faults compared to the commissions (Fig. 8). In both years, assessor B has more results with the three most common wine faults than commissions. Monitoring of all negative results confirmed that she has a lower tolerance for wine quality problems in general; both in frequency and in the type of descriptor. She found more samples, with different descriptors, compared to the commissions. The choice of the same descriptors concerning assessor/commissions was more pronounced in the case of assessor A than B. There are fewer differences in general, and assessor A presented a more moderate approach in the perception of wine faults (Fig. 7). He had a smaller number of negative results compared to the commissions. This analysis pointed that the reductive smells of hydrogen

sulphide and mercaptans are the main cause of disagreement between assessor B and commissions. At the same time, it was the descriptor with the smallest deviation between assessor A and the commissions. Furthermore, this analysis shows where the problem of disagreement is most pronounced. Assessor A presented more sensitivity to oxidation than commissions, and very good matching with the commissions in case of reductive smells and volatile acidity, while assessor B definitely has a different perception of hydrogen sulphide and mercaptans compared to the commissions. This kind of analysis can be a good tool, not only in monitoring the assessors, but also in targeting of thematic trainings, and in planning of periodical aptitude tests.

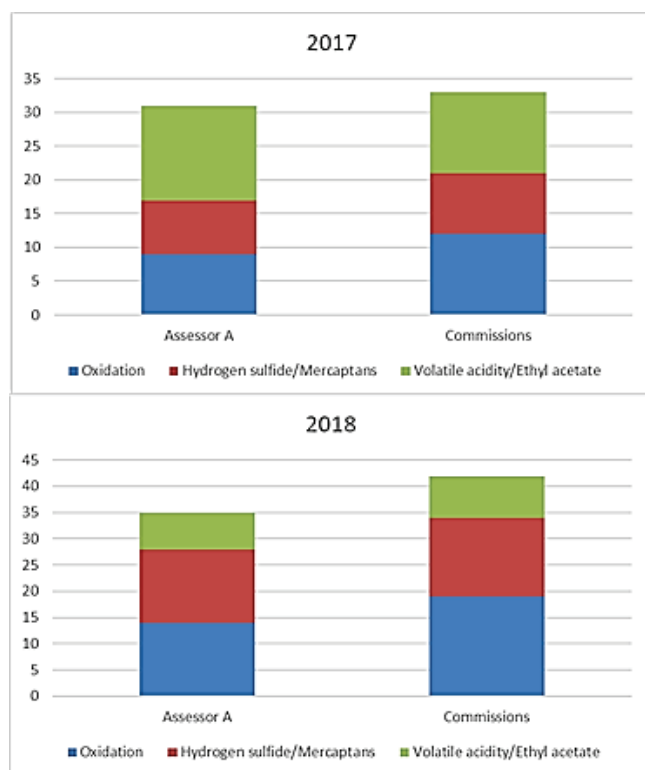


Figure 7 Descriptors of most frequency wine faults, assessor A and commissions

The assessors training, selection, and monitoring are subject of different international standards, like ISO 11037 (colour vision), ISO 3972 (taste perception), and ISO 5496 (odour sensitivity). However, using these standards, it is possible to establish a panel that can detect and describe sensory differences between the products, but not to establish and evaluate the quality of products. A problem with wine sensory testing is a huge variety of products and creations that are not accompanied by appropriate product descriptions, and it limits and complicates the training of assessors, as well as control of their work. Once the assessor selection is completed, other questions such as the monitoring of the assessors during their regular activities have to be dealt with [11]. Monitoring of assessors is not a problem when a permanent panel performs sensory analysis, or when the product has specification with detailed described sensory parameters, and different statistical models can be found in the literature [13, 26].

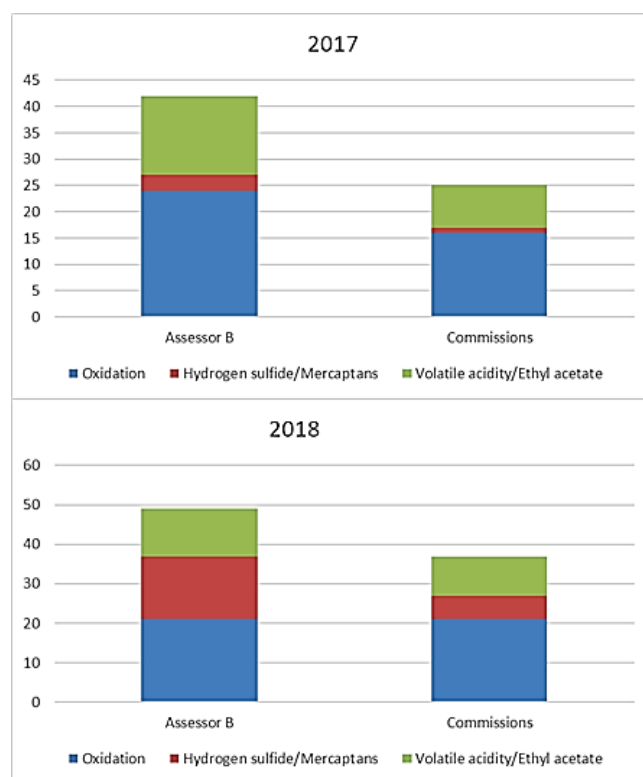


Figure 8 Descriptors of most frequency wine faults, assessor B and commissions

Unfortunately, we could not find the references on monitoring assessors in laboratory work when the testing is a part of pre-market wine certification performed with the larger number of external assessors in a different combination, as in our case. It is presented in this study that control charts can be used for the new statistical analysis thanks to intervals of result values. We chose the Chi-square statistic and showed how this test could be an excellent tool to upgrade the control charts; for assessors and or methods results. In similar sensory analysis, it is possible to use some other control methods, but when repetition of sample testing is performed. Pinto [27] recommended pre-testing panellists with the Triangle test and Cronbach's Alpha coefficient as a measure of assessor's consistency. Some other authors [17] also presented Cronbach's Alpha criterion for positional analysis because it allows a deeper look at the consistency and variability of assessors. This model could be useful when assessors rarely participate in testing or are limited to a specific product or parameter. However, in everyday laboratory work, as in our case, this model is complicated, increases the number of samples, and is not appropriate due to pre-expected error.

It is known that there is no repetition of samples in wine sensory analysis in certification. Sample replication is desirable but not considered in routine laboratory testing. According to Moskowitz et al. [28], when well-trained assessors are instruments with low variability, the data are still valid. Despite all the elements of quality assurance of these sensory tests based on which the result of one testing is reliable, the risk of error still exists. Replication has been found to be a good way to monitor the assessors and their mutual agreement in the case of descriptive sensory analysis. A replicated test is defined as a new independent

measurement taken under the same set of conditions as the original one. Repetition is helpful when it is possible to control samples for qualitative properties and limit the sample information presented to the assessor. For example, Pinto [29] proposed a new methodology to determine an estimate of uncertainty to be applied in interpreting results related to sensory testing of wines for the certification of some specific geographical names in Portugal. The basis of his research was the repetition of sample analysis. Unfortunately, in the case of this and older study [27], it is not clear what sample data were presented to the assessors in the test, which may be a source of error and limit the possibility of repeating the test in routine work.

However, there are test conditions when replication does not guarantee the reliability of the results. In our study, assessors in sensory testing have some crucial information about the sample; grape variety, vintage, specific technology, and some analytical data. We did some experiments, but they recognized the same sample in replicating. This is probably reinforced by the fact that the sample should be presented in the same group concerning the recommendation of the order of analysis [30]. In such conditions, the sense of applying repeat testing of the same sample is questionable due to an error that is difficult to determine, and we know it exists.

One of the assessor's work requirements is an agreement with the panel in quality descriptors and their intensity evaluation (reliability/repeatability). It is very difficult to find any information in the primary literature about the panel or assessor monitoring based on descriptors of wine faults. According to Hodgson's examination of judge reliability at a major US wine competition, the assessors tend to be more consistent in wine's sensory testing they do not like than opposite [2]. The detection of wine faults in interlaboratory tests is presented to support the improvement of the assessor's sensibility and performance [31].

According to Perez-Elortondo and Zannoni [32], if the results of one of the assessors fall outside the acceptable limit of performance, the re-training should be recommended and the assessor should provide adequate results. It is presented in this study that the analysis of attributes used in the description of wines with faults can be useful in routine laboratory work. Several aspects of control can be included; assessor sensitivity analysis, perception analysis, and consistency analysis. Sensitivity is incorporated in the final perception of the intensity of an attribute, which directly affects the result and is therefore very important in the training and selection of assessors.

The analysis of two assessors in our study raises the question of what is more important to use as a basis for evaluating the assessor's reliability; consistency or concordance. This analysis offers the thesis that the consistency of assessors is a better basis for estimating error, which is further important in their selection. Indeed, a deeper analysis (more results and a more extended period) can upgrade the graphical model with an appropriate and significant statistical tool that will define the limits of error acceptability.

Ensuring the reliability of the results requires constant monitoring of all factors that may be the source of the error. Although the reliability of the results in the analysis we have studied here is very high due to the numerous input insurance

references, including assessors who are wine experts, it is necessary to raise the scale of quality constantly.

4 CONCLUSION

Sensory analyses are burdened with many potential sources of error. Although the external causes of errors can be completely controlled, the measuring instruments (assessors) are a constant risk due to physiologically caused errors regardless of knowledge and experience. Monitoring the results is therefore not only a recommendation but also the "*Conditio sine qua non*". The monitoring of results aims to uncover the factors that produce the inconsistency. This paper presented the possibilities of some statistical tools in control of the wine assessor's work in the case of two methods of sensory analyses; "100 points" and "Yes/No". Control charts have been used as a basis for testing other options. It is presented that correlation analysis and non-parametric Chi-square test are effective tools to check the reliability and consistency of assessors. It was pointed out that graphical analysis of the type and frequency of selected wine faults descriptors is an excellent basis for monitoring the sensitivity and consistency of assessors and can be supported in the planning of training programs. These tests can be used to compare the assessors of the same panel, test the assessor's results compared to the panel, and test the assessor's consistency over time. This research needs to continue; it is proposed to include a larger number of assessors and a longer period of time in the analysis. The next level of development of quality assurance in laboratory work could include upgrading the software with selected statistical tools to monitor the results of assessors in routine work that will provide significant indicators of the quality of sensory analysis.

5 REFERENCES

- [1] Jackson, R. S. (2017). *Wine Tasting*. A Professional Handbook. Chapter 8 - Nature and Origins of Wine Quality. 337-370. Academic Press, Elsevier Ltd. <https://doi.org/10.1016/B978-0-12-801813-2.00008-2>
- [2] Hodgson, R. T. (2008). Evaluation of wine judges at a major U.S. wine competition. *Journal of Wine Economics*, 3(2), 105-113. <https://doi.org/10.1017/S1931436100001152>
- [3] Regulation (EU) No. 1151/2012 on quality schemes for agricultural products and foodstuffs.
- [4] Commission delegated Regulation (EU) No. 2019/33 supplementing Regulation (EU) No 1308/2013 regarding applications for protection of designations of origin, geographical indications and traditional terms in the wine sector. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019R0033>
- [5] Stone, H., Bleibaum, R. N., Thomas, H. A., & Sidel, J. L. (2012). *Sensory Evaluation Practices*. 4th edition. Academic Press, Inc., San Diego, CA.
- [6] Meilgaard, D., Civille, G. V., & Carr, B. (2007). *Sensory Evaluation Techniques*, 4th edition. CRC Press, Boca Raton, FL. <https://doi.org/10.1201/b16452>
- [7] Amerine, M. A., Roessler, E. B., & Filippello, F. (1959). Modern sensory methods of evaluating wine. *Hilgardia*, 28(18), 477-567. <https://doi.org/10.3733/hilg.v28n18p477>
- [8] Perez Elortondo, F. J., Ojeda, M., Albisu, M., Salmerón, J., Etayo, I., & Molina, M. (2007). Food quality certification: An

- approach for the development of accredited sensory evaluation methods. *Food Quality and Preference*, 18(2), 425-439. <https://doi.org/10.1016/j.foodqual.2006.05.002>
- [9] González, M. M., Navarro, T., Gómez, G., Pérez, R. A., & de Lorenzo, C. (2007). Sensory assessment of table olive: I. Set up of a panel test and use of standardised scales. *Grasas Aceites*, 58(3). <https://doi.org/10.3989/gya.2007.v58.i3.176>
- [10] López-Aguilar, J. R., Valerio-Alfaro, G., Monroy-Rivera, J. A., Medina-Juárez, L. A., O'Mahony, M., & Angulo O. (2007). Evaluation of a simple and sensitive sensory method for measuring rancidity in soybean oils. *Grasas Aceites*, 57, 149-154. <https://doi.org/10.3989/gya.2006.v57.i2.30>
- [11] Etaio, I., Albisu, M., Ojeda, M., Gil, P. F., Salmerón, J., & Pérez Elortondo, F. J. (2010). Sensory quality control for food certification: A case study on wine. Panel training and qualification, method validation and monitoring. *Food Cont.*, 21, 542-548. <https://doi.org/10.1016/j.foodcont.2009.08.011>
- [12] King, M. C., Hall, J. W., & Cliff, M. A. (2001). A comparison of methods for evaluating the performance of a trained sensory panel. *Journal of Sensory Studies*, 16, 567-581. <https://doi.org/10.1111/j.1745-459X.2001.tb00321.x>
- [13] Sipos, L., Nyitrai, Á., Hitk, G., Friedrich, L., & Kókai, Z. (2021). Sensory Panel Performance Evaluation - Comprehensive Review of Practical Approaches. *Appl. Sci.*, 11, 11977. <https://doi.org/10.3390/app112411977>
- [14] Cicchetti, D. (2006). The Paris 1976 Wine Tastings Revisited Once More: Comparing Ratings of Consistent and Inconsistent Tasters. *Journal of Wine Economics*, 1(2), 125-140. <https://doi.org/10.1017/S193143610000016X>
- [15] Gawel, R. & Godden, P. W. (2008). Evaluation of the consistency of wine quality assessments from expert wine tasters. *Australian Journal of Grape and Wine Research*, 14, 1-8. <https://doi.org/10.1111/j.1755-0238.2008.00001.x>
- [16] Hedozeau, H. C., Ballester, J., Chatelet, B., & Lempereur, V. (2015). Wine competition: From between-juries consistency to sensory perception of consumers. *BIO Web of Conferences*, 5, 03009-03015. <https://doi.org/10.1051/bioconf/20150503009>
- [17] Khalafyan, A. A., Temerdashev, Z. A., Abakumov, A. G., & Akinshina, V. A. (2021). Study of consistency of expert evaluations of wine sensory characteristics by positional analysis. *Heliyon*, 7(2), e06162. <https://doi.org/10.1016/j.heliyon.2021.e06162>
- [18] Pérez-Elortondo, F. J., Symoneaux, R., Etaio, I., Coulon-Leroy, C., Maître, I., & Zannoni, M. (2018). Current Status and Perspectives of the Official Sensory Control Methods in Protected Designation of Origin Food Products and Wines. *Food Control*, 88, 159-168. <https://doi.org/10.1016/j.foodcont.2018.01.010>
- [20] Etaio, I., Gil, P., Ojeda, M., Albisu, M., Salmerón, J., & Elortondo F. P. (2012). Improvement of sensory quality control in PDO products: An example with txakoli white wine from Bizkaia. *Food Quality and Preferences*, 23(2), 138-147. <https://doi.org/10.1016/j.foodqual.2011.03.008>
- [21] Pinto, M. M. (2021). Evaluating uncertainty in sensory analysis. A case study of the panel of tasters of the Dão Regional Wine Commission. *Ciência e Técnica Vitivinícola*, 36(1), 22-31. <https://doi.org/10.1051/ctv/ctv2021360122>
- [22] Tataridis, P., Kannelakopoulos, C., Gatselos, L., & Logothetis, S. (2013). Sensory Evaluation of Wines as a part of Certification of Protected Designation of Origin in Greece and the EU. *AMPELOS 2013, Trends in world vitiviniculture development*. Santorini Island, Greece. <http://ampelos2013.conferences.gr/ampelos2013-proceedings.html>
- [23] Djekic, I., Lorenzo, J. M., Munekata, P. E. S., Gagaoua, M., & Tomasevic, I. (2021). Review on Characteristics of Trained Sensory Panels in Food Science. *J. Texture Stud.*, 52, 501-509. <https://doi.org/10.1111/jtxs.12616>
- [24] Doty, R. L., Applebaum, S., Zusho, H., & Settle, R. G. (1985). Sex differences in odor identification ability: A cross-cultural analysis. *Neuropsychologia*, 23(5), 667-672. [https://doi.org/10.1016/0028-3932\(85\)90067-3](https://doi.org/10.1016/0028-3932(85)90067-3)
- [25] Corwin, J., Louri, M., & Gilbert, A. N. (1995). Workplace, age, and sex as mediators of olfactory function: data from the National Geographic Smell Survey. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.*, 50, 179-186. <https://doi.org/10.1093/geronb/50B.4.P179>
- [26] Bi, J. & Kuesten, C. (2012). Intraclass Correlation Coefficient (ICC): A Framework for Monitoring and Assessing Performance of Trained Sensory Panels and Panelists. *Journal of Sensory Studies*, 27(5), 352-364. <https://doi.org/10.1111/j.1745-459X.2012.00399.x>
- [27] Pinto, F. S. T., Fogliatto, F. S., & Qannari, E. M. (2014). A method for panelists' consistency assessment in sensory evaluations based on the Cronbach's alpha coefficient. *Food Quality and Preference*, 32, Part A, 41-47. <https://doi.org/10.1016/j.foodqual.2013.06.006>
- [28] Moskowitz, H. R., Muñoz, A. M., & Gacula, M. C. (2004). *Viewpoints and Controversies in Sensory Science and Consumer Product Testing*. Food & Nutrition Press, Connecticut, USA. <https://doi.org/10.1002/9780470385128>
- [29] Pinto, M. M. (2021). Evaluating uncertainty in sensory analysis. A case study of the panel of tasters of the Dão Regional Wine Commission. *Ciência Téc. Vitiv.*, 36(1), 22-31. <https://doi.org/10.1051/ctv/ctv2021360122>
- [30] OIV. (2009). Standard for international wine competitions and spirituous beverages of vitivinicultural origin OIV-CONCOURS 332A-2009 Retrieved from <http://www.oiv.int/public/medias/4661/oivconcours-332a-2009-en.pdf>
- [31] Mazzoni, K., Tirard, A., & Boubetra, B (2022). Detection of Wine Alterations by Sensory Analysis: Overview of Results obtained from Interlaboratory Tests. *Medicon Agriculture & Environmental Sciences*, 2(1), 04-13.
- [32] Perez-Elortondo, F. J. & Zannoni, M. (2021) Guidelines for Sensory Analysis of Protected Designation of Origin Food Products and Wines. *European Sensory Science Society (E3S)*. E3S-PDO Product working group. Retrieved from: <https://www.e3sensory.eu/guidelines-for-the-sensory-evaluation/>

Authors' contacts:

Ivana Alpeza, PhD
(Corresponding author)
Croatian Agency for Agriculture and Food,
Centre for Viticulture, Enology and Edible Oils,
10000 Zagreb, Jandričeva 42, Croatia
099 7054843, ivana.alpeza@hapih.hr

Ivna Linke
Croatian Agency for Agriculture and Food,
Centre for Viticulture, Enology and Edible Oils,
10000 Zagreb, Jandričeva 42, Croatia
Ivna.linke@hapih.hr

Biserka Runje, PhD, Full Professor
University of Zagreb,
Faculty of Mechanical Engineering and Naval Architecture,
10000 Zagreb, Ivana Lučića 5, Croatia
biserka.runje@fsb.hr

Tensile Strength of Carbon Fiber

Atanas Kochov*, Martin Spasovski, Filip Serafimovski

Abstract: Carbon fiber is a material with a vast area of utilization. When combined with a resin, we get a composite, which provides great mechanical properties, has low mass and can be utilized in many different applications. Carbon fiber cloth can be found in a wide range of forms and shapes, differing from weight to the shape of their weave. Manufacturing a composite part can be achieved with a variety of techniques and processes, each of them having advantages and disadvantages over the other. One of the most common process for producing composite parts is the open moulded hand lay-up process. This process is economically viable, has low cost for equipment, and can produce parts with satisfactory results. A composite sheet is manufactured using the hand lay-up process. Three testing samples are taken out from the composite sheet, and their tensile strength is determined using the ASTM D3039 standard.

Keywords: composite material; hand lay-up; high performance; light weight

1 INTRODUCTION

The composite material produced with the combination of polymer matrix and reinforcement fibers, provides a high-performance composite. The composite material obtained with combination of carbon reinforcement fibers and epoxy-based polymer matrix, can provide the same mechanical properties of conventional materials used in the industry, like aluminum or steel, with the benefit of being lightweight. Carbon fiber composite, compared to aluminum can have a weight reduction of more than 20%, while when compared to steel, the weight reduction exceeds 50% [1]. Because of the weight saving and the rest of the mechanical properties the carbon fiber is primarily used in aerospace and space industries. This material is taking a major progress as well as in the high-end car industry, civil engineering, wind turbine blades and sporting goods [1]. There are large number of companies and research facilities which are running research and development programs to develop processes and technologies for mass production.

To maximize the performance of the fibers in the composite structure, the orientation of the fibers has to be arranged in the direction of the loads, according to the geometry of the part. The fibers can be arranged as a unidirectional configuration in which all fibers are placed in parallel and as a quasi-isotropic in which fibers are placed at 0, 90 and 45°, as presented in Fig. 1 [1].

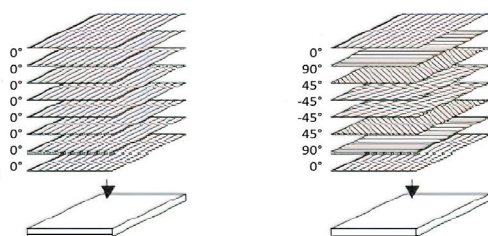


Figure 1 Unidirectional and quasi-isotropic laminates

To take full advantage of the mechanical properties of the reinforcement fibers, it is important to have a high-volume fraction of fibers in the material (55-60%), for the

purpose of avoiding fiber curvature or misalignment as well as limiting the void content in the resin (<3%) [1].

There are different types of technologies and processes to manufacture a part from carbon fiber material. These technologies and processes depend of the form of the fibers. The carbon fibers can be found and used in the form of long-fibers, short-fibers or woven in the form of sheets (Fig. 2 [3]). Most promising is the application of the carbon fiber weave in combination with thermoset resins, of which epoxy resins are the most commonly used. Epoxy resin have high mechanical strength, good chemical corrosion and exhibit high performances on elevated temperature of 120 °C [2]. Based on these custom-built sheets complex structures can be produced. The anisotropy and inhomogeneity offer the designer many aspects for structural optimization, but these characteristics also lead to complex stress and strain states within the material.

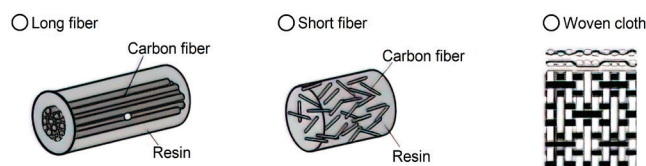


Figure 2 Various arrangements of carbon fibers

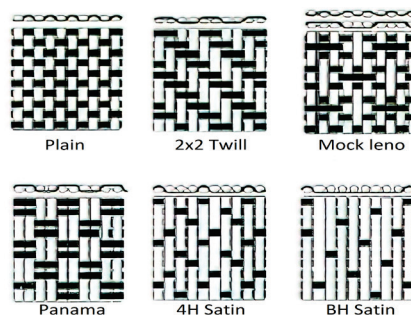


Figure 3 Various types of weaves

The easiest way to manufacture composite parts is hand lamination. For the purpose of reaching high quality its required simple tooling and good manual skills. This process

is performed at room temperature. The volume of reinforcement fiber in this process is limited to 40% due to the nature of the process. This process produces a composite part with many imperfections and high void content, due to the lack of vacuum assistance or other type of pressure on the composite part. The reinforcement fibers can be found in a wide variety of weaves (Fig. 3), with different weights and forms, each offering different benefits.

To achieve the best result in mechanical properties, a combination of weaves is required to be used.

2 TESTING METHOD AND MECHANICAL PROPERTIES

The test method determines the in-plane tensile properties of polymer matrix composite materials, reinforced by high-modulus fibers. The composite material forms are limited to continuous fiber or discontinuous fiber-reinforced composites in which the laminate is balanced and symmetric with respect to the test direction. A thin flat strip of material having a constant rectangular cross section is mounted in the grips of a mechanical testing machine and monotonically loaded in tension while recording load. Determining the ultimate strength of the composite material is carried out by testing the maximum load that the composite coupon can carry before failure. By monitoring the strain of the coupon with strain or displacement transducers through the process of testing, a stress-strain curve can be obtained, from which the ultimate tensile strength, Poisson's ratio, tensile modulus of elasticity and transition strain can be derived.

ASTM D3039 provides the minimum requirements for designing a specimen for composite coupons, presented in Tab. 1. These requirements are insufficient for producing a properly dimensioned coupon drawing with tolerances. Therefore, recommendations from Tab. 2 of ASTM D3039 are taken into accord for manufacturing a specimen [4].

Table 1 Tensile specimen geometry requirements [4]

Parameter	Requirement
Coupon Requirements:	
shape	constant rectangular cross-section
minimum length	gripping +2 times width + gage length
specimen width	as needed
specimen width tolerance	$\pm 1\%$ of width
specimen thickness	as needed
specimen thickness tolerance	$\pm 4\%$ of thickness
specimen flatness	flat with light finger pressure
Tab Requirements (if used):	
tab material	as needed
fiber orientation (composite tabs)	as needed
tab thickness	as needed
tab thickness variation between tabs	$\pm 1\%$ tab thickness
tab bevel angle	5 to 90°, inclusive
tab step at bevel to specimen	feathered without damaging specimen
See Table 2 from ASTM D3039 standard for recommendations	

This method of testing, provides valuable information for the materials specifications, which are beneficial for

research and development, analysis, quality assurance and structural design.

The tensile strength of the composite material is dependent on many factors: type of material, lay-up of the material, preparation, specimen stacking, specimen preparation, specimen conditioning, specimen alignment, the environment of which the testing is conducted, type of grips used on the machine, speed of testing, temperature, void content in the coupon and volume percent reinforcement.

A carbon fiber sheet is manufactured for the purpose of testing the mechanical properties of a composite, who is produced using the hand lay-up process. From the composite sheet, three samples are taken which are used for testing the tensile strength of the composite, dimensions are presented on Fig. 4.

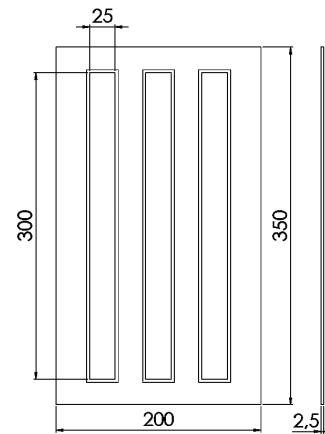


Figure 4 Cut scheme for test samples

3 CARBON FIBER MANUFACTURING PROCESSES

There are many different processes for manufacturing carbon fiber components, varying from manual process of manufacturing, with hand lay-up of carbon fibre cloth, to a more complicated process, using autoclave ovens with pre-impregnated carbon fiber. Each of these processes has its own advantages and disadvantages when compared to another process of manufacturing. A big consideration should be taken beforehand in the requirements of the manufactured part, and choosing the suitable process for the right project.

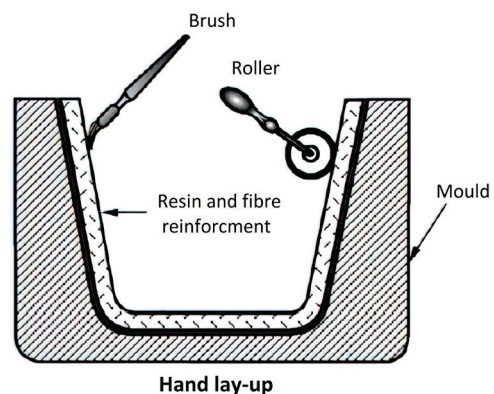


Figure 5 Description of a hand lay-up process

One of the oldest and most common technique for manufacturing carbon fiber components is the open moulded: hand lay-up process, presented on Fig. 5 [7]. This process, has low demands and requirements for tools and environmental conditions, but has a high requirement in the workers skill [6].

The final results achieved with this process, can vary from high grade produce with satisfactory results, to a low grade produce with many imperfections, weak areas (due to excess resin), voids, pin holes, etc., depending on the workers skill.

Hand lay-up process, consists of laying carbon fiber cloth in the mould, and soaking the cloth with a predetermined amount of resin. The wetting of the carbon fiber cloth, can be achieved with a brush, squeegee or a roller. This task is repeated, until the desired thickness of the laminate is achieved.

Before starting the process of hand lay-up of the carbon fiber cloth, the mould surface has to be thoroughly cleaned with acetone, to remove any traces of wax or grease. Also, the mould surface has to be dust free, to prevent inclusions of dust particle in the surface of the part.

To prevent sticking between the part and the mould, there has to be a barrier between these two parts. This barrier is attained with the help of a release agent, which can be: PVA (polyvinyl alcohol), wax, or a suitable chemical release agent. One of the oldest and most common release agents is PVA, which is a release agent dissolved in ethanol. This release agent is commonly used in combination with wax and provides excellent results in releasing the part from the mould.

Peel ply is a woven cloth typically made of nylon, glass or other synthetic materials. This layer of woven fabric, is applied as the final layer of the composite part to provide a porous surface, suitable for adhesive bonding to other surfaces. Other than that, peel ply is used for drawing excess resin from the part, in combination with a bleeder layer.

4 MANUFACTURING OF A CARBON FIBER SHEET

Manufacturing of the composite sheet occurs in a room with low humidity and temperature of 21 °C, which is in the range of the optimal room temperature for producing composite parts.

The carbon fiber sheet is manufactured using the open moulded hand lay-up process. The sheet is made of three individual layers of 12K carbon fiber cloth, who's weave is a 2x2 twill weave pattern, presented on Fig. 6. This weave pattern is composed of tows who go in directions of 0° and 90°, and give an illusion of a diagonal pattern. The 2×2 (also can be found in 4×4 weave) in the name stands for tows going under two tows, and then again over two tows of carbon fibers. This weaving is often used for complex shapes, because of its looser weave, which can be easily manipulated to fit more complex shapes in the mould.

The dimensions of the cloths are 350×200 mm, and approximately 0.6 mm thickness. In theory these layers should produce a ~1.8 mm thick carbon fiber sheet. The

fabric to epoxy ratio in this process is 50:50. The weight of the carbon fiber cloth used for manufacturing of the carbon fiber sheet, is approximately 130 g, so the amount of resin used will be about 155-160 g. The amount of resin used is higher due to the method of application with a brush. This is done to prevent a shortage of resin, because of the nature of the brush to soak a small amount of resin. Manufacturing of this carbon fiber can be done in moulds made of a variety of materials, ranging from wood, aluminium, composites moulds to 3D printed plastics moulds.

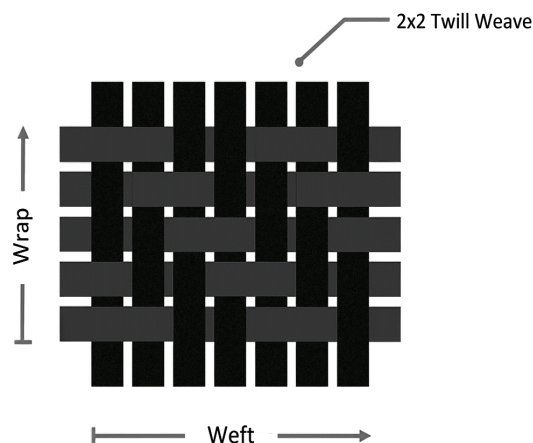


Figure 6 Representation of a 2×2 Twill Weave Pattern

4.1 Manufacturing Process

The manufacturing of the carbon fiber sheet is done on, glass surface which serves as a mould surface, because it provides a flat surface (which allows producing a carbon fiber sheet with a smooth finish), and has no porosity (no need for using a sealant for preventing a mechanical bond between the mould and the carbon fiber sheet). The glass surface is cleaned beforehand with acetone, to remove any traces of grease or dust. A layer of Rexco Formula Five Mould Release Wax is applied using a clean dry cloth, according to the instructions given in the manual. After applying the final layer of wax, the mould is left for one hour for the residual solvents to gas-out before applying PVA. A uniform layer of PVA is applied using a spray gun. After applying the release agent, the PVA is left to dry itself for 10 – 15 min, to create a very thin release film.

The resin used in this process of producing a carbon fiber sheet is SIK A CR-82 epoxy resin, with SIK A CH80-1 catalyst. This resin has low viscosity and is suitable for use in the process of hand lay-up and vacuum bagging where curing temperature more than 75 °C can't be achieved. The mixing ratio of this epoxy system per manufacturer instructions is 100:27 [9], when mixing the two parts by weight.

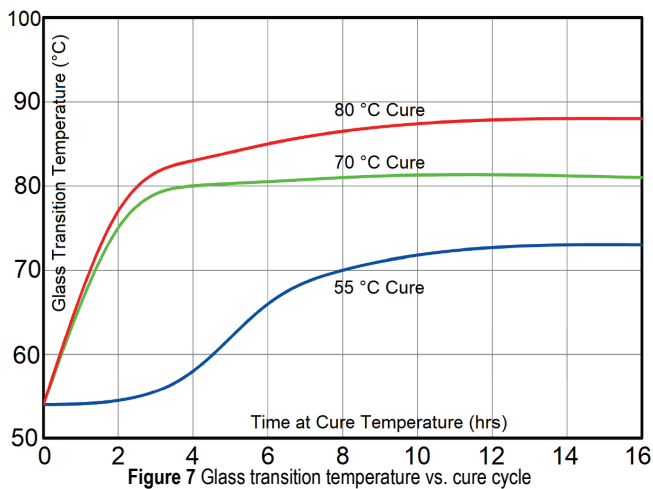
To reduce the number of voids and pinholes on the surface area on the manufactured sheet, a layer of resin is applied on the mould surface. After wetting the mould surface, a layer of carbon fiber cloth is applied on the surface. The cloth is wetted out using a brush, and the excess resin is removed with the help of a squeegee or a roller. This part is

repeated with the remaining layers of carbon fiber cloth. After applying the last layer of carbon fiber cloth, a final layer of peel ply is applied on the composite sheet, and using a roller, the excess resin is drawn out from the part.

After finishing with the stacking of the layers of carbon fiber cloth and finalizing with the peel ply, the carbon fiber sheet is left to cure for 24 hours at room temperature. Additionally, the composite sheet is post cured in an oven for eight hours at 80 °C, Fig. 7 [9]. This is done to acquire a fully cured part, and obtaining the best properties.

Table 2 Properties for HDT and GTT for Biresin CR82 with CH80-1 catalyst

Biresin CR82	ISO Standard	Hardener CH80-1	Unit
Heat distortion temperature	ISO 75A	72	°C
Glass transition temperature	ISO 11357	83	°C



With this epoxy resin system, different glass transition temperature can be achieved, depending on the curing conditions of the composite part. Curing the composite sheet at elevated temperature of 80 °C for additional 8 hours, raises the glass transition temperature of the composite sheet to a higher tolerance.

5 TESTING AND TEST RESULTS

From the manufactured carbon fiber sheet, three samples were cut-out with dimensions of 300 mm length, and 25 mm width, presented on Fig. 8. The three coupons are used to test the tensile strength of the manufactured carbon fiber sheet. The test is implemented using the ASTM D3039 standard, using a 250 kN tensile strength tester. The test of the three coupons was conducted in a room temperature environment.

The test sample is mounted on the tensile strength testing machine, using suitable grips, used for holding the specimen, and providing a sufficient and even distribution of pressure to prevent slippage of the specimen during the test. The two grips hold on combined length of 120 mm of the specimen (60 mm on each side), presented on Fig. 9. The displacement between the grips determined by the ASTM D3039 standard is 2 mm/s.

To calculate the tensile strength of the composite samples, Eq. (1) was applied:

$$\sigma_{\max} = \frac{F_{\max}}{A}, \quad (1)$$

where: $\sigma_{\max}$ - ultimate tensile strength (MPa); $F_{\max}$ - maximum load before failure (N); A - average cross section of the coupon (mm²).

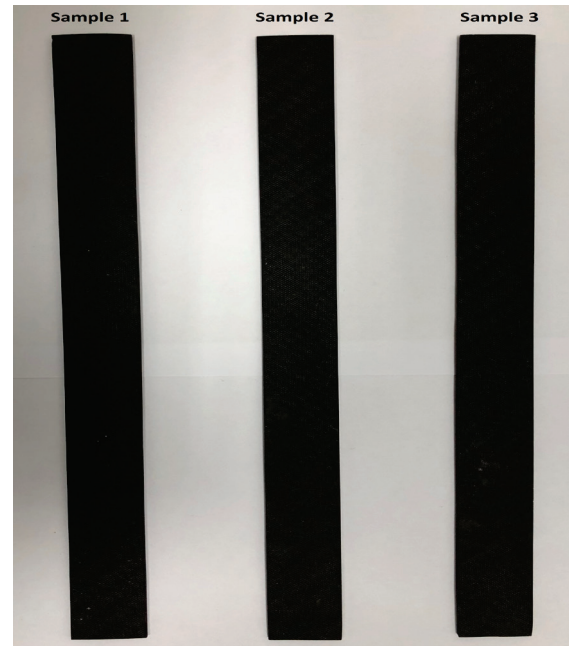


Figure 8 Representation of the test samples before the testing procedure



Figure 9 Mounted carbon fiber sample

The results of the ultimate tensile strength of the coupons are shown in Tab. 3.

Table 3 Results from the ultimate tensile test

Test Samples	$\sigma_{\max}$ (MPa)
Sample 1	539,24
Sample 2	512,62
Sample 3	488,58

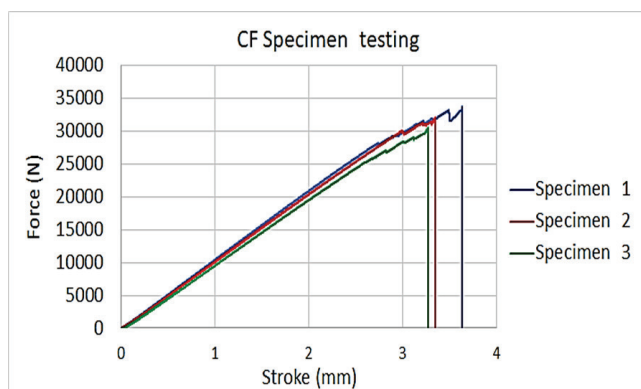


Figure 10 Graphic representation of the conducted tensile test on the three samples

Specimen 1 withstood the highest load of the three coupons, reaching fatal failure at 34.1 kN and extension of 3.6 mm. Specimen 2 withstood maximum load of 32.6 kN and extension of 3.3 mm before fatal failure. Specimen 3 withstood the lowest load of 3.1 kN and extension of 3.2 mm before fatal failure, shown on Fig. 10.

The specimens after testing are presented on Fig. 11. Specimen 2 and 3 have two break points.



Figure 11 Aftermath of the test samples

6 CONCLUSION

Open moulded hand lay-up process combined with its low cost for tools, provides adequate product, which almost entirely depends on the workers ability. The final result from the manufacturing of the composite sheet, is a product with some surface voids, pinholes and trapped air. The surface of the carbon sheet, due to the nature of the process (lack of vacuum environment) has many voids. To get a surface rid of any imperfections and voids, the sheet needs to undertake additional treatment to get a perfect smooth finish. The final thickness of the manufactured sheet is 2.5 mm. When compared to the thickness of the carbon fiber cloth, which is approximately 1.8 mm, it's concluded that the remaining 0.7 mm is gained from the resin. With this procedure, the final product is resin rich, which in turn gives us a part with an increased mass and lower mechanical properties in contrast to the other carbon fiber manufacturing processes, due to the excess resin.

7 REFERENCES

[1] Drechsler, K., Heine, M., & Mitschang, P. (2021). *Carbon Fiber Reinforced Polymers*.

- [2] Daniel, C. & Durand, C. (2016). Composites technologies for transport applications: A roadmap.
- [3] See <https://www.plastics.toray/trmx/en/torayca/index.html>
- [4] ASTM D3039, Standard Test Method for Tensile Properties of Polymer Matrix Composite Materials
- [5] See <https://coventivecomposites.com/explainers/tips-for-working-with-satin-weaves/>
- [6] Middleton, B. (2016). *Composites: Manufacture and Application*.
- [7] Swift, K. G. & Booker, J. D. (2013). Chapter 5 - Plastics and Composites Processing, *Manufacturing Process Selection Handbook*, 13(5.8) 141-174.
<https://doi.org/10.1016/B978-0-08-099360-7.00005-7>
- [8] Elkington, M., Bloom, D., Ward, C., Chatzimichali, A., & Potter, K. (2015). Hand layup: understanding the manual process. *Advanced Manufacturing: Polymer & Composites Science*, 1(3), 138-151.
<https://doi.org/10.1080/20550340.2015.1114801>
- [9] See <https://industry.sika.com/en/home/advanced-resins/composite-resin-systems/composite-systemsforwetlay-up/sikabiresin-cr82.html>

Authors' contacts:

Atanas Kochov, PhD, Prof.
(Corresponding author)
Faculty of Mechanical Engineering - Skopje
Karposh II bb, 1000 Skopje, Republic of North Macedonia
+389(0)71-299-299, atanas.kochov@mf.edu.mk

Martin Spasovski, MSc
Faculty of Mechanical Engineering - Skopje
St. 518 No. 2 Marino, 1041 Skopje, Republic of North Macedonia
+389(0)75-239-472, spasovskim.01@gmail.com

Filip Serafimovski, BSc
Faculty of Mechanical Engineering - Skopje
St. 534 No. 62 Marino, 1041 Skopje, Republic of North Macedonia
+389(0)78-299-006, filip.serafimovski96@gmail.com

Information Model Development – Attributes and Power of Information

Tomislav Mrčela*, Tamara Zadravec

Abstract: Starting from the general concept of a system-process, which is defined by the quality of relations between individual elements, we can single out attributes whose elements have a dominant character of the information system, all in the projection of the quality of the observed process. In any state of the process, information is undoubtedly significant as a subsystem, whose property cannot be ignored, because it is a dynamic process that bases its dynamism in the time domain. The paper precisely defines the information power model as well as its response to the occurrence of perturbations in a strictly defined time domain, all in the function of one process-event. The described complexity can be a function of horizontal complexity, whereby a hierarchical component is frequently lost, or it can be another component resulting from the overall information process. The paper deals with measurements of the response to an event that took place at the global level, that is, the finals of the UEFA Champions League 2021. A unique functional relationship of power of information in strictly defined marginal conditions was obtained. The paper presents the spectrum of the power of occurrence-opening of the portal of the mentioned event through the functional relationship of the dynamic process. All the described phenomena gave new insights in the development of the information model.

Keywords: attributes speed of information; information; model; power

1 INTRODUCTION

Using scientific methods in studying information systems and their characteristics, and considering other scientific attributes, an appropriate concept of the current information system situation can be defined. Sustainability of the information system depends on its current condition as well as on the initial or null condition. The information system structure is a complex set of state quantities that essentially influence its behaviour. By their nature, the state quantities can be measurable or non-measurable, discrete or continuous, current, past or future, as input and output parameters of the observed information system. The number of state quantities of the observed information system essentially tends towards infinity. For pragmatic reasons, the state quantities will be observed as a finite set – if they are discrete, and information systems have all the attributes of discretion; logic, set theory, design theory, information theory, complexity theory and probability theory. In this paper, a semi-closed interval of values is observed, without studying the state quantities that have properties of continuity. If the information system is observed through the function of time, it can be considered as a dynamic system that strongly changes its states as a function of time. Based on the principles of development, this information system and all the phenomena will be subject to changes, even though a strong reduction of reality was done. The changes in our information system require a change of every component as well as of every phenomenon. Real information systems have the property of inertia, that is, it is impossible to change their state in a very short time. The rate of change, that is, its gradient, changes within wide limits. The reaction time of the information system is called the transient phenomenon of the information system. Historically, information systems had properties that were defined by space and time. Generally, the two most complex systems are biological and information systems. At the moment of designing the state of the social system, the most

important thing is its axiomatic definition and that is the implication of its problem with reality. The general concept of the state of the social system has input values (of the past, current and future time) as well as output values (of the past, current and future time), which are connected by a set of mapped functions. Every scientific discipline must primarily start from the definition of the content that is the subject of its study as well as the definition of boundaries of the observed social system. This paper implicitly studies the special state of the social system, or more precisely, it measures the power of information in real time of an even at the global level. According to the logical postulate that each input change must change the parameters of individual elements of the observed system, it is up to the observer to reduce the insignificant responses of the mapped function, using their own scientific backgrounds, and try to bring the activated dynamic value system to a stable state. This algorithm for social systems is most easily implemented if the basic set is broken down into partial subsets and interconnected by internal connections – relations according to the horizontal and vertical principle.

A systematic analysis with the following steps is to be applied:

1. **Problem definition** is an activity of problem transformation in which an unadjusted event is limited to the observed phenomenon, that is, a sports event, whereby the power of information is measured.

2. **Designing the observation subject** – in order for this activity to come to life, the impact of a sudden event, the football competition finals, is reduced to important factors only.

3. **Transformation of the problem into a system – model** – on the basis of similarity and analogy, relations are set within the elements of the observed social system. This is the most important moment when the impact of the problem on individual elements of the observed social system is assessed and the final solution is made, aiming to completely

map the input unexpected changes to the output in the social environment.

4. Testing the obtained results in the designed model – it is the most criticized process in social systems because it is a consequence of the information credibility and it must contain the basic testing postulates, such as measurability and reliability of the obtained values – results.

5. Implementation of the selected designed model – real-time to the social system and its application through an additional factor, which is the final response of the social system as a complex dynamic and discrete system to a random variable of unknown aetiology. [1-4]

2 INFORMATION MODEL DEVELOPMENT

Modelling is one of the general scientific methods. The modelling method can be applied in all scientific research. Therefore, if a person realises that some information takes place in space and time, their desire is to master the information, in the context of predicting the course and time of the process they are observing. Conclusion by analogy belongs to the basic type of indirect reasoning. The essence of this way of reasoning lies in the assumption that two phenomena are similar in some attributes. Therefore, the observer may conclude that the other attributes are similar as well. In order to fully understand the modelling programme, we need to master the cognitive process. The assumptions of the cognitive process are the following: connecting similar attributes, environment, knowledge, information and perceptions. All these properties cannot be implemented into the cognitive process by the observer without emotions. Abstract reasoning is a series of related thoughts that have a sequence towards a specific goal, for the purpose of solving a problem. Abstract thoughts, regardless of their origin, intuitively and discursively must have a logical basis. The process of reasoning is the last stage in a series of defining the cognitive process, preceded by understanding and judgment. When defining an information model, we must start from several types of reasoning - three-ductive, deductive or inductive. Analysing the power of information through its appearance, we come to conclusions by analogy or similarity. Powerful information phenomena in space and time are impossible to understand and model without knowing the basic postulate of information flow (the power of information). The following figure shows the original form of an information flow model as well as its result (the power of information). In science, in addition to physical analogy, economic, biological and linguistic analogy are often used. Therefore, we can say that Titus Lucretius Carus (99 BC-55 BC), using linguistic analogy and poetic intuition, connected space and time based on the theory of relativity. [5]

Poetic thought gave depth to the struggle between the real and the imaginary world. This example cannot be explained in depth if we do not have enough information regarding time, speed and quantity.

The power of the information process is based on knowledge and memory at the time of observing the subject event. Interaction and understanding of the whole process

leads to the judgment and conclusion, controlled through the cognitive process. Therefore, it can be concluded that the power of information depends on three key attributes, namely: the quantity of information, speed of information and the observed time-moment.

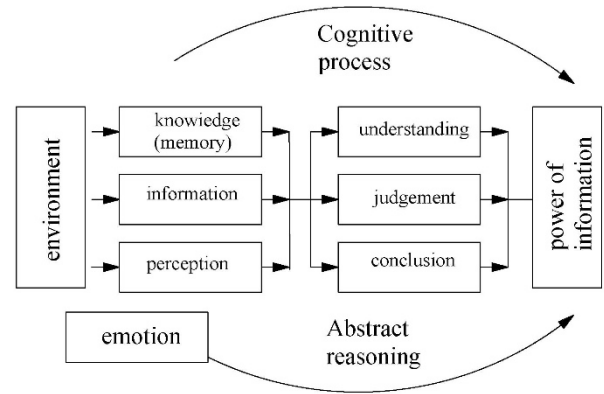


Figure 1 Information model / attributes /

The logical basis for modelling a social model is making conclusions based on similarity and analogy. The essence of conclusion is making an objective decision based on new judgment, all in line with the established boundary conditions of the observed social model. In order for the observed model to exist in the real world, we must start from the object of conclusion, the process of conclusion and the axioms with rules on the basis of which the conclusion is made.

On the example that follows, in the context of structure, similarities and analogies, we can set up information models that have in-depth properties of isomorphism or homomorphism. The power of information has unambiguous properties of isomorphism as a piece of information enters the space of observation, and has all the dynamic properties that such an isomorphic model can possess in the observed space. A general question may be asked: where is the end of a particular piece of information? A piece of information appears in the observed time, it has its lifespan, it asymptotically approaches one constant and remains in the observed space forever. The presented model of the information, cognitive process is an abstraction of a special kind, that is, it is a link between abstract and theoretical reasoning and objective reality. In other words, it is a means by which the regularities and relationship in objectively existing information systems are defined. An information model is at a higher level if it contains more isomorphic elements, that is, it reflects a higher level of information reality when the presentation of information power is more transparent. Information power (P_i) is the function of speed (v_i), quantity (q_i) and reliability (p_i).

$$P_i = f(v_i, q_i, p_i) \quad (1)$$

Fig. 1 shows that it is impossible to complete the cognitive process, that is, abstract thinking with the current power of information, and all implemented in the

environment, without a strong emotional component of the observer of the information process.

3 ANALYSIS OF THE OBSERVED INFORMATION EVENT

On 29 May 2021, a global event took place, that is, the UEFA Champions League final in 2021 was played in Porto at the stadium Estádio do Dragão. The measurement of information flow was performed by registering the number of openings in a certain time interval. A time interval of 120 minutes was chosen for the analysis. The beginning of the match was chosen as the boundary condition for the start of the measurement, which took place at 9 p.m. CET. The following web portal was used: <https://www.vecernji.hr/sport/uzivo-od-21-sat-finale-lige-prvaka-1496367>.

This portal gives public access to the number of active page openings in a unit of time. The units of time is one minute. The precise data for the observed interval are shown in the chart in Fig. 2.

If we start from the fact that the speed of opening the mentioned portal is also an attribute of the established model of information flow, we can conclude with a high degree of certainty that the power of information P_i has a dynamic character. The dynamic character of information power also reveals all the complexity of the observed social model.

The results are shown in the chart in Fig. 2.

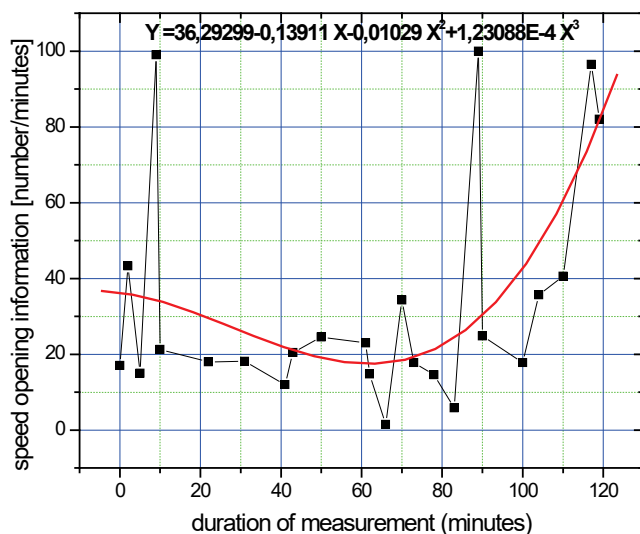


Figure 2 Speed of opening the portal in the observation period

The measurement of the described event on the web portal started on 29 May 2021 at 9 pm CET and ended at 11 pm. The results are presented in the chart, which shows an extreme speed of 100 pieces of information per minute in the 9th minute of the observed event. It is the moment at which a turnaround in the football match happened – one of the teams took the lead. The second peak happened in the 89th minute due to the fact that at that moment, the football match ended in terms of active play. In the further course of the measurement, an increase in the speed of opening the portal is noticeable, as a result of the power of information that are determined by the final result of the observed measurement.

Fig. 2 shows an approximation by third-degree polynomial, whose coefficients are given above the approximated curve. It can be concluded that certain properties of information power fully confirm the hypothesis that information power is a function of quantity, speed and time.

At the moment of measuring, that is, at 9 p.m., the number of openings was 350, and at the time of completion of the measurement of the defined time interval 3 370 openings. The total measurement interval ended on 30 May at 2 p.m. and the total number of openings was 10 700. By analysing the attribute of the established model, it can be undoubtedly concluded that social information models are complex models that can be simulated only if input information has high reliability. It is impossible to analyse a social model while reducing the attributes of the model; in other words, to transform properties of isomorphism into homomorphism. The measurement of results was performed in the interval from (0 to 1,060) minutes. Discrete data were approximated, all with the aim of confirming the thesis that the speed of access to the page is not constant. [6-8]

4 CONCLUSION

1. A general model of information process can be established that contains all the key attributes of the system theory, namely the influence of the environment as an attribute of space, the influence of information flow rate in the observed space and the influence of quality (reliability) of information as a modern sociological phenomenon.
2. The paper defines a special form of information model that has been tested on a significant global event.
3. Studying the obtained results of the research and the measurements, it can be concluded undoubtedly that the power of information changes in the process of occurrence of an event, that is, it is a function of time and as such contains all the dynamic attributes of occurrence.
4. In the observed event, as regards the spectrum – density of information, we can precisely define the moment at which a significant perturbation occurred – the opening of the defined portal (9th and 90th minute), in light of the events in the process of the described measurement (the change of result and the end of the match)

5 REFERENCES

- [1] Devide, V. (1984). *Matematička logika*. Školska knjiga, Zagreb. (in Croatian)
- [2] Athans, M. (1985). *Networks and Computations*. McGraw-Hill, New York.
- [3] Hubka V. (1984). *Teorie der Maschinensysteme*. Springer-Verlag, Berlin Heidelberg New York. (in German)
- [4] Špiranec, S. & Banek Zorica, M. (2008). *Informacijska pismenost: teorijski okvir i polazišta*. Zavod za informacijske studije, Zagreb. (in Croatian)
- [5] Mrčela, T., Dujmović Bocka, J., & Zadravec, T. (2021). Design of a social system decision-making model in crisis. *Interdisciplinary management research XVII*, Osijek: Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek, Croatia, pp. 218-239

- [6] <https://www.eric.ed.gov/ERICDocs/data/ericdocs2/content-storage-01/0000000b/80/23/4a/12>. Accessed: January 29, 2021.
- [7] <https://www.educause.edu/ir/library/html/erm/31231>. Accessed: January 29, 2021.
- [8] <https://www.vecernji.hr/sport/uzivo-od-21-sat-finale-lige-prvaka-1496367>. Accessed: May 29, 2021.

Authors' contacts:

Tomislav Mrčela, prof. dr. sc.
J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and Information Technology,
Kneza Trpimira 2b, 31000 Osijek, Croatia
mrcela.tomislav@gmail.com

Tamara Zadravec, dr. sc.
J. J. Strossmayer University of Osijek,
Faculty of Philosophy,
Lorenza Jäger 9, 31000 Osijek, Croatia
tzadravec@ffos.hr

Development of a Test System to Make Circularity Measurements of Rubber Bushing After Stress Relief

Muhammed Abdullah Ozel*, Selim Sefa Baysal, Cemil Sungur

Abstract: Rubber metal bushings are one of the most important parts of the suspension system. These parts are specially designed to minimize the vibrations and noises that vehicles are exposed to while driving. In this study, it is aimed to measure the circularity of rubber metal bushings after stress relief. A modular design is provided according to the tolerance values of the bushings in the inventory. Simultaneously with deep learning and image processing algorithms, laser circularity measurements are verified with reference to bush circularity and deformation points are determined. The detected deformation is transferred to the HMI screen. With this study, it was ensured that the circularity errors in the rubber metal bushings were taken under control and the problems that may be encountered after the production were eliminated, a more sensitive solution was provided, and a quality production and accurate results were obtained.

Keywords: bushing; circularity; deep learning; image processing; quality control

1 INTRODUCTION

The wishbone, which is one of the most important components of vehicle suspension parts, has important components within itself. Wishbone bush is one of these parts. The bushing, which enables the swing to work better, is one of the most important parts that absorb the effects such as vibration, shock and impact from the road and the moving parts of the vehicle, preventing their transmission to the vehicle body and providing the driving comfort requirement. The bushings, which can have various geometries according to the place and task used, are basically single-layer bushings consisting of inner bushing, outer bushing and rubber material filling between the bushings [1]. Fig. 1 shows various metal rubber bushing parts.



Figure 1 Bushings [1]

The rubber metal bushings shown in Fig. 1 generally consist of a metal outer bushing, a metal inner bushing and a rubber material filling between the bushings. Rubber parts are produced in special molds in rubber presses with a method called vulcanization method. In rubber bushings, diameter process is applied by giving a plastic shape to the outer metal along the radius from the outside to the center. This diameter process can be done with symmetrically closing jaws, or it can be done by passing it through a conical mold. In Fig. 2, there is a visual representation of the diameter process [2].

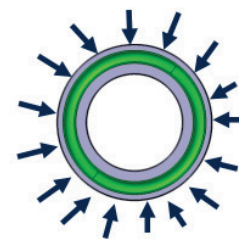


Figure 2 Diameter Processes [2]

When the production for rubber bushings is examined, the part coming out of the mold is left to the cooling process. During cooling, the rubber material in the regions close to the inner and outer bushings cools later and is pulled towards the edges. As a result of this shrinkage, internal stresses occur in the rubber material. For this reason, the produced bushings are produced slightly larger than the slot to be mounted, and the outer bushing is plastically deformed by the diameter process and the diameter value is reduced and mounted in the slot. After the diameter process, the outer bush may deteriorate due to errors such as circularity, rubber material behavior, insufficient pressure, mold error. The cases where this deterioration is ignored should only be within the tolerance values. In cases where the bushings are outside the tolerance values, the tight fit of the bushing surface cannot be ensured and the bushings come out of the body. For this reason, 100% control of the bushings is required after the diameter process.

1.1 Measuring

The measurements made before the measurement standards were created varied depending on the abilities of the person making the measurement, and there was no certain limit to these variations. In order to eliminate these differences, international measurement standards have been established. In order to ensure the conformity of the productions and to be within the tolerance limits, the measurements must be correct. The position of the part to be measured is one of the factors affecting the measurement.

Accordingly, the sensitivity of the measuring instrument and the environmental impact should be taken into account according to the limits of the parts to be measured. Circularity is the condition that all elements in the plane perpendicular to the axis of the part are equidistant from the axis of rotation. The condition that the elements in the plane perpendicular to the part axis are in two concentric circles within the tolerance distance is called the circularity tolerance [3].

2 MATERIAL AND METHOD

Within the scope of this study, firstly, the tolerance values of the bushings in the inventory were determined. Diameter, tolerance and product code values of the determined bushings are recorded. In order to provide modular design according to the bush variety in the recorded table, bushing positioning apparatuses in accordance with the bushing inner diameters were designed and produced. At the same time, the apparatus has been designed for the laser sensor in such a way that the height adjustment can be made without error according to the bushing inventory. Apparatuses made are exemplified in Fig. 3.

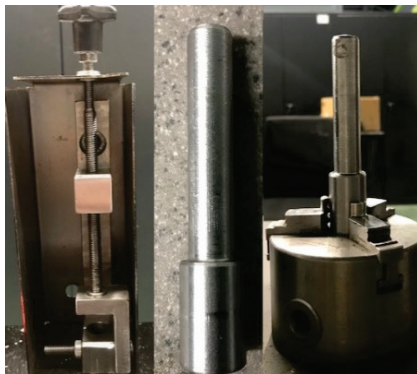


Figure 3 Apparatus [3]

2.1 Test System Preparation

Secondly, the preliminary design of the system, material selections, the laser sensor to be used, servo motor and motor driver, image processing interface, industrial type camera and PLC and HMI as controller were determined. Sensor sensitivity is selected according to bushing tolerance values. With deep learning and image processing algorithms, laser circularity measurements are verified with reference to the circularity of the bush, and the points where the circularity is corrupted or deformed are detected. The detected deformation is transferred to the screen as a faulty area. Tolerance values and bush diameter value specified in the calibration page must be entered. Afterwards, calibration should be done in order to use the test device and to obtain healthy measurements. The bushing inner diameter apparatus, which is selected according to the bushings to be tested, is fixed to the chuck. The bushing to be referenced is placed and at least 10 measurements are made by changing the direction. Measurements are made with the data collection button. After the data collection process is completed, the calibrate button is pressed. The recipe number

field on the main screen is filled. The calibrated recipe number for the bushing to be measured from the registered recipes is activated, the bushings are placed on the apparatus in order and the test process is started.



Figure 4 Machine Interface

If a bushing is not placed on the bushing inner diameter apparatus, the error screen displays a missing part warning and the green light does not appear. During the placement of the bush on the apparatus and the measurement, data is taken from both sensors. If there is a measurement that is not within the tolerance values, the red light representing that sensor lights up, and the inspection is repeated with the image processing algorithm. The lower and upper bases of the tolerance values are located next to the light that represents each sensor. If the measurement result is incorrect, the red warning lamp lights up and audible warning is given. The machine interface is shown in Fig. 4.

2.2 Measurement and Verification

The measurement system is as follows; A bushing, which is determined for calibration and has the correct dimensions, is taken and the bush is properly positioned in the measuring slot. Sampling data is collected from 10 different points randomly on the bush. The bushing tolerance value is entered and the calibration process is completed. These calibration data will be used in the verification of the bushing data to be measured, and the surface diameter values will be checked and compared with the results of image processing and deep learning algorithms. As a result of these comparisons, the instant diameter value of the bush and the surface points with diameter errors are detected and reflected on the screen. The verification made by image processing and deep learning method is shown in Fig. 5.

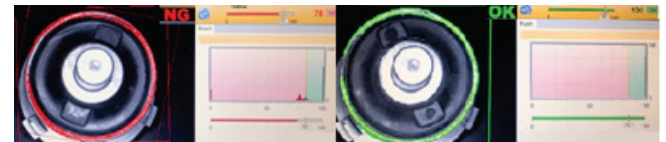


Figure 5 Control Process

3 RESEARCH RESULT AND DISCUSSION

With this study, circularity errors in rubber metal bushings are controlled, eliminating the problems that may be encountered after production and ensuring a quality

production, and also obtaining accurate results by taking the margin of error from the operator and offering a more precise solution. With 6 different recipe recording systems, bushing measurements in different tolerance ranges and different diameters have been accelerated. The distance values of the instantaneous laser values were measured for the accept bushing and reject bushings.

These measured values are shown in Fig. 6. For cases where the value difference between the center distance and the target distance values exceeds the tolerance range, the deformation zone was determined by comparing the values with deep learning and image processing algorithms.

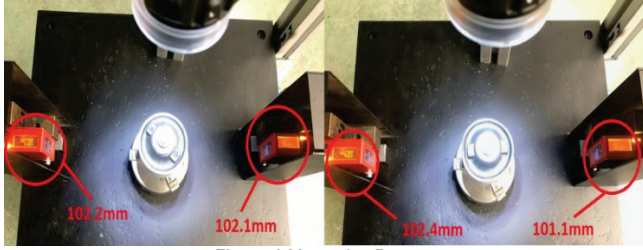


Figure 6 Measuring Process

3.1 System Measurement and Compliance Check

Measurement trials were carried out on a total of 450 bushings. As a result of these trials, there are 395 accepted and 55 rejected bushings after manual measurement. There are 406 accepted and 44 rejected bushings in the trials with system measurement.

When manually measured and rejected bushes are re-measured precisely, it is determined that they comply with the acceptance criteria and it has been determined that the system measurement works more stable than the manual measurement.

Table 1 Manual and System Measurements

	Ø20,5-25,5	Ø25,5-30,5	Ø30,5-35,5	Ø35,5-40,5	Ø40,5-45,5
Manuel Reject	27	15	9	1	3
Manual Accept	123	85	91	39	57
System Reject	22	12	9	0	1
System Accept	128	88	91	40	59

As can be seen in Tab. 1, the number of bushings with deformation detected has been reduced. In addition, deformation detection has been performed clearly for bushings where no deformation could be detected.

4 CONCLUSION

In this study, diameter measurements and bushing ovality approval processes were carried out after the diameter process in rubber metal bushes. With the automatic control system, the operator's initiative has been removed in control. The system seen in Fig. 7 can be controlled and monitored via HMI screen and SCADA.

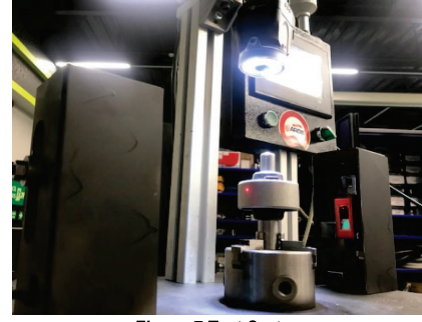


Figure 7 Test System

Acknowledgements

This work was supported by the project numbered AYD2101-06. In addition, this study was produced from the thesis of the first author. For their supports in the study, thanks to our advisor Mr. Prof. Dr. Cemil SUNGUR, AYD Otomotiv Endüstri A. Ş. R&D Team and AYD Otomotiv Endüstri A. S. R&D Manager Mr. Ahmet CAKAL.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

5 REFERENCES

- [1] Akgünoğlu, A. & Erzincanlı, F. (2019). Salıncak Kolu Burçlarının Çıkma Yükünü Etkileyen Parametrelerin İncelenmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 7(1), 729-747. (in Turkish) <https://doi.org/10.29130/dubited.432438>
- [2] Öncü, I. & Dova, A. (2018). Titreşim Sönümleme Elemanı Olan Kauçuk Burçta Çaplama İşleminin Ömre Etkisinin Sonlu Elemanlar Analizi ve Deneysel Çalışma ile İncelenmesi. *Journal of Engineering and Science*, 6(2), 11-18. (in Turkish) <https://doi.org/10.21541/apjes.364545>
- [3] Töre, C. (2005). *Geometrik Toleranslar*. Birleşim Yayıncılık, Ankara.
- [4] Hope, S., Conrado, G., Paula Pinto, A., & Quaqlia, J. (2003). Lower Control Arm Rear Bushing Development an Integrated Approach. *SAE Technical Paper*. <https://doi.org/10.4271/2003-01-3676>
- [5] Naiyi, L. & Peng, Z. (2014). Modeling of electric stress control in HV bushing using field grading material. *IEEE Conference on Electrical Insulation and Dielectric Phenomena*, 550-553. <https://doi.org/10.1109/CEIDP.2014.6995822>
- [6] Siwei, T. & Zou, J. (2021). Experimental study on fatigue performance of bushing repair with different fasteners. *Journal of Physics Conference Series*, 1885(3). IOP Publishing. <https://doi.org/10.1088/1742-6596/1885/3/032037>
- [7] Segura, M., Kanninen, M., & Suarez, D. (2006). Allometric models for estimating aboveground biomass of shade trees and coffee bushes grown together. *Agroforestry Systems*, 68(2), 143-150. <https://doi.org/10.1007/s10457-006-9005-x>

Authors' contacts:

Muhammed Abdullah Özel, Electrical Electronic Engineer
(Corresponding author)
AYD Automotive Industry Konya,
Vali İhsan Dede Cd. No: 7, 42280 Büyükkayacık Osb/Selçuklu/Konya, Turkey
+905073349721 / muhammedabdullahozel@gmail.com

Selim Sefa Baysal, Electrical Electronic Engineer
AYD Automotive Industry Konya,
Vali İhsan Dede Cd. No: 7, 42280 Büyükkayacık Osb/Selçuklu/Konya, Turkey
+905319557288 / ssbaysal@gmail.com

Cemil Sungur, Prof. Dr.
Konya Technical University,
Faculty of Engineering and Natural Sciences Electrical Electronics Engineering,
Akademi Neighborhood Yeni İstanbul Road No: 235/1, Selçuklu/Konya, Turkey
csungur@ktun.edu.tr

A New RF Satellite Link Analyzing and Antenna Effect on Satellite Communication

Bircan Çalışır*, Ayhan Akbal

Abstract: Satellite communications (SatComs) recently, have admitted as a new technological advances that attracted and raised special investment and ventures. Space communication technology has found many application areas from costly, one-of-a-sort structures, to utilize again of technology on sequential tasks, to the progress of canonical protocols accepted by space agencies of lots of countries. Satellite communication has accepted as one of the important technologies for 5G backhauling, particularly on bandwidth request increased in 5G mobile broad band (eMBB) applications. This paper presents the efficiency of the satellite communication system using 16 QAM digital modulation technique for the X band. This digital modulation method presents high data rate in transmission without raising the bandwidth compared with other digital modulation methods. Applications are realized in Matlab environment and obtained conclusions are discussed such as power spectrum diagrams, constellation schemas and BER ratio for several bands such as 4, 8, 12, 16 GHz.

Keywords: antenna arrays; satellite communications; 5G

1 INTRODUCTION

Satellite Communications (SatComs) have applications such as media broadcasting, backhauling, news gathering etc. areas. Recently via advancing of Internet basis applications, SatComs are passing a conversion stage to make centre again the system model on data services, known as broadband SatComs [1-4]. While especially the number of Low Earth Orbit (LEO) satellites are growth rapidly, ground basis communication systems have also evolved. The transmitting and receiving antennas are at the forefront of these systems. To make better the parameters and to optimize the performances, many antenna configurations have tested and continue to test in cases where it is necessary to follow specific conditions to make receiving from the satellite. [5-8]. Arrays of antennas are combine of groups of antennas adjusted to get aimed directivity properties. An array can formed with any combination of radiant elements but, in application usually classified equivalent elements in organized geometries are used. With progressing of satellite, communications, in order to make receiving or transmitting the signals as close as feasible to the space where the signals have needed, via not to disturb other spaces, the radiation model have been more and more precise. So, antenna arrays in phased [5, 7] have been highly improved. These antennas can alter the radiation model in array with respect to the phase shift that occurred among the elements supply voltages. This paper has organized as channel modelling, antennas, UHTS systems, data collection, application, and obtained results of this paper has explained in the conclusion part, respectively.

1.1 Channel Modelling

Characteristics of channel and propagation have dictating act in the designing of system. These statuses have originally defined by the frequency of the process besides the system configuration. Because the satellite communication systems use a large interval of frequency, results with various channel models. Fixed satellite and Mobile satellite are channel models that came across in satellite communications.

Since mobility means the existence of dispersed multipath constituents next to straight path, the channel characteristics of Mobile satellite systems are different from their Fixed Satellite equivalents. While following the satellite via the mobile station means the direct path, enhanced number of scatters in mobility caused variable environments, scattering from foliage, reflection from building defines the multipath term. In the literature, several of channel models such as narrowband and broadband have been suggested. In general, these models include a multiple condition Markov model, via every status defining the parameters and the character of the matching channel dispersion. As an example, a Markov model with 3 state narrowband interested state statistics have been presented by ITU Recommendation P.681 [9].

2 ANTENNAS

Antenna design is a necessary significant aspect to analyse. The principal properties of the passive and active antennas and highlight oncoming aspects in designing antenna for satellite communication systems will express in this section.

2.1 Passive Antennas

From radio transmission to broadband tasks signs a passing from contour ray coverage, which purposed serving a dedicated geographical area, to multi-beam antennas (MBAs) that by merit narrower ray makes possible both larger gains and use again of frequency, so spectral ability has maximized. GEO missions, which is the corresponding variation of conventional passive antenna structures, have directed the formed reflector antennas versus Multiple Feeds Per Beam (MFPB) MBAs [32] and Single Feed Per Beam (SFPB).

2.2 Active Antenna Arrays

Amplifiers of the active antennas integrate with the radiating elements. One of the differences between passive

antennas and active antennas, the obtained signal is a distributed amplification of the spreading signal in passive antennas. Distribution of the power of spatial RF and the decreasing of peak RF power levels supply reliability advance in active arrays (inclusive decreasing multi-process brinks) and ensure an elegant corruption. Structures, where Active antennas can place, are Direct Radiating Array (DRA) or Array Fed Reflector (AFR) structures [32]. This selection and the incorporated trade-offs are potentially related to the requirements of the system. When the wide area of sight joined with the decreased requests on gain preferred DRA solutions for LEO and MEO systems. The big electrical sizes necessary to reach the main aim of GEO HTS (High Throughput Satellite) missions are achieved preferentially with reflector basis structures that provide maintain an enlargement of the radiating interval, so there are still open questions about the best active array architecture [10-12].

3 ULTRA HIGH THROUGHPUT SATELLITES (UHTS)

In addition to numerical payloads, fore-coding techniques, and non-orthogonal multiple reach the main servers of the UHTS system are accession to the PHY layer view. Satellite communications have been obligated to satisfy the increment demands for the trustworthy and elastic links at higher yields due to the diversity of markets. The payload has accepted as a piece of the head on channel constantly and its attitude should measure continuously. In Orbit Test (IOT) process of the satellite payloads includes a specially modelled testing signal for the evaluation and subtracting of several important payload factors such as on board responses of filter answer of high power amplifier. The IOT is basic process throughout used the lifetime of the satellite to confirm and track the success and structural necessities of the satellite payload [13].

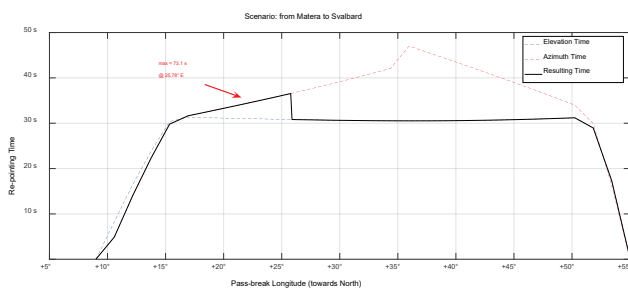


Figure 1 Re-pointing Time to Pass-break Longitude for Scenario: Svalbard to Matera

3.1 Data Collection

Collecting data from Earth or from space is also a significant subject worthy of taking into account. Satellite communication in commercial applications, such as observation and LEO satellites, stand on exceedingly available great rates in data along a short transition of the satellite. Thus, new station modems talented of the efficient process are developed at much great symbol rates recently. If the whole Ka band is used, the main challenge for the

designing modem, outcomes from a much wide signal spectrum that can raise up to 1.5 GHz. These days, the proposed terminal modems for commercial high data rates support up to 500 MHz [14-16].

Elevation and azimuth time are important parameters in antenna design. In our application there is two scenario to be evaluated. In Fig. 1 for pass break longitude is between +5 to +55 degree, maximum re-pointing time 73.1 s is obtained at 25.78 degree. For pass-break longitude is between +170 to 145 degree, maximum re-pointing time 84. s is obtained at 163.62 degree in Fig. 2.

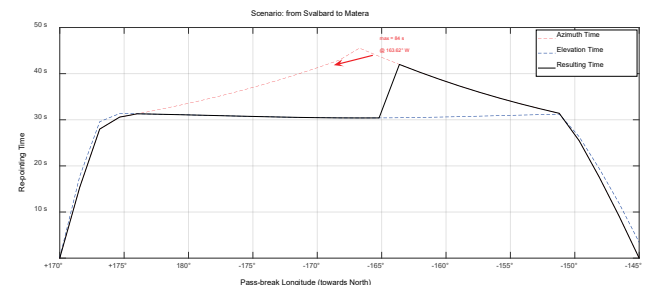


Figure 2 Re-pointing Time to Pass-break Longitude for Scenario: Matera to Svalbard

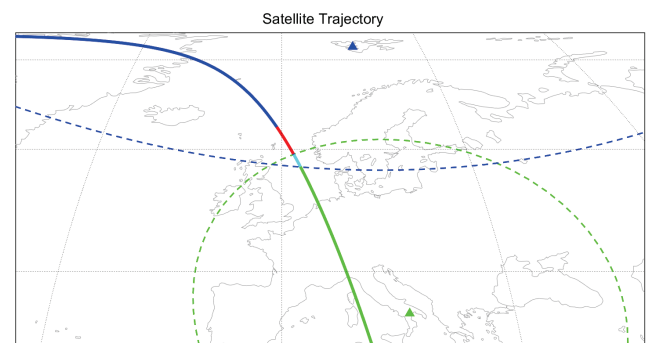


Figure 3 Satellite trajectory for Scenario: Svalbard to Matera

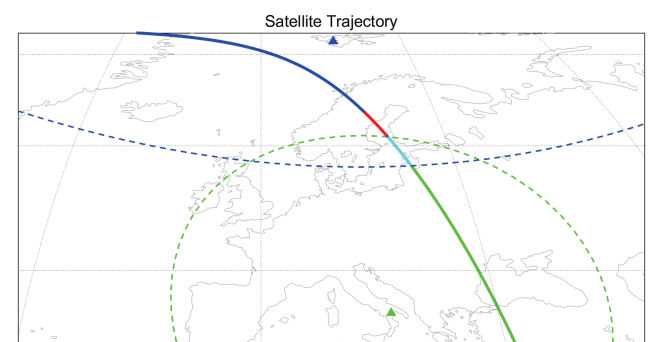


Figure 4 Satellite trajectory for Scenario: Matera to Svalbard

3.2 5G Networks

5G has purposed more than alteration of the former norms, to make please significant department of the market, such as the automotive and transporting lines, broadcasting and demonstration, industry. For 2020 and beyond, ITU-R International Mobile Telecommunications has defined several great structures for the enhancement of 5G [8]. Thus,

the role of satellites that is crucial in the 5G ecosystem has largely acknowledged.

Satellite trajectory plotting for Scenario: Svalbard to Matera has given with Fig. 3. Satellite trajectory plotting for Scenario: Matera to Svalbard has given with Fig. 4.

Recently, mid-layers of transmissions systems have emerged between terrestrial and conventional satellite slices via the technological progress of the air and miniature satellite structures. According to their operation elevation, these new structures can be categorized into three major groups have allocated High Altitude Platforms (HAPs) satellites, Low Altitude Platforms (LAPs) satellites and Very Low Earth Orbit (VLEO) satellites [33]. Elevation sets are admitted [17, 18] 100-450 km for VLEO satellites, 15-25 km for HAP satellites, and 0-4 km for LAP satellites respectively. A recent multi-floor transmissions structure [19] with multiple inter-floor linkages able to cope with the toughest scripts has enabled by the ad-vent of these new structures. The schematic approximation of this new multi-layer transmissions system has shown in Fig. 5.

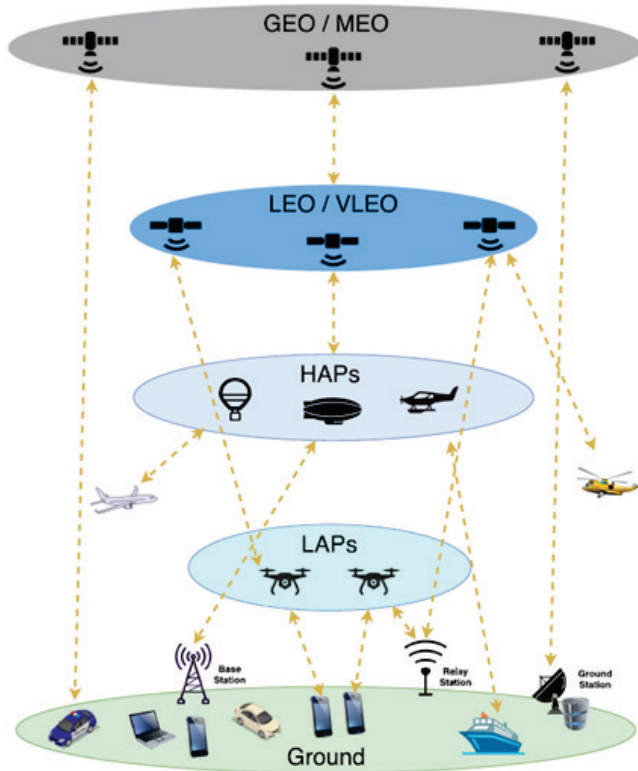


Figure 5 Multi-layer communications architecture

LEO satellites work more intimately with the Earth according to LEO satellites. For this reason, VLEO satellites are simple, small and cheap according to LEO satellites [17]. However, lower altitudes like VLEO cover a more intense section of the atmosphere, and so, larger aerodynamical powers. Although this has seen as a tough, but this can also point out a chance in check of trajectory and height [20]. In addition, in VLEO satellites the raised dragging indicates a shortening of the trajectory life span, which means a denser shipping exchange of smaller and cheaper space craft. At the

end of this situation, VLEO satellites become appropriate more sensitive to technology and changes in the bazaar [21]. Varied particular firms such as Telesat, SpaceX and OneWeb are purposing to initiate their Mobile Satellite Services (MSS) at VLEO. HAPs have features to supplement traditional satellite networks. In fact, they also acknowledged as High Altitude Pseudo Satellites [22, 23]. HAPs have the ability to supply telecommunications duties at a local scale be-cause of their working altitude. The most evident example of LAPs is Unmanned Aerial Vehicles (UAVs). A significant component of the immediate future wire-less networks is expected to be UAVs which can make ease wireless broadcast and assist high rate transfer potentially [24, 25]. UAVs (LAPs) have similarities to the HAPs ones in terms of the main advantages, but on the cellular plane: LAPs have fast and elastic propagation specifications, powerful line-of-sight (LoS) link, and independence in design with the tested mobility and autonomous. In addition, atmospheric base stations working with UAV can install, improve, and rescue coverage of cellular as real time for consumers in far away, crowded and catastrophic regions.

Satellite communications utilize the Extremely High Frequency (EHF) band, in particular among 1-50 GHz. In varied air circumstance, service and users types operate on different frequency bands. The frequency bands utilized for satellites defined by straightforward characters for ease. For example lower frequencies are determined as *L*, *S*, *X* and *C*, and larger frequencies are determined as *Ku*, *K*, *Ka*, *Q/V*. A diagrammatic demonstration spectrum of the satellite is shown in Fig. 6.

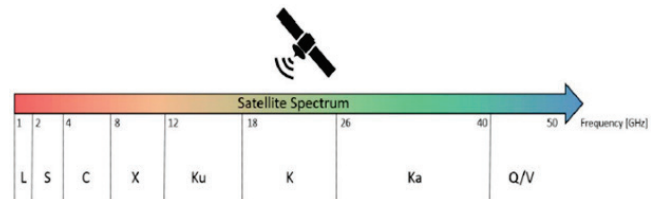


Figure 6 Satellite spectrum

In the *L*-band, navigation systems of radio, like GPS or Galileo work. Radar of weather and surface ship, and several satellites, and particularly satellites of NASA for transmission with the International Space Station (ISS) and the Space Shuttle [26, 33] use the *S*-band. In addition, *L* and *S* bands have utilized for Telemetry, Tracking and Control (TT&C). Especially, the space search, space process, and EO satellite services share the frequency bands among 2-2.3 GHz equivalently [27]. Since in the lower bands, there is not much bandwidth available so it has become costly work. Spatial variation of the traffic over the satellite orbit is practiced in MEO and LEO satellites. Therefore, there is a requirement for a reconfigurable antenna to make equal the coverage of satellite with the spatial spreading of the receiving and transmitting traffic. Telesat, Starlink and Akash have a compatible approach with the LEO constellations [28]. O3B admits a controllable beam approximation for the MEO constellation [29]. For small satellite structures such as CubeSats a review of antenna solutions can be found in [30].

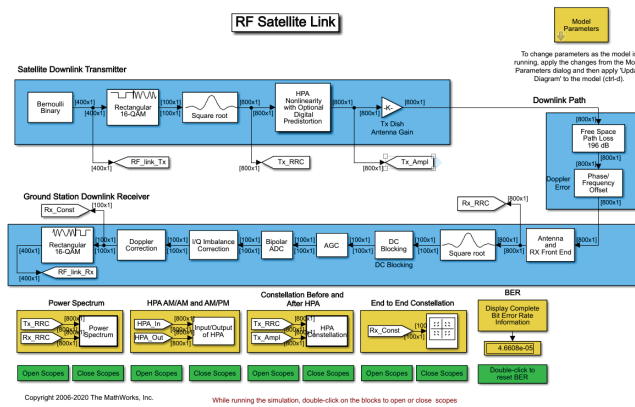


Figure 7 RF Satellite Link Scheme [32]

4 SIMULATION RESULTS

The model in Fig. 7 shows a satellite link to simulate several impairments such as nonlinearity of memoryless, path loss of free space, error of Doppler, thermal noise of receiver, noise in phase, instability in-phase and quadrature, and DC offsets [32]. The satellite connection sample and its signal areas have shown by the model, which includes Satellite Downlink Transmitter, Downlink Path, and Ground Station Downlink Receiver. This architecture models the gains and losses on the connection and calculates link budget to define bit error rate (BER). The Free Space Path Loss block and the Receiver Thermal Noise that named as the gain and loss blocks, specify to be supported data rate in an additive white Gaussian noise channel. In block named as Satellite Downlink Transmitter; Bernoulli Binary Generator forms binary data flow randomly, Rectangular QAM Modulator Baseband plans the constellation diagram of 16-QAM data stream. The signal modulated utilization the SRRC (Square Root Raised Cosine) pulse shape is up-sampled and developed by Transmit Filter of Raised Cosine. A mobile wave tube amplifier utilization the Saleh pattern choice of the Memoryless Nonlinearity is modelled by HPA (High Power Amplifier) Nonlinearity with Optional Digital Predistortion and selectively the AM to AM and AM to PM are corrected with a Digital Predistortion block. Tx Dish Antenna Gain known as gain block utilize gain of the parabolic dish antenna in the transmitter. Free space path loss reduces the signal by the free space path loss in Downlink Path. The signal is rotated by Phase/Frequency Offset (Doppler Error) to define Doppler error on the connection. Rx Dish Antenna Gain known as Gain block performs gain of the receiver parabolic dish antenna In Ground Station Downlink Receiver. White Gaussian noise that indicates the active temperature of the system in the receiver is added using the Receiver Thermal Noise block. Random phase corruption that outcomes from $1/f$ or phase chirps noise is presented using Phase Noise. The block of I/Q Imbalance Correction brings to the signal several impairments such as DC offset, amplitude or phase non-stabilization. In short known as LNA, Low Noise Amplifier supplies low noise amplifier gain. A corresponding filter to the modulated signal has obtained using Raised Cosine Receive Filter. DC Blocking block eliminates the DC offset in I/Q Imbalance

block. AGC block has used to adjust the desired level signal power. I/Q Imbalance Compensator has used for prediction and eliminating I/Q imbalance from the signal using an algorithm that presents solutions adaptively. Doppler Correction block and Carrier Synchronizer blocks together are using to eliminate the carrier frequency offset emerging due to Doppler Effect. Rectangular 16-QAM block is using to remap the data stream from constellation space.

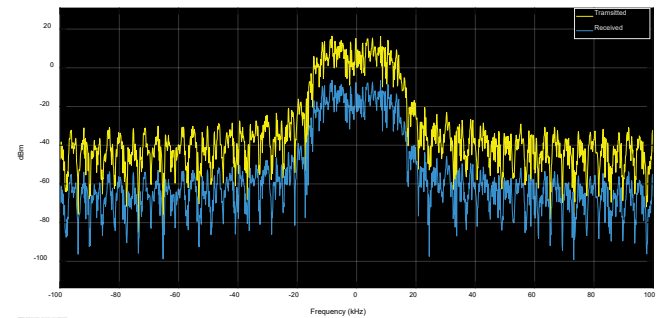


Figure 8 Power spectral density of transmitted (yellow) and received (blue) signal for RF Satellite Link [32]

Referring to Fig. 8 the main difference of the noise level among power spectral density of transmitted and received signals observed at 20 kHz is -6 dB approximately.

When compare the two spectrum the impact of the some RF corruptions are encountered such as regrowth in spectrum emerged due to HPA non-stabilities raised by the Memoryless Nonlinearity block, thermal noise occurred by the Receiver Thermal Noise block and Phase chirps occurred by the Phase Noise block. In Fig. 8 the value of Thermal Noise is selected 20 that is low noise.

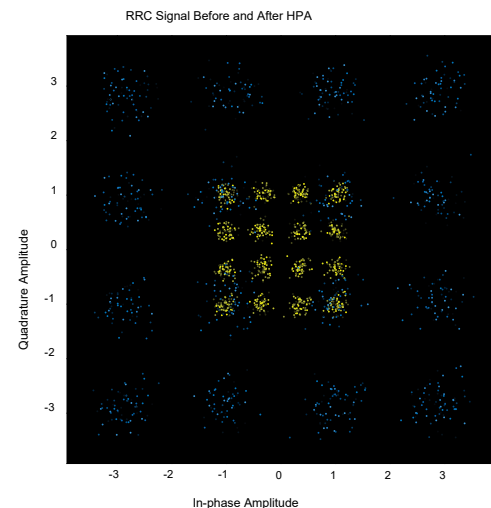


Figure 9 Modulated (yellow) and received (blue) signal for RF Satellite Link

Fig. 9 allows us to check the constellation diagram of the transmitter signal with yellow and after the HPA with blue. Output signal of the HPA is larger than input signal of the HPA due to the amplifier gain. This figure facilitates us to observe the united impact of both non-stability and digital pre-defect of the HPA. As seen in the constellation diagram

in Fig. 9 the received signal has affected by noise compared to the transmitted signal.

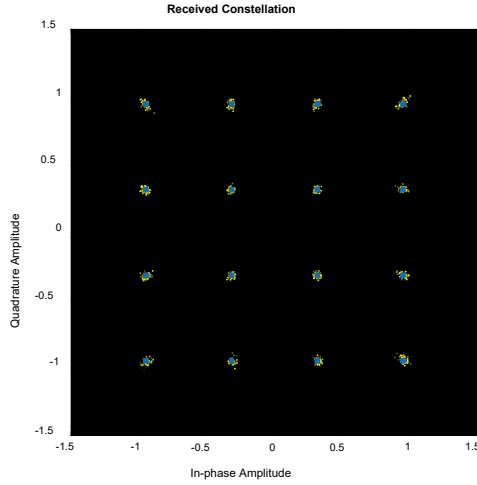


Figure 10 Received constellation diagram

Fig. 10 allows us to compare the 16-QAM reference constellation diagram (blue) with the received QAM constellation before demodulation (yellow). The effect over all the RF corruptions on the received signal and the influence of the compensations have observed by using these constellation diagrams. Satellite height (km) has defined as the margin among the satellite and the ground station. Modification of this parameter renovates the Free Space Path Loss block, which the default value has selected as 35600 [31]. Frequency defines the link carrier frequency, changing of this parameter renovates the Free Space Path Loss block, which the default value has chosen as 4000 kHz. Diameters of transmit and receive antenna are the first elements in the vector and are utilized to compute the gain of a block of the Tx Dish Antenna Gain. The other component expresses the receive antenna and is utilized to compute the gain in block of the Rx Dish Antenna Gain. Noise temperature lets us to choose from noise temperatures in four active receiver systems. The chosen temperature of noise alters the Noise Temperature of the Receiver Thermal Noise block.

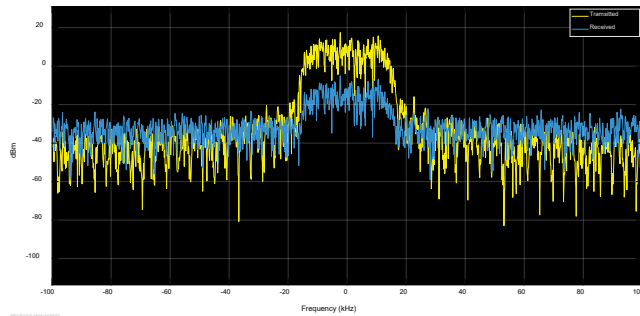


Figure 11 Power spectral density for Noise Temperature 290 K

The effect of chosen temperature noise 290 K is shown by Fig. 11, 12 and 13 for power spectral density, RRC signal before and after HPA and received constellation. In addition, Fig. 14, 15 and 16 present the effect of chosen temperature

noise 500 K for power spectral density, RRC signal before and after HPA and received constellation.

Referring to Fig. 11 the key difference of the noise level among power spectral density of transmitted and received signals monitored at 20 kHz is approximately -6 dB.

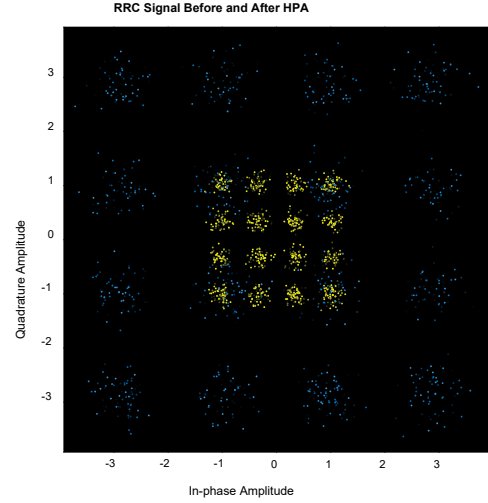


Figure 12 RRC signal for Noise Temperature 290 K

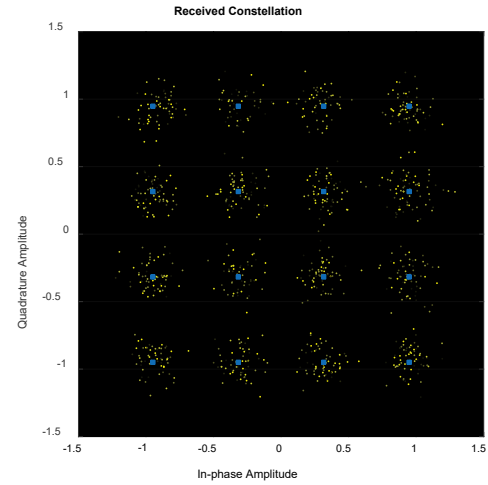


Figure 13 Received constellation for Noise Temperature 290 K

As seen in the constellation diagram in Fig. 12 the received signal has effected by noise compared to the transmitted signal.

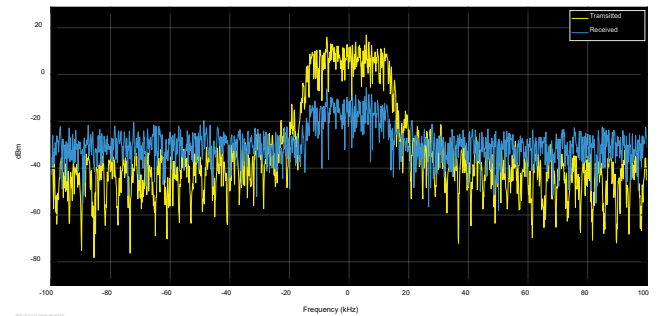


Figure 14 Power spectral density for Noise Temperature 500 K

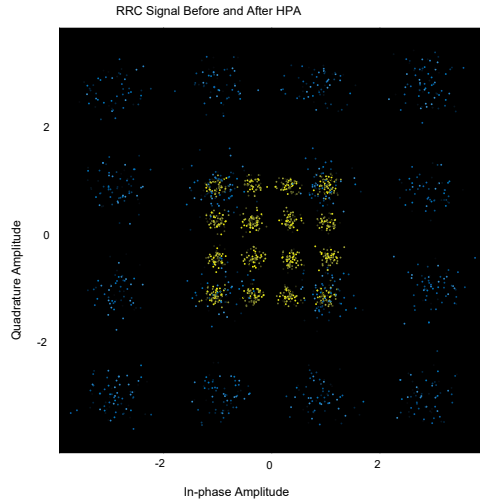


Figure 15 RRC signal for Noise Temperature 500 K

When the noise performance is compared, the constellation diagram of the signal at the output of the HPA is better than at the input of HPA.

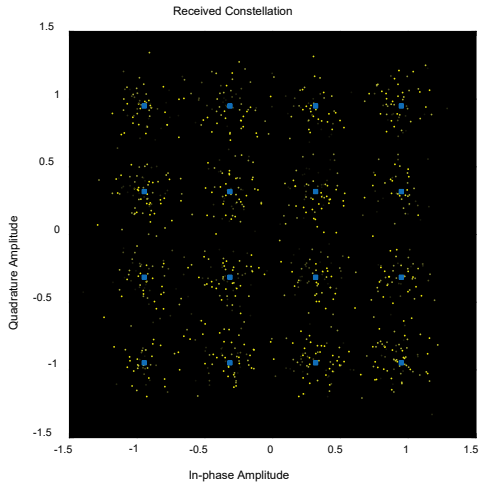


Figure 16 Received constellation for Noise Temperature 500 K

In Tab. 1 Variation of BER according to Carrier Frequency is reported. From this table, it is seen obvious the value of BER increases as the carrier frequency increase.

Table 1 Variation of BER according to Carrier Frequency

Frequency Band	BER
4 GHz	1.39e-6
8 GHz	2.99e-6
12 GHz	5.98e-6
16 GHz	2.98e-5

In this article, the ultimate technical progress in academic, industrial, and standardization investigations in the satellite communications domain have been examined. Particularly, the titles of channel modelling of satellite communication, active and passive antennas, UHTS systems, data collection, and layers of communications systems that have emerged between terrestrial and conventional satellites are explained in detail. After all the explanations,

applications realized and results have given with figures. Especially RF satellite link transmitter and receiver blocks have expressed separately.

In this paper, an RF satellite system using 16 QAM digital modulation technology is proposed, and the performance of the satellite system is explored and examined. The purposed model was implemented for various noise temperatures and obtained results such as power spectral density, constellation diagram. In addition, it was seen that the value of BER increases as the carrier frequency increase obtained as a result of simulations.

One of the future applications is named delay-tolerant networking (DTN) which should get knowledge from one area of the Solar System to another one. The DTN structure enables uniform communications over extended distances and delays in time. In fact here that there is the Bundle Protocol (BP), which has similarities to the Internet Protocol (IP) that is known as the base of the Internet on Earth [33]. Some researchers have been now studying to improve a recent link-floor structure for the Space Internet but have not observed applicable improvement yet.

Notice

The paper was presented at the International Congress of Electrical and Computer Engineering (ICECENG'22), which took place in Bandırma (Turkey), on February 9-12, 2022. The paper will not be published anywhere else.

5 REFERENCES

- [1] Sturza, M. (1995). The teledesic satellite system: overview and design trades. *Teledesic Corp.*, Kirkland, WA, USA.
- [2] https://en.wikipedia.org/wiki/SpaceX_Starlink, accessed 2021/11/12.
- [3] Antipolis, S. (2019). Technical specification group radio access network; study on new radio (NR) to support non-terrestrial networks; (release 15), version 15.2.0, 3GPP, France, Rep. 38.811.
- [4] Antipolis, S. (2019). Technical specification group radio access network solutions for NR to support non-terrestrial networks; (release 16), version 0.9.0, 3GPP, Sophia Antipolis, France, Rep. 38.821.
- [5] Bruninga, B. (2019). Antennas for APRS satellite Igates, <http://www.aprs.org/aprs-satellite-igate-antennas.html>
- [6] Done, A.: Contributions to the development of ground stations for satellite reception. PhD Thesis, "Ștefan cel Mare" University of Suceava, Suceava.
- [7] Lesanu, C. E., Done, A., Cailean, A. M., & Graur, A. (2018). Vertical polarized antennas for low VHF radio meteor detection. *14th International Conference on Development and Applications Systems*, Suceava, Romania. <https://doi.org/10.1109/DAAS.2018.8396078>
- [8] Stutzman, W. L. & Thiele, G. A. (2013). *Antenna theory and design*. 3rd New York: Wiley.
- [9] ITU. (2003). Propagation data required for the design of earth space land mobile telecommunication systems. ITU, Geneva, Switzerland, ITU Recommendation P. 681-6.
- [10] Rao, S. K. (2015). Advanced antenna technologies for satellite communications payloads. *IEEE Transactions Antennas Propagation*, 63(4), 1205-1217. <https://doi.org/10.1109/TAP.2015.2391283>

- [11] Palacin, B. et al. (2017). Multibeam antennas for very high throughput satellites in Europe: Technologies and trends. *11th Eur. Conf. Antennas Propag. (EuCAP)*, 2413-2417. <https://doi.org/10.23919/EuCAP.2017.7928493>
- [12] Fenech, H., Amos, S., & Waterfield, T. (2016). The role of array antennas in commercial telecommunication satellites. *Proc. 10th Europe Conference Antennas Propagation, (EuCAP)*, 1-4. <https://doi.org/10.1109/EuCAP.2016.7481857>
- [13] Perez-Neira, A. I., Vazquez, M. A., Shankar, M. R. B., Maleki, S., & Chatzinotas S. (2019). Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios. *IEEE Signal Processing Magazine*, 36(4), 112-131. <https://doi.org/10.1109/MSP.2019.2894391>
- [14] Kodheli, O., Andrenacci, S., Maturo, N., Chatzinotas, S., & Zimmer, F. (2018). Resource allocation approach for differential Doppler reduction in NB-IoT over LEO satellite. *9th Advanced Satellite Multimedia Systems Conference, 15th Signal Processing Space Communication Workshop, (ASMS/SPSC)*, 1-8. <https://doi.org/10.1109/ASMS-SPSC.2018.8510724>
- [15] Hofmann, C. A. & Knopp, A. (2019). Direct access to satellites: An Internet of remote things technology, *IEEE 2nd 5G World Forum (5GWF)*, 578-583. <https://doi.org/10.1109/5GWF.2019.8911640>
- [16] Hofmann, C. A. & Knopp, A. (2019). Ultra narrowband waveform for IoT direct random multiple access to satellites. *IEEE Internet Things Journal*, 6(6), 10134-10149. <https://doi.org/10.1109/JIOT.2019.2935909>
- [17] Llop, J. V., Roberts, P. K., Hao, Z., Tomas, L., & Beauplet, V. (2014). Very low earth orbit mission concepts for earth observation: Benefits and challenges. *Reinventing Space Conference*, 1441-1444.
- [18] Reynaud L. & Rasheed, T. (2012). Deployable aerial communication networks: Challenges for futuristic applications. *9th ACM Symposium Performance Evaluation Wireless and Hoc Sensor Ubiquitous Networking (PE-WASUN)*, 9-16. <https://doi.org/10.1145/2387027.2387030>
- [19] Li, Y., Wei, H., Li, L., Han, Y., Zhou, J., & Zhou, W. (2019). An extensible multi-layer architecture model based on LEO MSS and performance analysis. *IEEE 90th Vehicular Technology Conference (VTC-Fall)*, 1-6. <https://doi.org/10.1109/VTCFall.2019.8891397>
- [20] Virgili-Llop, J., Polat, H. C., & Romano, M.: Attitude stabilization of spacecraft in very low earth orbit by center of mass shifting. *Front. Robot. AI*, 6(7). <https://doi.org/10.3389/frobt.2019.00007>
- [21] Dakka, S. M. (2018). VLEO satellites a new earth observation space systems commercial and business model. *36th Int. Communication Satellite System Conference (ICSSC)*, 1-11. <https://doi.org/10.1049/cp.2018.1718>
- [22] Iskandar, M. A., Wibisono, S., Priatna, T. & Rachmana, N. (2017). On the design and development of flying BTS system using balloon for remote area communication. *The 11th International Conference Telecommunication System Services Applications*, 1-5. <https://doi.org/10.1109/TSSA.2017.8272902>
- [23] Qiantori, A., Sutiono, A. B., Hariyanto, H., Suwa H. & Ohta, T. (2010). An emergency medical communications system by low altitude platform at the early stages of a Natural Disaster in Indonesia. *Journal of Medical Systems*, 36, 41-52. <https://doi.org/10.1007/s10916-010-9444-9>
- [24] Li, B., Fei, Z., & Zhang, Y. (2019). UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet Things Journal*, 6(2), 2241-2263. <https://doi.org/10.1109/JIOT.2018.2887086>
- [25] 5G-SKY: Interconnecting the Sky in 5G and Beyond a Joint Communication and Control Approach (2019), https://www.fr.uni.lu/snt/research/sigcom/projects/5g_sky_interconnecting_the_sky_in_5g_and_beyond_a_joint_communication_and_control_approach, accessed 2021/11/30.
- [26] ESOA-EMEA Satellite Operator Association. Satellite Spectrum. <https://www.esoa.net/spectrum/satellite-spectrum.asp>, accessed 2022/01/01
- [27] Li, Y., Wang, Z., & Tan, W. (2013). The frequency spectrum management for aerospace TT & C system. *5th IEEE International Symposium Microwave Antenna Propagation EMC Technology Wireless Communication*, 595-600. <https://doi.org/10.1109/MAPE.2013.6689909>
- [28] Tubau, S., Vidal, F., Legay, H., Palacin, B., & Goussetis, G. (2019). Novel multiple beam antenna farms for megaconstellations. *40th ESA Antenna Workshop (ESA-ESTEC)*, Noordwijk, The Netherlands, 2-31.
- [29] https://www.itu.int/dms_pub/itu-r/md/12/iturka.band/c/R12-ITURKA.BAND-C-0010!!PDF-E.pdf, accessed 2021/11/25.
- [30] Gao, S., Rahmat Samii, Y., Hodges, R. E., & Yang, X. (2018). Advanced antennas for small satellites. *IEEE*, 106(3), 391-403. <https://doi.org/10.1109/JPROC.2018.2804664>
- [31] Alshepane S. H., Alhaddar K. F., & Bishi H. (2021). Investigation the Performance of RF Satellite Link Model Using 32 QAM. *2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences a Techniques of Automatic Control and Computer Engineering MISTA*. <https://doi.org/10.1109/MI-STA52233.2021.9464443>
- [32] MathWorks, UK. (2022). RF Satellite Link - MATLAB & Simulink - MathWorks, United Kingdom. <https://uk.mathworks.com/help/comm/ug/rf-satellite-link.html>
- [33] Kodheli, O. et al. (2021). Satellite Communications in the New Space Era: A Survey and Future Challenges. *IEEE Communications Surveys & Tutorials*, 23(1), 70-109. <https://doi.org/10.1109/COMST.2020.3028247>

Authors' contacts:

Bircan Çalışır, Research Assistant
(Corresponding author)
Firat University,
Department of Electrical and Electronic Engineering,
23119 Merkez/Elazığ, Turkey
05437368234, bkamislioglu@firat.edu.tr

Ayhan Akbal, Dr.
Firat University,
Department of Electrical and Electronic Engineering,
23119 Merkez/Elazığ, Turkey
ayhan_akbal@firat.edu.tr
ayhanakbal@gmail.com

Prioritizing the Researchers' Important Competencies in Research Centres

Ahmad Farmahini Farahani

Abstract: Changes caused by the knowledge economy, including the emergence of new idea flows in management, methods and structure of organizations, have led to a change in the roles and skills needed for researchers in organizations. As new age organizations focus on intellectual property, organizational aspirations and organizational change, the researchers, as the wealth creators, in order to quickly adapt to new situations and develop their competencies in the competitive market, need to constantly change and develop a new identity for themselves. Since competencies have a prudential feature through describing skills and behavioural approaches, identifying and explaining researchers' competencies in research centres is of particular importance. Accordingly, the present paper seeks to identify the factors and indicators of researchers' competencies in research centres using scientific methods and surveys and then identify, classify, and prioritize researchers' important competencies using statistical methods. According to the results obtained from the present study, creativity and innovation, integration, accountability and customer orientation competencies have higher priorities; however, all identified important competencies have a significant positive distance to mean. With the help of the results of this study, researchers and managers can clarify expectations about each other.

Keywords: important competencies; industrial research centres; researchers' competencies prioritization

1 INTRODUCTION

Because competencies have a prudential feature, and describe the skills and behaviours that researchers and research managers in the research centers needed to create a new culture and skills to meet future challenges, defining and explaining the researchers' competencies is one of particular importance both for research Centres and researchers in research centers.

Clarifying the expectations of research centers and explaining the research activities requirements to assist the future development of research centers regarding researchers' engagement and developmental projects are among the benefits of identifying competencies and determining their importance in this industry.

The researchers' important competences in the research centers can be seen as a robust basis for establishing consistent performance standards, which are formed by a common language about what is needed and expected by research centers [1].

Because of applying the results of its findings to specific and intra-organizational problems, and for addressing the researchers' important competencies (including the components and dimensions of competencies), the present study is considered as applied research. The research is also considered as a descriptive-survey in terms of the data collecting approach.

The present study also considers the research competencies, not the management ones, and other non-executive competencies, not technical and specialized ones, so this institute is selected as the statistical population.

Library studies and surveys, interviewing relevant academic and research experts, Delphi method, examining related models, questionnaires, analyzing findings based on mathematical models, and statistics, competency software, SPSS, Excel and Lisrel software and other future studies methods have been used in the present study and in accordance with the procedure shown in the Fig. 1.



Figure 1 The research procedure

2 LITERATURE REVIEW

2.1 Definitions Competence

The ability to utilize or apply a set of relevant knowledge, skills, and abilities needed to successfully execute "critical aspects of work" or tasks defined in a work environment. This should not be confused with qualification, but competence describes a behaviour rather than trying to describe the performance level.

Competencies often considers as a basis for skill standards that provide a "level" of knowledge, skills, and competence needed to succeed in the workplace as well as a potential indicator to evaluate the competence achievement.

2.2 Competency Model

One of the challenges facing enabling an organization to compete effectively is adapting employees to their job needs. Using job analysis is a traditional approach to formulating job performance models, in which different procedures are used to identify and describe the job critical requirements and differentiate the high and low performance [18]. Job analysis focuses on issues such as what employees are really doing in their jobs, perceptions of what people are doing and what should be done in the future [8]. Two problems may arise when using job analysis to describe and describe jobs. Firstly, the description of traditional written jobs is only for describing the activity and may not clearly explain the outcomes needed for the organization to succeed; and the second is to describe traditional jobs rapidly in today's fast-changing and dynamic environment [6]. Criticism of job analysis indicates that performance models are very specific to jobs and are more focused on the job rather than the individuals and values and goals of the organization. Hence,

a relatively new approach is proposed to define and determine the performance called the competence model [13].

Competency models often contain a brief description of jobs, values, and key organizational goals for preparing a list or selecting from an existing list of the most important competencies required for performance that align with organizational values [10].

The competency model is based on the approach that competencies are defined as effective and visible behaviours that are tied to organizational outputs [15]. 'Introducing the competencies needed for senior executives in key positions in the organization', 'trying to eliminate any gap from competence through effective selection', 'training and development', 'ensuring recognition and reward for good performance' are among the goals of the competency-based approach [5].

According to Afkhami et al. [1], by identifying and emphasizing employees' skills that coordinate and support core strategies and competencies, the competency model approach can equip organizations to the competitive advantage [22].

A complete picture of job requirements is provided in competency models that will increase the chances of those who will be hired, and ensure about a systematic and orderly interview process. It can also help to distinguish between competencies that improve with training and those that are more difficult to improve [24]. A set of competencies is also defined in the competence model, by definitions and descriptions of behavior for a particular or a group of jobs and it is a narrative description of the competencies needed for a job or the job category that results from observing the employees' satisfying or phenomenal performance. A competency model identifies a set of competencies and behaviors, directly related to the work to be done, and the skill levels for each behaviour [3].

As the ability to perform activities within the work domain according to established standards, the competency model is in fact a measurable model of the abilities and behaviors needed to perform the job successfully [9].

The competency model is at the heart of organizational performance because it guides learning and training processes and strategies, content development, and performance evaluation [17]. In order to develop competencies, individuals need to know both the knowledge and opportunities for learning skills in practice.

The competency model is a certain behavioral description of skills and features needed by employees to be effective in a given job and it is also useful to determine the required abilities and features to meet current and future employees' requirements and focus on employees' development efforts to eliminate the gap between the available and needed capabilities [16].

The competency model is a detailed description of the competencies needed for the full success or an exemplary performance in a job, role or organization [23]. It is a decision-making tool in the human resource management system in the processes of recruitment, selection, training and development, performance management and succession

planning in order to meet the organizational requirements including a list of competencies, definitions, and behavioral indicators [4].

In addition, as it translates business requirements to training and operational needs for HR professionals, the competency model in many organizations helps leaders to develop within the organization to enhance the organization's performance from good to excellent and to provide a clear picture of what managers need to be successful in their careers. For many researchers, the competency model is the basis of any human resource system. According to [12], formulating a competency model has a positive relationship with individual and organizational performance. The competencies required for each managerial position are identified when formulating the competency model, and consequently, the organization can appoint managers who are more likely to have optimal performance [2].

2.3 Reviewing Competency Models

The competency models are discussed in two parts in the following: the general models and the models designed and used for research centers.

2.3.1 General Models

2.3.1.1 Beyham's et al. Model

Introduced in 2002, this model has four dimensions interpersonal skills, leadership skills, management or business skills, and personality traits. The characteristics of this model are listed in Tab. 1 [11].

Table 1 Beyham's et al. model

Index	Dimension	Index	Dimension
Entrepreneurship	Business management	Effective communication	Interpersonal skills
Business Frost		Effectiveness of Interpersonal Culture	
Important orientation		Customer-orientation	
Global intelligence		Developing important relationships	
Job handling		Persuasiveness	
Equipping Resources		Creating organizational talent	Leadership
Operational decision making	Personality features	Leadership for Change	
Proper self-control		Coaching and Delegation	
adaptation		Influencing others	
Result-orientation		Team development	
energy		Expert	
Learning-orientation			
optimist			
Performance oriented			
Understanding the environment			
Value for diversity			

2.3.1.2 Zenger & Folkman Model

The leadership competencies are classified in five dimensions in Zenger and Folkman [23] that to differentiate the organization: personality, personality capacity, result-orientation, interpersonal skills, and organizational change leadership. The model characteristics are listed in Tab. 2 [19].

Table 2 Zenger & Folkman model

Index	Dimension	Index	Dimension
Communications	Inter-personal skills	High integrity and loyalty	Personality
Motivating others to perform well		Technical and professional expertise	Personality capacity
Building relationships		Problem analysis and solution	
Being a team and working as a team		Innovation	
Partnership and Collaboration		Pay attention to own development	Results-orientation
Developing important vision and vision	Organizational change leadership	Result-orientation	
Champion of Change		Formulating flexible goals	
Creating internal group contacts with the organization's external environment		Responsiveness about results	

2.3.1.3 Boyatzis Model

The Management Association in the US delegated competency modelling research to Mac Company. The research was resulted to a model presented by Boyatzis in 1982 as follows. He has divided management competencies into six dimensions, including goal and operations management, leadership, human resource management, marketing leadership, focus on others, and specialized knowledge which indicators are presented in Tab. 3 [14].

Table 3 Boyatzis model [4]

Dimensions	Index	Dimensions	Index
Objective and Operation Management	Efficiency-orientation	Staff management	Using the power of command
	Being passive		Developing people
	Diagnostic use of concepts	Focus on others	Perception of reality
	Aware of the consequences of achieving the goal		Self-control
leadership	Self Confidence		Endurance
	Verbal presentation	Specialized knowledge	adaptation
	Perception		Specialized knowledge
Human Resource Management	Social power		
	Group management process		

2.3.1.4 Stone's Management Competency Model

This model, which is presented in Tab. 4, was formulated in 2001 and categorizes competencies into six dimensions: communication, development and partnership, core practice, technical expertise, organizational effectiveness, and personal effectiveness.

Table 4 Stone's Management Competency Model

Index	Dimension	Index	Dimension
Verbal communication	Communications	Building relationships	Organizational effectiveness
Written communication		Public Relations	
Conflict management		Evaluation and Accountability	
Introversion		Organizational review	
Educational Skills		Resource Management	
Presentation skills		Balancing work and personal life	Personal effectiveness
Teamwork and partnership	People Development and participation	Flexibility	
Facilitate teamwork		Individual education and personal learning	
Volunteer development and management		Job Management	
Control and coaching		Result-orientation	Action-orientation
Delegation of power		Change implementation	
Customer-orientation			
Technical expertise	Technical expertise		
Technology integration			
Problem solving			

2.3.1.5 Mansefield Management Competency Model

Introduced in 2003, this model states the manager's competencies in three areas: employees with two competencies, leadership, communication and influence, job domain with two competencies of achieving the results and follow-up and problem solving and self-management domain. The model indicators are presented in Tab. 5 [7].

Table 5 Mansefield Management Competency Model

Dimension	Item	Index
Employees	Communication and influence	Paying attention to communication
		Verbal communication
		Written communication
		persuasiveness
		Interpersonal awareness
		Penetration skills
		Establishing collaborative relationships
		Customer orientation
		Empowering others
		Change Management
	Others direction	Development of others
		Performance Management

Dimension	Item	Index
Job	Achieving results	Create a focus of attention
		Motivational support
		Teamwork nurturing
		Innovation
		Entrepreneurial orientation
		Nurturing innovation
		Result-orientation
		Precise and scrutiny
		Decisiveness
	Problem follow-up and solving	Gathering diagnostic information
		Analytical thinking
		Leading thinking
		Conceptual thinking
		Important thinking
		Technical expertise
	Self-management	Self confidence
		Stress management
		Individual trustworthiness
		Flexibility

2.3.1.6 Zenger et al. Model

The management competencies triangular with four dimensions are provided by Zenger and Folkman (2002), which includes important business skills, individual capability, cultural processes skills, and change and human resources skills. The model indicators are listed in Tab. 6.

Table 6 Zenger et al. Model

Index	Dimension	Index	Dimension
Change and Culture Management	Cultural processes and change skills	Important thinking and planning	Important Business Skills
Organizational Design		Customer-orientation	
Consultation		Leadership	
Facilitator		Important Sector Analysis	
Project management		Communications	Personal capability
HR planning	HR skills	Coaching	
Recruitment, training and development		Self-awareness	
Performance management, rights and benefits		Professional training	
Labor relations, conflict management		Participation	
		Integrity	

2.3.1.7 Managerial Assessment of Proficiency Tool (MAP)

The dimensions of this tool include executive skills, communication, monitoring and diagnosis. The model components and indicators are listed in Tab. 7 [4].

Table 7 Managerial Assessment of proficiency Tool (MAP) [4]

Dimension	Item	Index
Executive	Time management and prioritization	Prioritizing
		Practicing self-discipline
		Controlling interruptions
		Doing important things instead of everything
	Setting standards and goal-setting	Support activities, goals, standards
		Reducing barriers
		Results evaluation
	Work planning and scheduling	Complex task analysis
		Source selection
		Work planning and scheduling
Communication	Listening and organizing	Setting up inspection and control points to monitor the process
		Identify and test assumptions and inferences
		Overcoming the obstacles to listening and organizing
		Refining and rearranging a message for reminder
		Detection and control of assumptions and inferences
	providing clear information	Overcoming barriers to information clarity
		Goals confirmation
		Maintaining a climate of mutual benefit and trust
	Getting undistorted information	Using guiding, non-guiding, reflective, and reasoning questions
		Understanding confirmation
		Get agreement
	Training, guidance and delegation	Choosing the Right People
		Reaching agreement on plans to perform
		Maintaining balance between input and output
		Assignment of responsibility
		Providing effective feedback
Monitoring	Individual and performance evaluation	Providing appropriate rewards
		Maintaining evaluations
		Focus on performance, not personality
		Reaching agreement on future expectations
		Following the executive plans
	Discipline and consolation	Maintaining performance
		Determining the scope of individual responsibility
		Enhanced performance enhancement
Diagnosis	Problem identification and solving	Diagnosing and distinguishing between symptoms and problems
		Data collection
		Identifying the root cause
		Weighting options
	Decision making and risk weighting	Take appropriate action
		Identifying the constraints and risks considered
		Assigning weight to each option
		Selecting the best option to achieve the desired goal
	Explicit and analytical thinking	Identifying valid assumptions
		Drawing logical results
		Recognizing misconceptions, false assumptions and overview without evidence

2.3.1.8 American Society for Training and Development (ASTD) Model

The American Society for Training and Development has considered three dimensions of competence, interpersonal competencies, management competencies and personal skills into its model. The model indicators are provided in Tab. 8 [21].

Table 8 American Society for Training and Development (ASTD) Model [21]

Dimensions	Index
Inter-personal skills	Trust
	Effective communication
	Influencing stakeholders
	Leveraging diversity
	Networking and Partnerships
Managerial skills	Needs analysis and solutions
	Business intelligence
	Result-orientation
	Scheduling and executing assignments
	Important thinking
Personal skills	Modelling personal development
	Adaptation

2.4 Application of Competency Models

As a set of applicable standards of skill in recruiting, leading jobs, evaluating employees, designing scientific programs and issuing industry certifications, competency models are a bridge between coaches, the business sectors and other stakeholders to prepare students and workers to deal with the challenges of the business environment where they have invested. Competency models can also be used by employees as an effective option and a tool for career development. These models are also able to assist human resources staff to adapt specific professional and work needs to different jobs in recruitment, promotion and career development while developing training programs. The competency models can be used to evaluate individuals' job performance, as well as the role of managers, reporting, customer and team members, and as a means of conveying performance expectations to employees in business environments.

These models can also be used to measure the gap between staff requirements and current training programs, and the system for providing training, reviewing existing courses, and designing competencies, and also as a path to identify gaps. The competency model can be seen as a benchmark when evaluating existing education and training programs or when designing new programs by program providers, resulting in courses that meet workplace needs and trends.

These models can be utilized by training program providers to obtain valid industry certificates. Such certificates prove that the graduates of a particular curriculum have earned the necessary competencies and dominance in a given field.

As a guide to investment boards and professional centers, competency models can match job requirements with the skill sets that employers determine for potential candidates. These services can even be offered to larger groups such as

high school, non-high school youth, workers transfer from one unit to another and current workers, as well as groups with special needs, and consequently, causes increasing the range of available workforce talent. Since all of these main partners work together through sharing assets and resources, the competency model plays an effective guiding and important role for government investment in the preparation of workforce strategies in an area or province.

3 RESEARCH METHOD

The present study is an applied research in terms of the purpose, which uses a combined research strategy that can have multiple applications in industry and organization. The study aimed to identify and prioritize the researchers' important competencies in research centers. The Delphi method (qualitative) method and the structural equation modelling (SEM) are used in this research to discover the most important researchers' important competencies and prioritize important competencies, respectively. The Delphi method can be used in qualitative research with exploratory dimensions to identify the nature and fundamental elements of a phenomenon. In this regard, Delphi method is utilized to "identify" and "screen" the most important decision-making indicators. The use of the professional knowledge and expertise of a set of experts is very helpful in making decisions about problems with a qualitative nature. Generally, Delphi method is defined as a research approach to reach consensus using a series of questionnaires and providing feedback to participants who have the related expertise. There is no standard statistical methods in Delphi method and expert opinion survey. However, data analysis methods depend on the research purpose, the type of questions, and the number of respondents. The data collection tool in this study are interviews and questionnaires. People participating in the Delphi method must be knowledgeable about the subject matter of the study and are referred to as "panellists" or "experts". Non-probabilistic sampling with the purposeful sampling and snowball sampling methods are used to select the experts. The study population was selected among the faculty members in the field of management and human resources, and the managers and researchers in the field of research and technology to reach the valid results. About the sample size, a sample of 10 to 20 qualified professionals have been estimated as a valid sample. In this regard, a list of individuals with high research, knowledge and expertise was selected to participate in the study who included faculty members and senior managers in the field of research and technology. The other experts were introduced during the interview by experts in the early stages of the interviews. Interviews continued until the content saturated that is, new components or dimensions were not mentioned by the interviewees. Finally, the interview would continue for 20 people. The participants' profile are presented in the Tab. 9.

Firstly, the components of researchers' important competencies were extracted through the exploratory phase by interviews. Then, a semi-structured questionnaire was used in the first Delphi round to limit the competencies, and

a closed questionnaire was used in the second and third Delphi round to converge the findings and discover the valid consensus. The expert agreement is reached in the present research when the average expert opinion on each item is above 4 (agreement). Kendall's coefficient was used to control the validity of the research instrument which was estimated at 0.763 in Delphi final stage which indicates appropriate validity. Generally, it should also be said that the content validity is guaranteed if the panel members have the needed expertise and quality as a qualified representative in the related field of knowledge. On the other hand, confirmatory factor analysis and construct validity are used to prioritize competencies and Cronbach's alpha was used to determine the questionnaire reliability which was estimated at 0.77 indicating acceptable instrument reliability.

Table 9 The expert panel profile

Variable	Group	Frequency	Percentage
Gender	male	14	70
	Female	6	30
Education	Ph.D.	12	60
	Master degree	8	40
Experience	Lower 10 years	5	25
	10-20	8	40
	Over 20 years	7	35
Position	Faculty member	11	55
	researcher	9	45
Total		20	100

4 RESEARCH FINDINGS

4.1 Qualitative Stage (Delphi)

The most important researchers' important competencies was provided to the experts in the first phase. The question is raised in the presence of most members and through open-ended interviews. Next, they were asked to suggest the components that were not considered by communicating with each of the members and explaining some of the elements extracted in the theoretical bases. After receiving the responses and examining the experts' views, similar items were merged and a number of non-relevant components were removed from the present study. The list of researchers' important competencies factors was compiled and analysed in the first phase of Delphi with the help of panel members and advisor and consultants. Finally, the initial questionnaire form, which was developed based on the literature using survey methods from experts and professors and initial analysis, was revised and modified according to the experts' suggestions and the final questionnaire was confirmed after two rounds of the Delphi method. The final questionnaire with 59 competency components was developed in the second phase and provided to the panel members to determine the importance of each component. The results were presented to the members in the next stage and the members were asked to review their responses and revise them if needed and state their reasons for disagreement. Finally, the consensus was achieved regarding 48 components of competence according to the following table and the final model was obtained.

Table 10 The Delphi results on the components of the researchers' important competence

Empowerment and competence	Average comments	Kendall coefficient	Rank
Professional scientific knowledge	4/76	0/768	1
The desire for continuous learning	4/56	0/762	2
Ability to teamwork	4/56	0/762	3
Innovation creativity	4/56	0/763	4
Analytical thinking	4/55	0/762	5
Critical Thinking	4/53	0/760	6
Decision making	4/51	0/764	7
Partners' passion / motivation	4/51	0/750	8
Integration	4/50	0/751	9
Formulating a research topic / problem solving	4/49	0/762	10
Moral competency	4/48	0/760	11
Time Management	4/48	0/764	12
Result-orientation	4/47	0/750	13
Systematic thinking and planning skills	4/45	0/762	14
Knowledge of management and organization	4/44	0/748	15
Foreign Language Skills	4/44	0/750	16
Developmental learning/ Future studies in learning	4/33	0/739	33
Water crisis	4/33	0/751	34
Crisis Management	4/32	0/748	35
Sustainable Development and Climate Change	4/32	0/750	36
Stress management	4/32	0/739	37
Self-confidence	4/31	0/751	38
Listening skills	4/30	0/748	39
The art of Influence	4/30	0/752	40
Analysis and use of ICT	4/44	0/760	17
Leadership	4/43	0/764	18
Adaptability / compatibility	4/43	0/750	19
Networking development capacity	4/43	0/762	20
Effective negotiation and communication skills	4/42	0/748	21
Project management	4/42	0/750	22
Important thinking	4/41	0/739	23
Risk management	4/39	0/751	24
Communication skills	4/38	0/748	25
Business intelligence and entrepreneurial spirit	4/36	0/760	26
Developing people	4/36	0/739	27
An open minded approach / horizonology	4/36	0/751	28
Challenge decision-making capacity	4/36	0/748	29
Challenge	4/36	0/750	30
Accountability and customer orientation	4/35	0/739	31
Resource management	4/35	0/751	32
Capacity of evaluating others	4/29	0/748	41
Conflict management	4/24	0/739	42
Financial management	4/21	0/751	43
Globalization	4/20	0/748	44
Science and Technology Policies	4/19	0/753	45
Changes	4/18	0/739	46
Regional turbulence	4/18	0/751	47
Transnational market	4/14	0/748	48

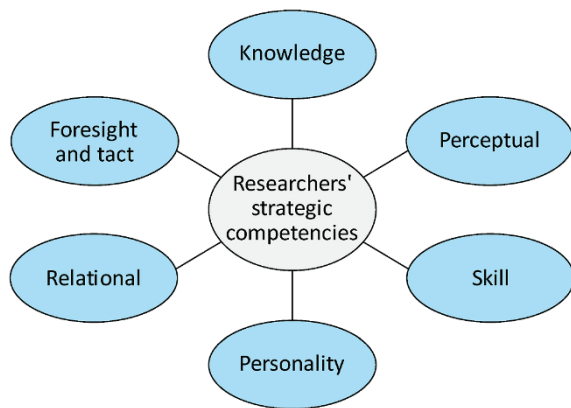


Figure 2 The final model of researchers' important competencies

4.2 Quantitative Stage (Structural Equation Modelling)

The validity of the model construct is tested in this section using structural equation modelling techniques by LISREL software.

4.3 Confirmatory Factor Analysis

The confirmatory factor analysis is mainly aimed to determine the power of a pre-defined factor model with a set of observed data. That is, confirmatory factor analysis aimed to determine whether the number of factors and variables measured on these factors correspond to what was expected by theory and theoretical model. In other words, this type of factor analysis tests the degree of conformity between the theoretical and empirical construct. Hence, validating the model is one of the functions of confirmatory factor analysis. The standardized loads are of particular importance in confirmatory factor analysis when interpreting factor analysis results. These loads indicate the correlation between each observed variable and its associated factor. The strength of the relationship between the factor (hidden variable) and the observable variable is represented by the factor load. The higher the factor loadings of an index in relation to a particular construct, the greater its contribution in explaining that construct. Also, if an index factor loading is negative, it indicates its negative impact on the explanation of the relevant construct. That is, the relevant question is designed inversely. The factor loading is a value between zero and one that if it is less than .3, the relationship is weak and it can be ignored. A factor loading between .3 and .6 is acceptable, and it is highly acceptable if it is greater than .6 [20].

4.4 Factor Analysis of Knowledge Construct

The factor loading of all the competencies of the knowledge construct is at the least acceptable status and the highest factor loading is related to the financial management variable (q3) with .83 according to Fig. 3.

In order to examine the significance of the relationship between variables, the T-test or t-value is used and since it is significant at the 0.05 level of error. So, if the factor loadings observed by the t-value test are lower than 1.96, the relationship is not significant and if it is greater than 1.96 the

relationship is significant and the model validity is confirmed.

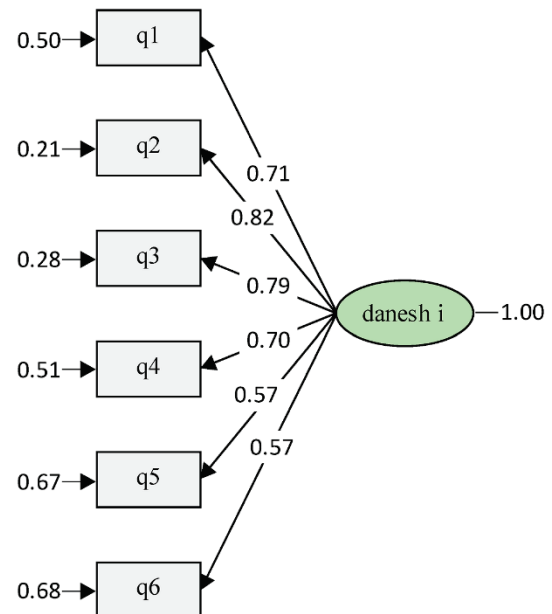


Figure 3 Factor Analysis of Knowledge construct (Standard Estimation Model)

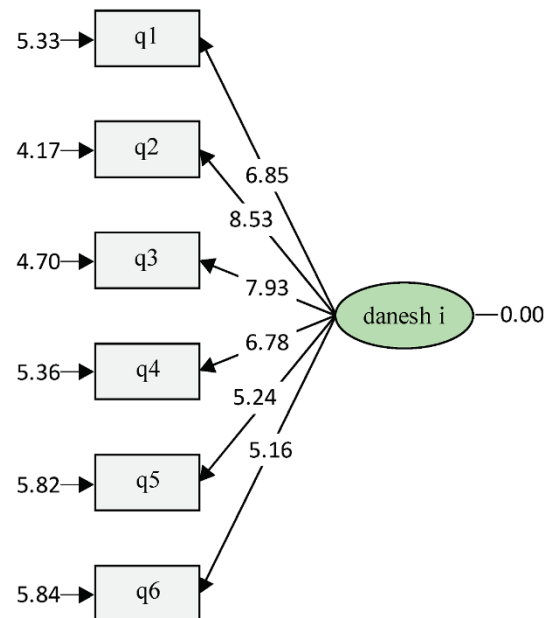


Figure 4 Factor Analysis of Knowledge construct (significance model)

Since all relationships between variables and constructs (knowledge competence) were greater than 1.96 according to Fig. 4 all relationships are significant and confirmed.

4.5 Second-Order Factor Analysis

The first-order Confirmatory Factor Analysis evaluates the relationship between factor(s) (latent variables) with items (observable variables). No relationship between latent variables is examined in this method. This type of measurement model is simply to ensure that latent variables are measured correctly. One can examine the relationship of

one factor with multiple items or factors or multiple items in the first-order confirmatory factor analysis.

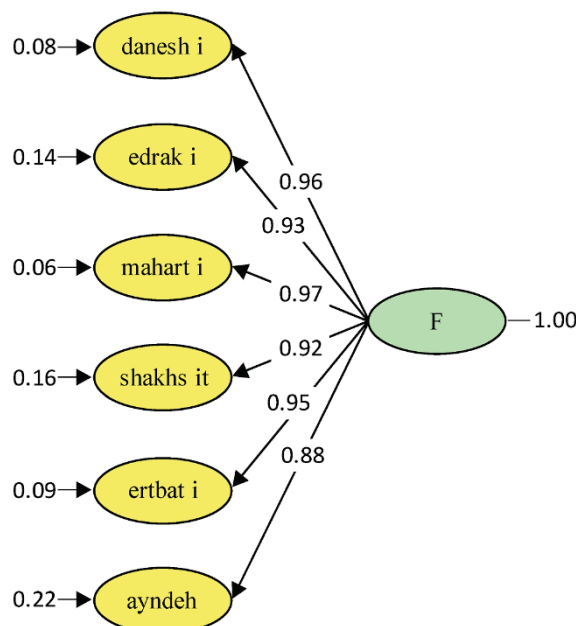


Figure 5 The second-order Factor Analysis (Standard Estimation Model)

However, when a large construct is composed of several latent variables, the second-order confirmatory factor analysis is used. In addition to examining the relationship of observable variables with latent variables in the second-order confirmatory factor analysis, the relationship of latent variables with their main construct is also examined.

The factor loading of all the competence groups (6 groups) is in optimal status according to Fig. 5, as all factor loading are greater than .70. It is also seen that the highest factor loadings belong to the skill class with 0.97

The final questionnaire was distributed among the researchers of the research and the score of teach index is obtained based on the Likert classification that the mean of each index was determined and confirmed based on rank and score using the following statistical methods.

5 CONCLUSIONS

The important competencies needed by researchers in the research centres are listed in the table below according to the average score of experts and researchers. As can be seen, the creativity and innovation, integration and accountability, and customer orientation competencies have higher priorities followed by the women presence in the labour market, the water crisis, financial management, the use of ICT and self-confidence. However, all the identified important competencies have a significant positive distance to mean. The researchers were also asked to provide any other mention to incorporate in the second phase of the questionnaire.

The important competencies needed by researchers according to the mean score of the Research Institute researchers are provided in the table below. According to this table, the competencies of scientific and professional

knowledge, the tendency for continuous learning, the ability to perform, creativity and innovation, analytical and critical thinking have top priorities followed by women's presence in the labor market.

The final questionnaire was distributed among the researchers of research centers and two indices were selected based on Likert classification so that the mean of each index was determined by rate and score using the specified statistical methods and confirmed by the statistical methods.

6 REFERENCES

- [1] Afkhami Ardakani, M., Baba Shahi, J., & Tahmasebi, H. R. (2016). Providing a Tool for Identifying and Measuring Knowledge Jobs, a Mixed Approach Research. *Journal of Human Resource Management Research*.
- [2] Baba Shahi, J. et al. (2017). 360 Degree Assessment of Exploration Management Employees and Managers. Institute for International Energy Studies.
- [3] Boyatzis, R. E. (1982). *The Competent Manager: A Model for Effective Performance*. John Wiley & Sons, New York, NY.
- [4] Beyhan, O. (2002). İngilizce dersinde yansıtıcı düşünme etkinliklerinin öğrencilerin akademik başarılarına ve dersle yönelik tutumlarına etkisi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 1(2), 128-142.
- [5] Boyatzis, R. E. (1982). *The Competent Manager: A Model for Effective Performance*. John Wiley & Sons, New York, NY.
- [6] Cochran, G. R. (2009). Ohio State University Extension Competency Study: Developing a Competency Model for a 21st Century Extension Organization. *Doctoral dissertation*, Ohio State University. OhioLINK Electronic Theses and Dissertations Center.
- [7] Corallo, A., Lazoi, M., Margherita, A., & Scalvenzi, M. (2010). Optimizing competence management processes: A case in the aerospace. *Business Process Management Journal*, 16(2), 297-314. <https://doi.org/10.1108/14637151011035615>
- [8] Dubiose, D. D. & Rothwell, W. J. (2004). *Competency-Based Human Resource Management*. Davies-Black Pub., Business & Economics, 291 p.
- [9] Farmahini Farahani, A. (2015). The Competency Models of Employment and Training Administration. *Department of Labor Education and Training*, 16(1), 290-295.
- [10] Onet Competency Model. (2017). (www.onetonline.org)
- [11] Hsieh, S., Lin, J., & Lee, H. (2012). *Analysis on literature review of competency*. Department of international Trade & Logistics, Overseas Chinese University.
- [12] Farmahini Farahani, A. (2013). *Hypertobs & Future Skills in the Field Energy*. Institute for International Energy Studies (IIES), Tehran.
- [13] Klett, F. (2012). The design of a sustainable competency based human resource management: A Holistic Approach. Fraunhofer Institute Digital Media Technology, Germany. *Knowledge, E-Management & E-Learning*, 2(3).
- [14] Liu, P. & Tsai, C. (2008). *A study on R&D competence for R&D Management personnel in Taiwan's High-tech Industry*. Department of Industrial Engineering and Management.
- [15] Lievens, F., Sanchez, J. I., & De Corte, W. (2004). Easing the inferential leap in competency modeling: the effects of task-related information and subject matter expertise. *Personnel Psychology*, 57, 881-904. <https://doi.org/10.1111/j.1744-6570.2004.00009.x>
- [16] Dreyfus, C. R. (2008). Identifying competencies that predict effectiveness of R&D managers. *Journal of Management Development*, 27(1). <https://doi.org/10.1108/02621710810840776>

- [17] Markus, L. H., Cooper-Thomas, H. D., & Allpress, K. N. (2004) Confounded by competencies? An evaluation of the evolution and use of competency models. *New Zealand Journal of Psychology*, 34(2), 117-126.
- [18] Mansfield, R.S. (1996). Building Competency Models: Approaches for HR Professionals. *Human Resource Management*, 35(1), 7-18. [https://doi.org/10.1002/\(SICI\)1099-050X\(199621\)35:1<7::AID-HRM1>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-050X(199621)35:1<7::AID-HRM1>3.0.CO;2-2)
- [19] Sinnott, G. C., Madison, G. H., & Pataki, G. E. (2002). *Competencies: Report of the competencies workgroup workforce and succession planning work groups*. New York State Governor's Office of Employee Relation and the Department of Civil Service.
- [20] Lawler, E. E., III. (1994). Performance management: The next generation. *Compensation and Benefits Review*, 26(3), 16-19. <https://doi.org/10.1177/088636879402600303>
- [21] Schippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B., et al. (2000). 23. The practice of competency modeling. *Personnel Psychology*, 53, 703-740. <https://doi.org/10.1111/j.1744-6570.2000.tb00220.x>
- [22] Macharzina, K. & Wolf, J. (2008). Unternehmensführung. Gabler. *Wiesbaden*, 33(2), 117-1234
- [23] Zenger, J. H. & Folkma, J. R. (2011). An Overview of Key Insights from the Extraordinary Leader, Extraordinary Performance. ATEL50.13-1.1, 1-8. <https://www.clemmergroup.com/wp-content/uploads/2014/08/ExtraordinaryLeaderInsights-TCG.pdf>
- [24] Zemke, R. (1982). Job Competencies: Can they help you Design Better Training? *Training*, 19(5), 28-31.

Author's contacts:

Ahmad Farmahini Farahani, PhD in Technology Foresight
 Research Institute of Petroleum Industry,
 Bolvar-e-Gharbi-e-Stadium-e-Azadi, Tehran, Iran
 E-mail: farmahiniaahmad@yahoo.com

Computer Game Elements and its Impact on Higher Education

Andrija Bernik*, Damir Vusić, Voravika Wattanasoontorn

Abstract: This review paper analyses authors who used gamified e-learning systems in higher education. Their thoughts, problems and conclusions are accompanied and observed. Positive guidelines are provided for the inclusion of computer game elements in the education system. Review paper includes a comparison of the gamified and traditional education system with an incentive to conduct further research activities. The list of researchers is provided who had experiments on students in the higher education system. Duration of those experiments is shown as well as the number of used game elements for their gamified systems. The list of survey questions for teaching staff (N=98) is provided as well as their thoughts on gamified systems and its use. The conclusions are that it is necessary to pay attention to the visual appearance and environment in which teaching materials are presented and studied.

Keywords: e-learning; game elements; gamification; higher education; teachers

1 INTRODUCTION

This article lists the authors of papers from gamified e-learning systems, and their research that is accompanied by observed problems and conclusions. Positive guidelines are provided for the inclusion of computer game elements in the education system. Manuscript concludes with a comparison of gamified and traditional education systems with an incentive to conduct further research activities. There are several keywords that should be familiar to the reader such as gamification, e-learning, game elements and flow theory. Strmecki et al. [1] explains that gamification is not so new trend of including video game elements into non-game environment such as education systems. Bernik et al. [2] examines best video game elements for implementation into e-learning systems such as score, feedback loops, collaborations, badges, levels, etc. Most of them are available within Moodle CMS system and are free to use. Furthermore, Bernik et al. [3] states that E-learning represents digital (both synchronous and asynchronous) approach where teaching and tutoring materials are online and users are connected with internet connection. If the E-learning system is designed correctly, the student creates a sense of comfort that can lead to a state of flow in which the learning process is optimal. This is one of the conclusion by Bernik [4] within the manuscript where gamification framework was described.

2 RESEARCH OVERVIEW

Landers and Callan [5] explore techniques for motivating students through the development of a one-semester educational social network. The study included 385 students who accepted an invitation sent via e-mail. Introduction to the aim and purpose of the research was conducted through a five-minute lecture. Landers and Callan have included in their e-learning system, in addition to teaching materials, the possibility of using a profile picture and posting public announcements about own achievements. The system supports asynchronous communication as well as badge and certificate assignment. Their system contained several knowledge tests with closed-ended questions for which certificates were awarded. The knowledge test

contained 10 questions for the answers of which the student had 10 minutes. Access to the knowledge test was provided every four days. Particularly interesting in their approach is the mentoring system, which, based on the certificate, enables the pairing of students with regard to individual achievements and learning needs. A student who has difficulties mastering a teaching unit may request the assistance of a proposed student who has successfully mastered a certain teaching activity and justified knowledge through the previously mentioned 10-minute knowledge test. Landers and Callan point out that at the end of the study, 113 out of a total of 385 students had won a total of 546 certificates. They also point out that internal (university) social networks can have a positive effect on student motivation. Rewarding students should be explicit and recognized in the social network, because the clearer the reward and the more valuable it is to students, the greater the motivation. The reward should be matched to the difficulty of the desired activity or knowledge test. Difficulty should be built gradually. Feedback for activities and knowledge tests should be current. It is necessary to pay attention to visual appearance and the environment in which the teaching materials are studied. Muntean [6] links the state of flow with the research of B. J. Fogg who proposed the so-called Fogg's Behavior Model, which is based on (a) motivation, (b) abilities and (c) drivers of certain activities. According to Muntean, motivation can be caused by feelings of comfort or pain, fear or hope, and social status, which can be acceptance or rejection. If a student has a high level of motivation, but lacks a certain level of ability, a change in behavior and / or an effective learning process will not be achieved. In addition to the above, it is necessary to include the triggers of certain activities, i.e. elements that suggest the student to do a certain activity at a certain time. Triggers can be simple signalling, all the way to a narrator who reacts at a crucial moment and directs student's activity. A pre-requisite for this is the level of education of a student in which a student is able to recognize, learn and reason based on the information presented. The state that occurs when all three elements are activated is called the flow state. It is pointed out that there is a problem in working with an e-course when conveying stimulus or emotion that is present in the interaction between

a teacher and a student. It is for this reason that e-learning systems need to be expanded with technological solutions that affect motivation and encourage students to engage in teaching activities.

The elements of mechanics from computer games that can be used to stimulate intrinsic motivation (IM) are listed as follows: Points (IM: Status, Achievement, Competition, Altruism), Levels (IM: Achievement, Competition), Challenges (IM: Rewarding, Status, Exploration, Competition, Altruism), Virtual currency (IM: Rewarding, Status, Achievement, Competition), Top list (IM: Status, Achievement, Altruism), Gifting (IM: Status, Achievement, Competition).

Muntean states that for the correct gamification of e-learning it is necessary to gradually discover teaching content, and put emphasis on repetition, practice, and knowledge testing. For correct solutions, a student needs to be provided with points. The goal to be achieved by gamifying e-learning is to ensure higher motivation among students as well as the involvement of students in teaching activities that are available in a gamified system. Muntean states that if students stay in the system longer, it is very likely that they will achieve a better result on the knowledge test. As a result, students will master teaching activities in the future faster with better results. In traditional e-learning systems, a teacher, administrator, or e-course designer should include some of the following elements of computer games: avatar, notification, separate chapters, and separate teaching levels. It is concluded that a student, after mastering certain teaching units, progresses to higher teaching levels where the process of mastering new teaching units is repeated with difficulty corrections of repetitive tasks and the correction of abstraction of teaching materials.

The e-course should reward students not only for their academic achievements, but also for a certain type of behavior, such as helping and communicating with other e-course participants. In addition, it is important to ensure that students know what their responsibilities are and what awaits them in the next course. In addition, it is important to create a sense of surprise and anticipation in order to influence motivation of the participants at the time of discovering new elements of the gamified system. Barata et al. [7, 8] conducted the research for two years. Elements of computer games (points, levels, top lists, challenges, and badges) are included in the Moodle system with the aim of positively influencing the interest of higher education students. The research was focused on motivation and learning outcomes, but due to the duration, there are significant differences in the implementation of the experiment: unequal number of respondents in groups, unequal teaching materials, and unequal conditions in which classes were held. The result of the research was a greater interest of students for lectures and for participation in the e-course, proactivity, and greater use of teaching materials. Barata et al. state that the research does not show a direct impact of motivating elements of computer games on the achieved course grades, and that is suggested for future consideration and additional research.

Jayasinghe and Dharmaratne [9] explore learning through educational computer games and learning through elements of computer games without the presence of a game in the field of programming. The research was conducted on

60 students who previously had no education in the field of programming, more precisely program-sorting algorithms. The experiment lasted 60 minutes and contained points and feedback on the correct answers that the students received. Students were divided into 4 groups of 15 respondents. Two groups were experimental, while the other two were control. The statistical significance of point difference between the groups was not calculated, but the groups were divided according to previously obtained grades in mathematics. The sorting algorithms used in the research were sorting by replacing adjacent elements (Bubble sort) and segmental sorting (Bucket sort).

The results of the study show that both of their experimental groups had higher average scores. In the case of sorting by replacing adjacent elements, the average score of both experimental groups was 21.3% higher compared to both control groups. In the case of segmental sorting, the average result in both experimental groups was 25.3% higher compared to both control groups. In addition to average points, the time in which students studied teaching materials in the field of programming was also measured. The conclusion is that both experimental groups mastered teaching materials in almost twice less time than both control groups of respondents, which concluded the research.

Ortega-de-Marcos et al. [10] developed a gamified add-on for Black Board LMS system that allowed tracking teaching activities of a total of 371 students, as well as mutual collaboration, but also competition. Teaching activities are related to four areas: word processing, spreadsheets, presentations, and databases. Each activity was presented as a challenge and was divided into several levels to help students achieve a sense of progress. The e-course was open to students for one semester. It was active 6-10 hours a week. After each level, a student is awarded a badge in the form of a trophy. Several badges were hidden to keep students surprised and further motivated. In the research, collaboration and sharing of achievements was ensured in the form of a forum. In addition to a gamified supplement for monitoring teaching activities, a web platform was developed that provided students with access to video materials and blogs. Students were able to follow other users through the so-called follow function. Students had the opportunity to post short messages, mark agreement with other people's opinions through a like function, as well as commenting on other activities.

The research uses three research groups, where the first group (consisting of 114 respondents) was exposed to a gamified e-course, the second group (consisting of 184 respondents) was exposed to a social network, and the third group (consisting of 73 respondents) was a control group which used a classic e-course. The experiment was conducted during 2012. The place of research were various university centres depending on the group of students. de-Marcos et al. used a pre-test and post-test methods to test student knowledge and a survey questionnaire (consisting of 10 questions) that was presented only to the first and second groups of respondents.

A pre-test was used to test knowledge between all three groups. The results of the pre-test indicate that there was no statistically significant difference between the groups. Based on the post-test knowledge test, it was determined that both

experimental groups achieved better results in the practical part of the knowledge test compared to the control group in all four teaching units (word processing, spreadsheets, presentation, and data bases). The group that used social networks achieved better results in the first two teaching units (word processing, spreadsheets) compared to the group that used a gamified supplement within the Black Board environment. Other results did not differ statistically.

Interestingly, control group students scored better than both experimental groups in terms of written knowledge testing, where the group that used social media scored better than the group that used the gamified supplement. Within the written test, all three results provide a statistically significant difference. It is concluded that both experimental groups place more emphasis on practical skills, where theoretical knowledge is neglected. The authors (de Marcos) [10] noted a problem where both experimental groups showed very little interest in teaching materials. Approximately 20% of students actively participated in this research, which puts emphasis on the design of the research. It is necessary to repeat this research, which would focus on the social component instead of achievements, badges, and competition.

Iosup and Epema [11] created two courses that were conducted over four semesters on over 450 students. In addition, gamification is considered as a set of tools that can influence the motivation and behavior of users. The researchers list seven gamification elements that are divided into two categories: mechanics and dynamics. Mechanics defines how a game functions as a system, while dynamics defines the interaction of a game and a user. Within the mechanics of computer games are points, levels, and scales of success. Within the dynamics of computer games, badges, an explanation for beginners, social loops and gradual unlocking of content are included. It was observed that over 75% of students passed the knowledge test in the first exam period. A positive correlation was also observed between students who passed the knowledge and satisfaction test attributed to the gamified elements. It was observed that almost 100% of students accessed bonus teaching materials even though those materials did not score for the overall grade. Students who have not successfully passed the knowledge test retake the test in 90% of cases. Students who have passed the knowledge test reapproach the test in 5% of cases with an aim of achieving even better result.

Laskowski and Badurowicz [12] conduct research within the first year of graduate study lasting one semester. The sample of respondents consists of 62 students who are divided into four groups. Two groups are experimental, while two are control. The experimental groups were in a system that assigns points and badges to participants for achievements and allows for a ranking scale. When analyzing the results, at the end of the semester, the following indicators were taken into account: the number of attendances at regular classes, the number of completed assignments, the average grade and similar information such as homework and project assignments collected during the semester. The aim was to examine the impact of using elements of computer games, without the presence of a game, on student motivation. The results of the research show that the experimental groups of respondents have a higher rate of attendance at classes (97%)

compared to the control groups of respondents (85%) as well as the number of solved individual tasks (1.7) compared to the control groups of respondents (1.37). However, the average grade at the end of the semester in the experimental group is 3.83 out of 5, while in the control group it is 4.53 out of 5. Laskowski and Badurowicz [12] confirm that gamification can be used to motivate students but cannot find a link between higher grades in the control group of respondents and lower attendance at teaching activities. They conclude the research with a remark that the research should be repeated with a larger number of respondents, with a pre-test and post-test of knowledge.

González and Carreño [13] research which elements of computer games can be used in traditional education, e-learning and hybrid learning in terms of student involvement in educational process. González and Carreño cite Werbach's 6D model as well as the individual elements of dynamics, mechanics, and aesthetics used within this paper. The sample used in the research numbers 100 third-year undergraduate students of computer engineering. The research activities were conducted for a total of four weeks. The research methodology is reduced to observation method, interview method and a survey questionnaire. All three instruments were taken from other researchers. For research purposes, they create their own platform within which points, ranking scale, user groups, avatars, virtual objects, etc. are implemented. Moodle platform was rejected because it did not meet the software architecture of the server used in the research. There are no more detailed explanations for rejecting the Moodle platform. They conclude that gamification or elements of computer games without the presence of the game itself can satisfy basic human desires or needs such as: identity, reward, achievement, competition, collaboration, self-expression, and altruism. Students showed how these elements are important to them in everyday life as well as in the virtual environment in which they are presented with teaching materials. The social component is particularly important, distinguishing between cooperation and competition as well as independent and teamwork. In addition, gamified solutions encourage students to a higher level of involvement whether it is formal or nonformal education, due to a phenomenon that is described as a state of flow.

Stokes [14] warns that the education system, in its present form, operates on the principle of a kind of "deduction of points" given the number of mistakes. A student must not make mistakes, which leads to stress and additional pressure, which negatively affects his/her motivation. The education system should recognize the degree of motivation and guide students on the right path through the elements of computer games. Deterding [15] points out that elements of computer games directly affect motivation and a sense of inclusion, which is widely used by marketing companies and employees working in the field of human resources. Ortega de-Marcos et al. [10] confirm a positive attitude and better motivation of respondents, with the achievement of better learning outcomes during the presence of elements of social networks in the educational system. Hamari et al. [16, 17, 18] come to a conclusion that confirms the research of Deterding and partly the research of de-Marco et al. Within academic education, researchers

Smole, Diniz and Milani [19] announce that in working with elements of computer games. Students can learn and develop their self-confidence in various ways, such as interaction and collaboration, critical and problem-oriented thinking, self-analysis of work and activities, and current feedback on success or failure, where failure is approached as a basis for new guidance and learning.

Dicheva et al., [20] within the FP7 project, investigate which teaching content can gamification be applied to, and which are the most commonly used elements of computer games in the education system. The analysis by Dichev et al. includes 34 empirical papers found within some of the leading scientific databases (ACM Digital Library, IEEE Xplore, ScienceDirect, SCOPUS, Springer Link, ERIC, and Google Scholar). Dicheva et al. state that points, badges, levels, rankings, progress status, virtual currencies and avatar are the most used computer game mechanisms used in the education system, of which points, badges, levels, and top rankings are much more common elements than the others are. The design of a gamified system is based on visual status, social inclusion, and social media, as well as freedom of choice, a positive approach to failed attempts and constant feedback.

The educational style that is most covered within the above 34 empirical papers refer to Blended Learning Courses. In second place are courses with or without online support, in third place are massive open online courses (MOOC), followed by educational websites, and gamified customer support platforms. In addition to the above, the most frequently covered educational areas are computer science and IT, computer game programming and STEM areas.

In addition, it is stated that clear indicators explaining the impact of computer game elements on motivation are still lacking. There are no reports on the sensitivity of computer game elements and the implementation of practical educational gamified solutions. There are no reports on the need for teachers to effectively conduct, monitor and evaluate students within the gamification approach. There is also a negative experience in which students were not ready for autonomy, and the project approach was not taken seriously enough, as was the approach to knowledge testing, which concludes the paper.

Schreuders and Butterfield [21] explore opportunities to increase student involvement in the teaching process and increase motivation, and to improve the experience that students have when going through the educational process. The research was conducted within the university course on computer security. The research lasted two years and involved 32 students. For the purposes of the research, a special open source software solution was created that ensured the transparency of points, the allocation of various points for student activities as well as providing feedback on the success of student activities.

The results of the research indicate positive indicators in terms of qualitative and quantitative results in knowledge tests. There was an increase in the satisfaction that students had when going through the educational process, but the results were not statistically significant. It is concluded that the research, despite the small number of respondents, is in line with other research that speaks about the positive effects

of gamification on increasing motivation and improving the experience that students have when using e-learning systems. Further research is needed to make the higher education system more receptive to computer game elements at all levels.

The conclusions that can be summarized from these studies are that it is necessary to pay attention to visual appearance and environment in which teaching materials are presented and studied, and that gamified systems could include avatars, news reminders, separate teaching units as well as separate teaching levels. These mechanisms will be taken into account when selecting elements of computer games to be included in the gamified pre-study system as well as in the first and second main surveys. The problem observed in the research of de-Marcos et al., in which the experimental groups of respondents showed very little interest in teaching materials and where approximately 20% of students actively participated in the study, will be particularly monitored in this study in order to avoid repeating the same. Also, the seven elements of computer games (points, levels, success scale, badges, explanation for a beginner, social loops, and gradual unlocking of content) that Iosup and Epema list, were considered for inclusion in the gamified system, along with other elements described in Bernik [4], with the aim of confirming the results cited by Iosup and Epema [11].

Special emphasis will be placed on the situation seen in Laskowski and Badurowicz [12], where the control groups of respondents at the end of a semester achieved higher average grades than the experimental groups of respondents, where the control groups of respondents had lower attendance at teaching activities. Laskowski and Badurowicz [12] could not explain such a result and emphasize that the research should be repeated with a larger number of respondents, which was done in Bernik [4] with a significantly larger number of respondents (309 in total) and a larger number of included computer game elements (24 in total)

3 DISCUSSION: IMPACT OF VIDEO GAMES ON TEACHERS

Parreno et al. [22] investigate the impact of the use of computer game elements, without the presence of the game, on teachers and list the following series of issues that are analysed through their research:

- 1) Do teachers in higher education show a positive attitude towards gamification?
- 2) Do teachers in higher education use gamification in their courses?
- 3) Does gender influence teachers' attitudes regarding the application of gamification in higher education?
- 4) Does gender of teachers affect the very application of gamification in higher education?
- 5) Does age affect teachers' attitudes regarding the application of gamification in higher education?
- 6) Does age of teachers affect the very application of gamification in higher education?
- 7) Does the type of educational institution influence the attitudes of teachers regarding the application of gamification in higher education?
- 8) Does the type of educational institution affect the application of gamification in higher education?

The research was conducted through an online survey, and the teachers were encouraged to forward the survey to their colleagues via a web link. A total of 98 teachers participated, with an average age of 44 years. In the sample, 56.25% of teachers were males.

Research shows that teachers who participated in the research (Q1 and Q2) had a positive attitude regarding the use of gamification in teaching, but only 18 out of a total of 98 teachers (11.30%) apply gamification in practice. The result is expected given that the term is quite new and insufficiently documented in the professional and scientific literature.

The answer was also negative to the third question: "Does age affect teachers' attitudes regarding the application of gamification in higher education?", as well as to the fourth question: "Does the gender of teachers affect the very application of gamification in higher education?" Gender does not affect attitudes regarding gamification as well as the application itself. The answers are based on the analysis and calculation of a t-test in which there was no statistically significant difference between participants.

Regarding the fifth question: "Does age affect the attitudes of teachers about the application of gamification in higher education?" As well as to the sixth question: "Does the gender of teachers affect the very application of gamification in higher education?", the answer was negative. Age is not related to attitudes or the application of gamification. In the analysis, the authors did not find a significant positive correlation regarding respondents' age.

Regarding the seventh question: "Does the type of educational institution influence the attitudes of teachers regarding the application of gamification in higher education?", the answer was positive. It has been shown that teachers of private higher education institutions have a more positive attitude towards the application of gamification, in contrast to teachers of public higher education institutions. Regarding the eighth question: "Does the type of educational institution affect the application of gamification in higher education?", the answer was negative. The results indicate the absence of statistically significant differences in responses between private and public educational institutions.

It is stated that there is a positive impact of the use of computer game elements on teachers of higher education institutions, but the application of this approach in practice is missing. Only 11.30% of respondents used gamification methods or educational computer games in their work environment. It is concluded that the age, gender, and type of educational institution do not statistically significantly affect the attitudes of teachers. The main problems that teachers face are related to information and communication technology and IT solutions that exist in a particular work environment, as well as the possibility of proper and timely teacher education. This is due to the lack of time for education and the lack of financial resources. Additional research in this area is suggested.

Good to mention is the relationships between the most common mechanisms of computer games and players

(system users) that serve teachers, designers and / or administrators of gamified education systems as recommendations in the planning phase. This topic mentions the types of players where the division cited by Richard Bartle is most used [23, 24]. If users are accessed through the Bartle matrix, then there are four groups of players to which system users can belong, and these are Winner, Collector, Socialite, and Explorer. It is necessary to know that not every one of the listed players is interested in every activity that is in the system, and accordingly, one should take care of mechanisms of computer games that are built into the educational system.

4 CONCLUSION

This paper presents a number of studies that include research of the higher education population in an online or offline environment. Guidelines are provided on how teaching staff should adapt their teaching material. Teachers should ensure the use of badges and points and enable the realization of certain achievements and the display of status. It is advisable to use a leader board, but also a top list of the top scoring students. Feedback to students must be available at all times. It is also useful to include occasional virtual meetings in the e-course to get students back into the system independent of the teaching materials. The gamified system should be as socially oriented as possible. Sociability is important for the type of players seeking recognition from other e-course participants. It is good to include in the e-course additional teaching and non-teaching elements (bonus) that are made available to students in special circumstances. The reward should be aligned with activity difficulty and activity difficulty should be built gradually.

An average teacher, who uses elements of computer games in educational process, should have advanced knowledge of computer tools for managing elements of computer games, especially if working with large groups of students. It takes a lot of energy to explain a new system to students, as well as getting approval from an organization that is used to working differently. An additional obstacle may be the lack of technical support of a higher education institution, which is not offered until the stated approach shows positive results in a certain situation and is accepted by a larger number of teachers.

When creating a gamified e-course, teachers, designers and / or administrators of gamified education systems should provide a minimum of one week for reviewing the elements of computer games included in the e-course, and one day for creating teaching materials of one teaching unit. It is necessary to set aside two hours for question analysis for one teaching unit, two days for entering the results of each knowledge test, and one week for pass rate analysis. Educational systems supported by elements of computer games will be increasingly used in the years to come due to the growing number of researchers and users of gamification around the world, especially considering the situation with the global health situation. Limitations of this study lies within one STEM field. It would be interesting to see how gamification could enhance similar university courses.

5 REFERENCES

- [1] Strmecki, D., Bernik, A., & Radošević, D. (2015). Gamification in E-Learning: Introducing Gamified Design Elements into E-Learning Systems. *J. Comput. Sci.*, 11(12), 1108-1117. <https://doi.org/10.3844/jcssp.2015.1108.1117>
- [2] Bernik, A., Radošević, D., & Dvorski, J. (2020). Gamification after Almost a Decade: Is it Still Relevant? A Case of Non-STEM Hybrid E-learning University Course. *Journal of Computer Science*, 16(5), 626-631., <https://doi.org/10.3844/jcssp.2020.626.631>
- [3] Bernik, A., Bubaš, G., & Radošević, D. (2018). Measurement of the effects of e-learning courses gamification on motivation and satisfaction of students. *Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics*, May 21-25, IEEE Xplore Press, Opatija, Croatia, pp: 0806-0811. <https://doi.org/10.23919/MIPRO.2018.8400149>
- [4] Bernik, A. (2021). Gamification Framework for E-Learning Systems in Higher Education. *Tehnički glasnik*, 15(2), 184-190. <https://doi.org/10.31803/tg-20201008090615>
- [5] Landers, R. N. & Callan, R. C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. *Serious Games and Edutainment Applications*, London. https://doi.org/10.1007/978-1-4471-2161-9_20
- [6] Muntean, C. I. (2011). Raising engagement in e-learning through gamification. *The 6th International Conference on Virtual Learning ICVL*.
- [7] Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Engaging engineering students with Gamification. *The 5th International Conference on games and virtual worlds for serious applications, VSGAMES 2013*, pp. 24-31, Bournemouth, UK. <https://doi.org/10.1109/VS-GAMES.2013.6624228>
- [8] Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving Participation and Learning with Gamification. *ACM International Conference Proceeding Series*, 9-16. <https://doi.org/10.1145/2583008.2583010>
- [9] Jayasinghe, U. & Dharmaratne, A. (2013). Game Based Learning vs. Gamification from the higher education students' perspective. *IEEE International Conference on Teaching, Assessment and Learning for Engineering*, University of Colombo School of Computing, Sri Lanka. <https://doi.org/10.1109/TALE.2013.6654524>
- [10] Ortega, de M. L., Domínguez, A., & Saenz-de-Navarrete, J. (2014). An empirical study comparing gamification and social networking on e-learning, Computer Science Department, University Alcalá, *Computers & Education*, 75, 82-91. <https://doi.org/10.1016/j.compedu.2014.01.012>
- [11] Iosup, A. & Epema, D. (2014). An Experience Report on Using Gamification in Technical Higher Education. *CM Technical Symposium on Computer Science Education*, pp. 27-32, USA. <https://doi.org/10.1145/2538862.2538899>
- [12] Laskowski, M. & Badurowicz, M. (2014). Gamification in higher education: a case study. *Human Capital without Borders: Knowledge and Learning for Quality of Life*, Portorož, Slovenia.
- [13] González-González, C. S. & Carreno, A. M. (2014). Methodological proposal for Gamification in the Computer Engineering Teaching. *International Symposium on Computers in Education*, University of La Laguna, Tenerife, Spain. <https://doi.org/10.1109/SIE.2014.7017700>
- [14] Stokes, Z. (2014). Integration of Gamification into the Classroom and the Reception by Students. Marshall Digital Scholar: Theses, Dissertations and Capstones, Paper 8562014.
- [15] Deterding, S. (2015). The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction*, 30, 294-335. <https://doi.org/10.1080/07370024.2014.993471>
- [16] Hamari, J. & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale. *Computers in Human Behavior*, 40, 133-143. <https://doi.org/10.1016/j.chb.2014.07.048>
- [17] Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? - A Literature Review of Empirical Studies on Gamification. *The 47th Hawaii International Conference on System Sciences*, pp. 6-9, USA. <https://doi.org/10.1109/HICSS.2014.377>
- [18] Hamari, J. & Parvinen, P. (2016). Introduction to Gamification: Motivations, Effects and Analytics. *The 49th Hawaii International Conference on System Sciences*, Koloa, USA. <https://doi.org/10.1109/HICSS.2016.165>
- [19] Smole, K. S., Diniz, M. I., & Milani, E. (2007). Math games from 6 to 9 grade, Math computer games. New York: Guilford Press, USA.
- [20] Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society*, 18(3).
- [21] Schreuders, Z. C. & Butterfield, E. (2016). Gamification for Teaching and Learning Computer Security in Higher Education. *USENIX Workshop on Advances in Security Education*, Austin, Texas.
- [22] Parreno, J. M., Mas, D. S., & Mas, E. S. (2016). Teachers' attitude towards and actual use of gamification. *The 2nd International Conference on Higher Education Advances*, València, Spain, *Social and Behavioral Sciences*, 228, 682-688. <https://doi.org/10.1016/j.sbspro.2016.07.104>
- [23] Bartle, A. R. (1985). MUD Advanced Project Report, an internal memorandum (CSM-73). *Department of Computer Science's series*, University of Essex, UK.
- [24] Bartle, A. R. (1996). Players Who Suit MUDs, a refereed paper. *Journal of MUD Research*, 1(1).

Authors' contacts:

Andrija Bernik, PhD, Assistant Professor
University North,
104. brigade 1, 42000 Varaždin, Croatia
andrija.bernik@unin.hr

Damir Vusić, PhD, Full Professor
University North,
104. brigade 1, 42000 Varaždin, Croatia
damir.vusic@unin.hr

Voravika Wattanasoontorn, PhD, Lecturer
Prince of Songkla University,
80 Vichitsongkarn road Kathu Phuket, Thailand
voravika.w@psu.ac.th

A Working from Home (WFH) 2022 - Best Practices Implementation and Case Study of Company Agreement

Stefan Schmidt*, Benjamin Stefan Godwin Schmidt

Abstract: While four per cent of employees in Germany worked from home before the covid crisis, this figure was around 30 per cent in the first lockdown in April 2020. Now working from home and mobile working are part of the everyday life of many employees. Experience has shown that many activities can also be carried out from other locations than the classic office in the company. Location-flexible work will also still be standard in some sectors after the crisis. This article shows employees' experiences with working from home in corona times, with serious advantages and disadvantages being reported. Best practices as implementation of company agreements are presented based on an Original Equipment Manufacturer (OEM) example: Example of working from home (WFH) / office work, life-phase oriented working time instruments as a personnel management success factor, promotion of cross-border cooperation, efficient use of resources, questions on learning from the pandemic, working through the corona period with employees, recommendations for surviving this crisis and culture of trust versus control.

Keywords: best practices; company agreements; case study; covid crisis; employees' experiences; working from home

1 ARTICLE POSITIVE AND NEGATIVE RESULTS OF WORKING FROM HOME

In the case of working from home, the place of work is defined and restricted to the employee's own home. Mobile working describes all forms of location-flexible working. Working from home became a daily reality for many during the covid crisis. The Hans Böckler Foundation has been researching the opportunities and risks of mobile working for many years [1].

Corona was a catalyst for mobile work. The percentage of employees working from home increased significantly from just four percent before the outbreak of the pandemic to 27 percent in April 2020, dropping to just 14 percent in November 2020 during the "lockdown light" period. Many employees with work-from-home jobs were encouraged to work face-to-face. At Siemens, for example, 140,000 employees could work from home for two to three days a week with the approval of their supervisor.

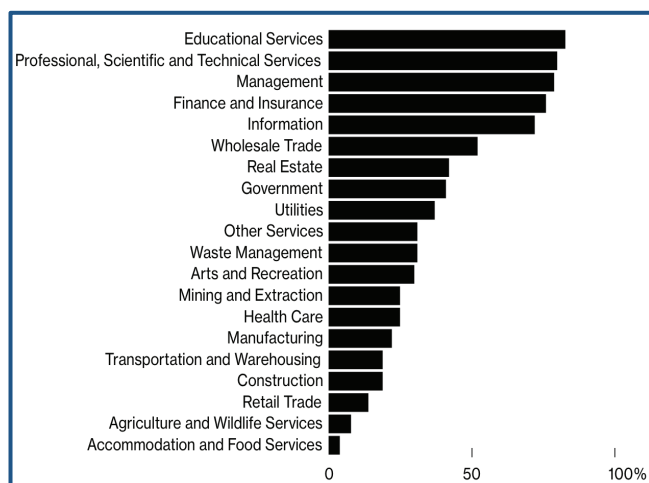


Figure 1 "Work from Home" Works for Some, Percentage of jobs that can be done at home, by industry [2]

Working from home will be more prevalent in the future according to 71 % of respondents to the June survey. However, it must be recognised that working at home also comes with risks, such as psychological overload, loneliness, or career disadvantages. The most significant disadvantage of working from home among respondents in Germany is the lack of direct contact with colleagues. In 2020/2021, 75/74 percent agreed with this statement. The risks can be averted with clear company rules and compliance with the necessary framework conditions [1].

In companies with co-determination, employees report an above-average number of positive experiences with working from home. Almost half of those surveyed, who work from home, would also like to work from home in the future. This article shows the experiences of companies and employees with working from home in corona times, with serious advantages and disadvantages being reported.

This article shows the experiences of companies and employees with working from home in corona times, with its reported significant advantages and serious disadvantages and negative impact on health.

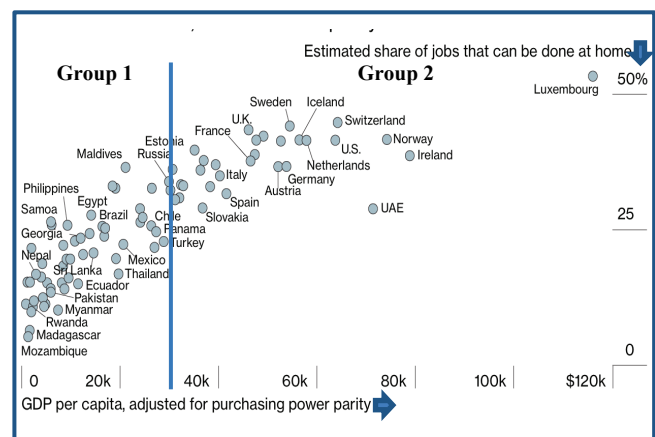


Figure 2 WFH Is for the Rich, Where countries rank in terms of wealth and capacity for remote work [2]

Administrative, organisational and information-based activities are better suited for working from home. In contrast, the creation and distribution of products as well as some services are less suitable for working from home to a greater or lesser extent, ranging from 'Accommodation and Food Service' with four percent to 'Educational Services' with 83 percent (see Fig. 1) [2].

WFH is mainly a topic of wealthy countries. Up to a Gross domestic product (GDP) of \$30,000, only countries from Asia, Africa and Latin America are in 'Group 1'. Above a GDP of \$30,000, only countries from Europe, the U.S. and the United Arab Emirates are in 'Group 2' (see Fig. 2) [2].

1.1 Positive: Is Your Boss Ending Remote Work? As a Chief Executive Officer (CEO), I tell You Why They are wrong

The 200 employees of the credit card processing company Gravity Payments have been working mobile since the beginning of the pandemic. Gravity Payments CEO Dan Price cut his salary by one million dollars in 2015 to enforce a minimum wage of 70,000 dollars in the company. For him, the shift to mobile working turned out to be very positive [3].

The shift to mobile working can be life-changing for employees. Buying a house and raising a family in cheaper areas is easier and employees are given the opportunity to decide how they want to work. Daily commuting requires a lot of time and incurs high costs. In addition, by avoiding commuting, companies make a contribution towards environmental protection. When commuting in the morning, employees experience more negative feelings than at other times of the day [4]. How can productivity be increased? According to nationwide research, mobile working led to a 5% increase in productivity [2], mainly because employees were not burnt out from commuting. Companies had been having it too easy for too long. Commuting was demanded by the company managers, who were in control.

Instead of the Chief Executive Officer (CEO) making the decision alone, the Gravity Payments employees were asked how they would like to work:

- 7 % wanted to return to full-time office,
- 32 % wanted a mix of office/telecommuting,
- 62 % wanted to work from home full-time.

Many companies need to work face-to-face, for them face-to-face interaction is a necessity. For Gravity Payments, office work makes less sense with regard to fostering innovation and increasing productivity. Employees know their work better and have shown in the last year that they can do their work just as well or better from home. Therefore, they can be allowed to decide for themselves how they want to work. It is advantageous for some companies to leave this decision to their employees. Mobile work also avoids additional short-term costs (such as turnover) as well as long-term costs (such as low morale).

1.2 Post-Covid Work Patterns must not be imposed by Bosses with an Eye on Cost

The pandemic sparked more discussions about the future of work than ever before [5]:

Some employers expected employees back in the office after the lockdown, while others welcomed the prospect of new ways of working and to this end invested heavily in mobile working technologies. Although companies are able to adapt quite quickly to the widespread shift to mobile working, this will, in contrast, have a negative impact on the development of shopping streets, homes and business parks built in an earlier era and will inevitably lead to change. Some governments are considering making mobile work the default option by law. Hybrid working (workers are in the office part of the week) seems to be becoming the norm, at least for workers in large companies.

For individuals, the impact will be profound and long-lasting. Workers' representatives noted a growing gap between a privileged elite of more flexible working professionals of large employers and those workers, who do not have the same flexibility and are required to work according to the demands of their employers, for example, in different shifts at the workplace. Workers and the growing number of workers in the gig economy will be even more bound to the shift schedules dictated by their employers in the workplace. The flexibility mainly benefits the employer with arbitrary decisions more likely to be made based only on the need to cut costs and increase productivity in the short term. However, it is essential that employment practices are developed in discussion with employees rather than dictating standard options from top management, as the new working patterns may otherwise develop in a misguided and costly way to the disadvantage of all concerned [6].

1.3 Negative: The Evidence is in: Working from Home is a Failed Experiment

People working from home are more stressed and less happy according to study results [7]. This is also supported by the results of a large study by Microsoft with more than 31,000 full-time or self-employed workers in 31 markets [8]:

- Nearly 2/3 of workers "crave" more face-to-face time with their teams.
- 37% of the global workforce complains that their companies "demand too much" when they are not in the office.
- Some 54% of these people feel overworked, 39% are simply exhausted.
- With mobile working, meetings are significantly longer, "chats" are up 45%.
- In February 2021, 41 million more emails were sent than in the same month last year.
- Older workers and bosses seem to have things under control (61% are doing well at the moment).
- Those without decision-making power: only 38 % are doing well at the moment.
- Younger generations - especially those aged 18 to 25: they have difficulty balancing work and their lives, have

problems to feel engaged or excited about work and are more exhausted than their office-based colleagues.

- Even worse is the death of innovation.
- According to Microsoft, companies are even more isolated from each other than before the pandemic.
- Interactions in their own networks are more common, but these close team interactions diminish over time.
- When connections are lost, people are less innovative and group-think increases.

Other companies make similar statements [7]: Citigroup introduced a series of "Zoom Free" Fridays to get employees away from their screens. LinkedIn employees were given a paid week off to deal with the stress of mobile work. According to Chief Executive Officer of Goldman Sachs, home working is an aberration that needs to be corrected as soon as possible. A client wants their employees back in the office to see them, talk to them, brainstorm and connect.

The trend towards working from home is unfortunately too strong to resist. There is an illusion of more independence, flexibility, and control over one's life. Of the employees, who responded to the Microsoft survey, 70 per cent want some kind of flexible working options in the future.

However, big tech companies such as Microsoft, Google, Amazon, Apple have powerful interests in mobile work. They want to tie their customers more closely to their cloud models. Huge media providers expect more traffic to their sites. Large retailers expect an increase in their online sales. Furthermore, savings in real estate costs as well as an offer of better (cheaper) talent can be expected [7].

According to studies [9, 10], people in home offices are more stressed and less happy. Mobile working leads to more disruption, less productivity and reduced innovation:

- 70% of Americans working from home during COVID-19 are struggling to maintain a healthy work-life balance.
- 65% are working longer hours than ever before
- 56% are more stressed about work than ever before
- 63% feel their company discourages them from taking time off
- 67% feel pressured to be available 24/7 while WFH

What is stressful about working at home [10]? In the home office, the boundaries between work and private life become blurred. Problems arise regarding resting one's mind / with taking a break. Employees are surprised by the stress once the novelty of working from home wears off and they realise the challenges. Common sources of stress faced by many home-based workers: lack of structure, too many distractions, social isolation, less physical effort, and difficulties with setting boundaries. The Work Trend Index 2021 Annual Report (The Next Great Disruption is Hybrid Work –Are We Ready?) represents the following trends [8]:

1. Flexible work is here to stay
2. Leaders have lost touch with employees and need a wake-up call
3. High productivity masks an exhausted workforce
4. Generation Z is at risk and will need to be re-energized
5. Shrinking networks are a threat to innovation

6. Authenticity will drive productivity and wellbeing
7. Talent is everywhere in a hybrid world of work

Mobile work can be a good and productive benefit for employees. Studies show that Millennials, who make up about half of the workforce, want more mobility and independence. But face-to-face contact in the office is also very important and working two to three days from home seems to be the ideal compromise [11].

As a result, 41 percent of workers are considering leaving their current employer this year and 46 percent are likely to relocate because they can now work remotely.

Amazon's shipping centres, on the other hand, are a different world: Welcome to a dystopia where Amazon employees are fired via an app [12] and Amazon.com founder Jeff Bezos seems to believe all workers are inherently lazy. Amazon's systems encourage high turnover and fear and despair; employees are suddenly mysteriously fired, treated as disposable entities.

1.4 Is Remote Working better for the Environment? Not necessarily

There are arguments for as well as against the environmental impact of mobile working. Before long, four times as many people could be working from home compared to before Covid. Mobile work is not necessarily better for the environment, while hybrid working could lead to a "worst case scenario". [13]. Companies now have a rare opportunity to redesign work models. Post-pandemic work does not necessarily have to take place in the office. According to the intuitive assumption, mobile work is de facto better for the environment. But the climate models of mobile and office work are complex [13]:

- Transportation is the largest contributor to greenhouse gas emissions in the United States.
- Mobile work will reduce emissions from the daily commute.
- More energy is used to run air conditioning or heating in homes.
- According to the International Energy Agency (IEA), in some cases the average residential weekday electricity use increased by more than 20 percent during lockdown.
- When employees use public transportation and short commutes, their total emissions may even be higher when they work from home.
- Sprawling suburban homes use three times more energy on average.
- Saved emissions from the daily commute are more than offset by flights to quarterly meetings.

According to the Carbon Trust and Vodafone's Homeworking report [14], working from home leads to a reduction in carbon emissions in six countries. Office emissions account for the largest share and the savings achieved by avoiding commuting are only secondary.

The International Energy Agency (IEA) notes that in an average year, the total energy saved because of the reduction

in commuting is still about four times the increase in household energy consumption due to working at home. The longer-term impact of a trend toward more home-based work on energy and emissions is uncertain [15].

2 HEALTH EFFECTS – WORKING WITHOUT END – RISKS

Remote work and then the pandemic of increased working from home leads to health impacts on people, who are inherently social beings. A study back in 2017 by the International Labour Organization on the impact of working remotely found, based on data from 15 countries, that although workers are more productive outside of a traditional office, they also show higher levels of employee stress and insomnia. This is associated with the risks of "longer working hours, higher work intensity and overlap between work and home. Two to three days working from home seems to be the ideal scenario [16].

Numerous statistics (number 50) on the impact of the corona virus (COVID-19) on the psyche are presented with key figures relating to the corona pandemic, the situation before its outbreak and the emotional state, fears, concerns, stress, psychological distress, and loneliness that people have experienced [17].

According to the Nationale Kohorte (NAKO) Health Study, the Corona pandemic leads to greater psychological distress. Young adults and middle-aged adults reported significantly more severe symptoms of anxiety, stress, and depression. With about 200,000 participants, the Nationale Kohorte (NAKO) Health Study is the largest long-term study in Germany and unique in Europe in its dimension. It is conducted jointly by universities and research institutes. The participants, who are between 20 and 69 years of age, have been studied nationwide since 2014 and form a representative cross-section of the population [18, 19, 20].

In the "Health after COVID-19" multidisciplinary study in Tyrol/South Tyrol, one fifth of the participants reported a poorer quality of life following a SARS-CoV-2 infection and recovery at home. Mental stress is the number one risk factor for psychological complaints after COVID-19, with an increase in depression and anxiety disorders [21 22].

Cases of mental disorders rose sharply during the Covid-19 pandemic. Younger people and women were hardest hit with depressive and anxiety disorders globally in 2020 [23].

Many life trajectories are set in place during adolescence and young adulthood. The COVID-19 pandemic has affected adolescents and young adults across several life domains. The pandemic and response have negatively affected their social connectedness and has led to a worsening of their psychological distress [24].

United Nations Children's Fund (UNICEF) State of the World's Children 2021 report on mental health describes the mental distress caused by Covid-19 as only the "tip of the iceberg". Even before that, a significant proportion of girls and boys were mentally distressed. Children and young people could feel the impact of the Covid-19 pandemic on their mental health and well-being for many years to come. One in seven young people between the ages of ten and 19 are estimated to be living with a diagnosed mental health condition or disorder [25].

According to UNICEF's 'State of the Children in Germany 2021' report, multiple constraints on public and private life not only jeopardise young people's educational achievements, but also have far-reaching consequences for their overall well-being and development [26]. In principle harm caused by Covid-19 in relation to age is shown in Fig. 3 [Author].

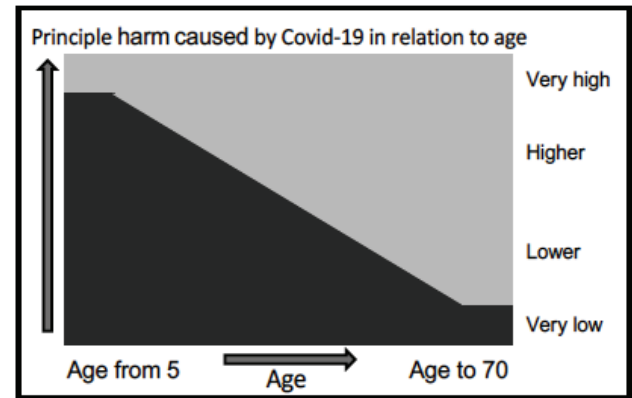


Figure 3 Harm caused by Covid-19 according to age [Author]

The impairment of employees by Covid-19 in relation to sustainable productivity is shown in Fig. 4 [Author].

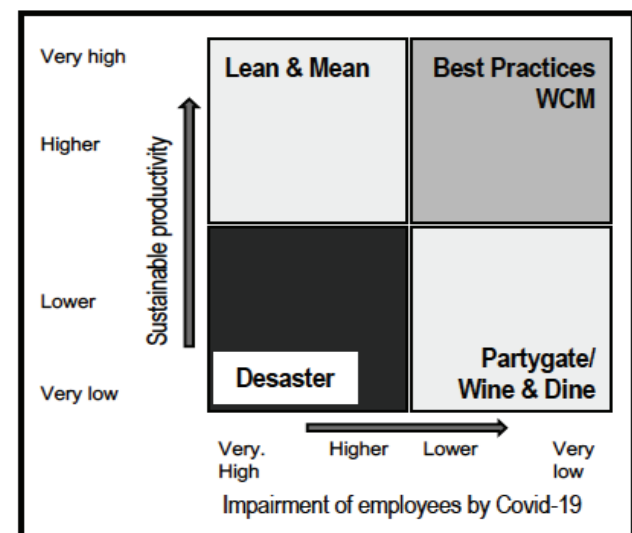


Figure 4 Sustainable productivity versus the impairment of employees by Covid-19 [Author]

The unlimited world of work with "flexible" being the new "productive" can also lead to the danger of working without end with constant accessibility and insufficient non-accessibility. Misuse is also possible, as the example of a call centre used by British companies shows. It is accused of monitoring home workers via teleperformance, providing biometric and medical data and using lie detectors [27].

Pandemics have changed the world with diseases that are now known to be transmitted from farm animals to humans, such as through industrial chicken farming. Therefore, risk management and appropriate home office management must be implemented for pandemic readiness. Many long-term effects of COVID-19 are not yet known and, like the impact

and extent of surveillance software for remote workers, need to be carefully monitored and further researched.

3 COMPANY AGREEMENT ON MOBILE WORK

3.1 Example of Company Agreement on Mobile Work [28]

The company agreement on mobile work concluded between the employee representatives and the management on mobile working is intended to achieve a number of balanced goals as well as improved, more flexible working practices and to promote health preservation. The high flexibility is also particularly important for the overcoming of crisis and managing future risks.

The compatibility of private and professional life is promoted through mobile working. The flexibility of working hours and of the place of work will be increased. The extent of constant accessibility and non-accessibility is fixed. All professional activities can be carried out externally both online and offline. An allocation of different work locations and times of day combined with part-time office work is possible. Mobile work is possible for the whole day or for part of it or in combination with office work or leisure time. It can be undertaken by employees, provided their assigned tasks are appropriate for mobile work, and by managers. Mobile work is carried out voluntarily with the agreement of the company management. The frequency and distribution of office and mobile work are determined by means of a regulated exchange of information.

There are several examples of the possible allocation of mobile work and/or office work:

- Exclusively office work.
- Exclusively mobile work.
- Working day divided into two parts combining office work and mobile work.
- Part-time employees performing mobile work and/or office work.
- Working day divided into three parts combining mobile work, office work and finally mobile work
- Any combination of mobile work, office work and free time.

Mobile work is not only online work on the PC/laptop, but also work offline, using phone, mobile/smartphone or on paper. The necessary work and communication tools are provided by the company. Mobile work can also take place without access to the company network. The provision of a statutory accident insurance or considering mobile work as a self-employed activity must be considered. In the event of an accident, a regular accident report must be made. The design of the workplace must comply with the usual ergonomic regulations. Liability for work equipment exists only in cases of intent and gross negligence.

3.2 Implementation of Company Agreement on Mobile Work

The example of an automobile manufacturer is used to present its implementation of the company agreement on mobile working as part of the life-phase-oriented working time instruments and a human resource management success factor. The company not only wants to expand its leadership in technology, but it also wishes to underpin its employer

attractiveness by improving working conditions through innovative and individual working time solutions: Mobile work - Flexible working, consciously switching off [28, 29].

Mobile working is a progressive working time instrument. It allows a flexible organisation of the work location and working time across national borders and time zones. Open dialogue with an individual agreement between manager and employee allows maximum freedom for customised solutions. Working time and place of work can be optimally adapted according to the individual employee's phase of life. Mobile working means a culture of trust and is based on constructive dialogue rather than rigid guidelines. The responsible handling of unlimited communication in the age of mobile and smartphone is supported. To this end, the topic of availability versus unavailability is evaluated as a focal point: "Work flexibly, switch off consciously". Possible qualifications support the dialogue in the company. Advantages of mobile working [29] are:

- Contribution to work-life balance,
- Promoting the long-term health and performance of employees,
- Increasing the competitiveness of the company.

Mobile work promotes the efficient use of resources in cross-border cooperation [30]:

- **Time** - through the avoidance or use of waiting and commuting times.
- **Location** - through the flexible use of different work locations and the private home office, savings in travel and operating costs are made possible. The conditions for new concepts for offices with the option of desk sharing are created.
- **Personal performance** - through the design of an individually optimal working environment and the consideration of personal biorhythms.

Examples of the share of content-related performance in total weekly working time are shown in Fig. 5.

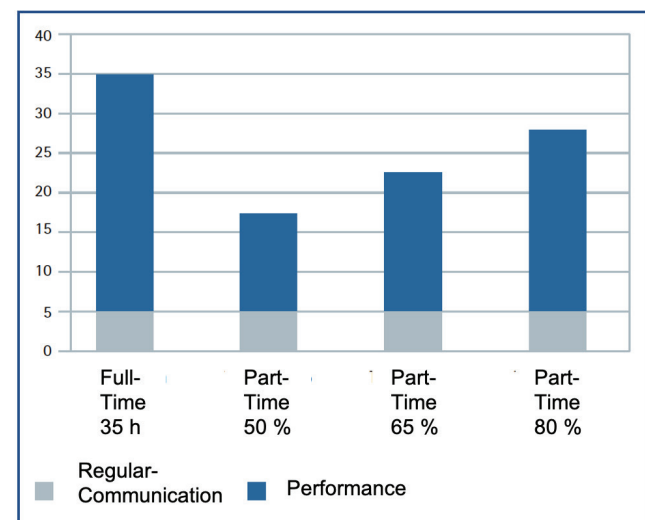


Figure 5 Share of content-related performance in total weekly working time [30]

To intensify the implementation of the company agreement, the managers are measured by the quota of their employees in mobile work. The chairperson of the works council wants a cultural change in the company [28]:

A company agreement on mobile work was already adopted before the pandemic. With this flexibility, for example during the day could also emails be processed off the office. This is working time as is working on the train. Mobile working is one step further than home office. However, an automobile company is not a commercial bank or insurance company, but an industrial company. Complex production plants or measuring machines cannot be operated at home. Employees from production have to be on site. The same applies to planning engineers, software engineers or auditors. Generally applicable regulations do not make sense. Where possible, a reasonable mix is strived for. For example, three days in the office, two days at home are preferable.

The existing company agreement on mobile working has not been redrafted. What is needed is a corresponding cultural change of mindset, not of documents. This is a question of leadership behaviour and drawing up appropriate criteria. There is consensus on sick leave: high levels of sick leave are a sign of poor leadership. A low share of mobile work should be understood as a sign that the leader has not yet understood the possibilities of digitalisation. Good leaders have a high proportion of mobile work, bad ones a low one. With this recognition, change can then come quickly. There are no rules set by the works council for this cultural change.

The monthly ten-minute video podcasts by the works council chair should contribute to this change. Because of Corona, no large works meetings could be held. In these podcasts, the issue of leadership culture and mobile working is clearly addressed. This is also communicated at the works meetings. These measures have already had an impact internally: managers will be evaluated on how they deal with this issue. As a result of the additional home offices, staff equipment has also been improved. With the provision of notebooks, all employees were equipped with appropriate technical equipment, which may also be used privately. More importantly, all company trainees have also been provided with notebooks.

Another effect of home office is the reduced need of companies for office space. If employees work two out of five days a week in all-day mobile work, this has a considerable impact on local transport, office space in the company or car park occupancy. The two days working from home also led to 40 percent less carbon dioxide CO₂ pollution from commuting. Federal Labour Minister Hubertus Heil plans to establish working from home permanently in Germany. In future, employees will be able to work from home 24 days a year if there are no operational reasons to the contrary [31]. The automobile company in question has already made a lot of progress in their goal to have mobile working two days a week. Any problems that arise are solved in consultation with the responsible department heads and board members and are not discussed publicly. This corporate culture is the more successful way [28].

3.3 Trust Creates Competitive Advantages - Culture of Trust versus Control

Trust plays a key role in strategic management as a 'cost-free' resource. Trust [32]:

- opens up potential that is otherwise unavailable,
- is essential in situations such as the home office,
- is a competitive advantage especially in times of shortage of skilled labour,
- is often the decisive differentiating factor in securing a company's future,
- promotes the development of competencies,
- promotes identification and long-term loyalty to the company,
- is the only alternative when no control is possible
- and promotes employee loyalty as a complex, long-term valuable resource that is difficult to replace.

Trust is, by definition, a 'risky advance' (Luhmann 2010). By taking a leap of faith, team members are assumed to have a fundamental will to perform, with the new contexts of home office and digitalisation inevitably lead in this crucial direction anyway. Moving forward with trust requires courage above all. Complete control is pure illusion!

4 WORKING THROUGH ISSUES WITH EMPLOYEES.

4.1 Review Learnings from the Pandemic/Remote Work with Questions.

It is recommended to work through the problems encountered during the pandemic and the lessons learned from the pandemic and the resulting increase in mobile work by asking questions of the staff [28]:

- What worked well in leadership during the lock-down?
- What did not work?
- How was closeness established with the team? What should be retained from this?
- What maintained performance in the team? What of this would be useful to continue?
- Where can leadership continue to build on the self-responsibility of its employees?

One thing is certain: the new normal needs empathy and mutual understanding, because post-Corona life is fragile. The next virus outbreak in the family, at school, day care or sports club shatters the newly won freedom and drives those affected back into the uncertainty of a quarantine.

4.2 Review with Employees the Corona Period/Remote Work with Questions

Managers must provide for the possibility of returning to mobile work at any time. Collaboration rather than competition is important [28].

- What communication and meeting routines do we keep, what do we need to do differently?
- Which tools have proven themselves and where are there still gaps?

- Which competencies have we developed and where do we continue?
- Which processes and decision-making channels have we reorganised and how do we proceed?

However, the transition to the new normal also brings with it the potential danger of falling back into old patterns. Much of what has been achieved in the last two years points in the direction of the future.

Recommendations to help the leadership and the team get through this crisis [28]:

- Drop the 9-to-5 mentality.
- Allow staff to create their most productive schedules.
- Get organised.
- Shorter working days are fine.
- Provide written updates instead of check-in calls.
- Being a good person makes you a better leader.
- Transparency in communication.
- Provide support for the "employee, who is down".

Communication, flexibility of mobile working and especially the extensive freedom of employees to decide on their working methods are particularly important for overcoming the crisis and managing future risks.

5 CONCLUSION

The company agreement on mobile work between the employee representatives and management on mobile working goes beyond working from home and proved its worth in the pandemic. The company has come a long way, with the implementation of the company regulations now tailored to the employees and the company. Voluntary mobile working online/offline has achieved many balanced goals and improved flexible working in a sustainable manner: improving the work-life balance with the regulation of both accessibility and non-accessibility, improving employees' long-term health and performance, increasing the company's competitiveness through flexible working hours and location, and enhanced resilience to crises. Employees create their own most productive schedules, whereby their equipment requirement for working at home is basic.

Mobile working requires a culture of trust and is based on constructive dialogue rather than rigid guidelines. As a 'cost-free' resource, trust plays a key role in strategic management. Achieving the required cultural shift in mindset is a question of leadership behaviour and appropriate criteria. This change is supported by the assessment of how leaders deal with this issue. Problems are solved between board members and employee representatives in dialogue and not discussed publicly. A culture of trust and cooperation rather than control is the more successful way forward for the company.

Mobile work can be performed by employees, for whom tasks appropriate for it can be assigned, and by managers. Production staff, on the other hand, must be on site. The expansion of mobile work requires a suitable corporate and trust culture, supportive leadership behaviour and embedding

in the life-phase-oriented working time instruments of human resource management, whereby health management and pandemic considerations must be included in the risk management process.

The WFH-literature shows that there are not only many significant advantages in working from home, but also serious disadvantages; for example, employees, who work from home, are more productive, but they also show negative health impacts such as a higher level of employee stress.

COVID-19 (Coronavirus) can cause damage to the heart, lungs, and brain, which brings with it an increased risk of long-term health problems. These effects can sometimes remain for months or may even be permanent, which will in turn increase the sickness rate for companies as well as health care costs. Many long-term COVID-19 effects and how they will impact those affected over time are still unknown, but research is necessary and ongoing [33].

Finding the right balance between on-screen and in-person contact is a delicate problem, as is finding the right balance between in-office work and working from home. At the moment, we do not know how to best quantify and develop this new situation, in order to achieve the optimum balance for both employee and employer.

Occupational health management needs reliable data for appropriate health measures and for preventive medicine. In Germany, medical data are considered to be highly sensitive and confidential and, although a more convenient and quicker means of communication, are therefore never sent by email, but by regular mail only; in addition, employers do not have the right to request any medical information relating to their employees. This makes it very difficult for employers to make inquiries about and obtain information on their employees; for this, they must first obtain permission from the unions, which takes a very long time.

In recent years, the use of technology to monitor the work force has sharply increased. The software - 'Bossware' - affects almost every employee, who may not even realize that they are being watched. This computer monitoring software supports companies, who wish to spy on their employees, for example, to measure their productivity, which often takes place without the employees' knowledge or approval [34].

A survey of 1,250 US employers with remote employees found that 60% of them are using work monitoring software, which is generally used to track web browsing and application use. 88% of employers said they had terminated workers' employment after implementing monitoring software [35]. The expansion of this so far unregulated market involving such products erodes labor standards for workers and exploits weak labor protections.

The effect and extent of monitoring software for remote workers needs careful observation and further research.

Notice

The paper was presented at MOTSP 2022 – 13th International Conference Management of Technology – Step to Sustainable Production, which took place in

Primošten/Dalmatia (Croatia) on June 8–10, 2022. The paper will not be published anywhere else.

6 REFERENCES

All documents retrieved March 25 2022.

- [1] Hans.Böckler-Stiftung (2021, February 16). Studien zu Homeoffice und mobiler Arbeit. Retrieved from <https://www.boeckler.de/de/auf-einen-blick-17945-Auf-einen-Blick-Studien-zu-Homeoffice-und-mobiler-Arbeit-28040.htm>
- [2] Source National Bureau of Economic Research in: Curran, E. (2021, April 22). Work from Home to Lift Productivity by 5% in Post-Pandemic U.S. Retrieved from <https://www.bloomberg.com/news/articles/2021-04-22/yes-working-from-home-makes-you-more-productive-study-finds>
- [3] Price, D. (2021, July 7). Is your boss ending remote work? As a CEO, let me tell you why they are wrong. *The Guardian*, Retrieved from <https://www.theguardian.com/commentisfree/2021/jul/07/is-your-boss-ending-remote-work-as-a-ceo-let-me-tell-you-why-they-are-wrong>
- [4] Kahneman, D. & Krueger, A. D. (2006). Developments in the Measurement of Subjective Well-Being, *Journal of Economic Perspectives*, 20(1), 3-24. <https://doi.org/10.1257/089533006776526030>
- [5] Partridge, J. & Makortoff, K. (2021, June 18). Office, hybrid or home? Businesses ponder future of work. *The Guardian*, Retrieved <https://www.theguardian.com/business/2021/jun/18/office-hybrid-or-home-businesses-ponder-future-of-work>
- [6] Post-Covid work patterns must not be imposed by bosses with an eye on cost (2021, June 20). *The Guardian online*, Retrieved from <https://www.theguardian.com/business/2021/jun/20/post-covid-work-patterns-must-not-be-imposed-by-bosses-with-an-eye-on-cost>
- [7] Marks, G. (2021, April 8). The evidence is in: working from home is a failed experiment. *The Guardian*, Retrieved from <https://www.theguardian.com/business/2021/apr/08/working-from-home-is-a-failed-experiment>
- [8] The Work Trend Index: The Next Great Disruption Is Hybrid Work—Are We Ready? (2021, March 22). *Microsoft.com*, Retrieved from <https://www.microsoft.com/en-us/worklab/work-trend-index/hybrid-work>
- [9] SWNS (2020, September 3). The stress of working from home is getting to most Americans. Retrieved <https://nypost.com/2020/09/03/the-stress-of-working-from-home-is-getting-to-most-americans/>
- [10] Scott, E. (2021, April 7). 8 Tips to Handle the Stress of Working From Home. Retrieved from <https://www.verywellmind.com/the-stress-of-working-from-home-4141174>
- [11] Marks, G. (n. D.). Study: Working From Home Causes More Stress - For some remote employees, it's not such a great thing. Retrieved March 24, 2022, from https://www.inc.com/gene-marks/study-working-from-home-causes-more-stress.html?cid=sf01002&sr_share=facebook
- [12] Crispin, J. (2021, July 5). Welcome to dystopia: getting fired from your job as an Amazon worker by an app. *The Guardian*, Retrieved from <https://www.theguardian.com/commentisfree/2021/jul/05/amazon-worker-fired-app-dystopia>
- [13] Schupak, A. (2021, August 2). Is remote working better for the environment? Not necessarily. *The Guardian*, Retrieved from <https://www.theguardian.com/environment/2021/aug/02/is-remote-working-better-for-the-environment-not-necessarily>
- [14] Acerini, L., Bordat, S., Fitzpatrick, L., Jennings, T., Stephens, A. (2021, June). Homeworking report, *Carbon Trust and Vodafone Institute for Society and Communications*, Retrieved from https://www.vodafone-institut.de/wp-content/uploads/2021/06/CT_Homeworking-report-June-2021.pdf
- [15] Crow, D. & Millot A. (2020, June 12). Working from home can save energy and reduce emissions. But how much? *International Energy Agency*, Retrieved <https://www.iea.org/commentaries/working-from-home-can-save-energy-and-reduce-emissions-but-how-much>
- [16] Messenger, J., Vargas Llave, O., Gschwind, L., Boehmer, S., Vermeylen, G., & Wilkens, M. (2017, February 15). Working anytime, anywhere: The effects on the world of work. Retrieved from https://www.eurofound.europa.eu/sites/default/files/ef_publication/field_ef_document/ef1658en.pdf
- [17] Pawlik, V. (2021, July 7), Statistiken zu den Auswirkungen des Coronavirus (COVID-19) auf die Psyche. *Statista*, Retrieved from <https://de.statista.com/themen/8237/auswirkungen-der-corona-pandemie-auf-die-psyche/>
- [18] Aktuelle Meldungen 2020 (2020, December 4). NAKO Gesundheitsstudie: Stärkere psychische Belastung durch Corona-Pandemie. *Bundesministerium für Bildung und Forschung*, Retrieved <https://www.gesundheitsforschung-bmbf.de/de/nako-gesundheitsstudie-starkere-psychische-belastung-durch-corona-pandemie-12564.php>
- [19] Dragano, N., Reuter, M., & Berger, K. (2022, March). Increase in mental disorders during the COVID-19 pandemic—the role of occupational and financial strains. An analysis of the German National Cohort (NAKO) Study. *Dtsch Arztebl Int*, 2022(119), 179-187. <https://doi.org/10.3238/arztebl.m2022.0133>
- [20] Dragano, N., Reuter, M., & Berger, K. (2022, March 9). Zunahme psychischer Störungen während der COVID-19-Pandemie – die Rolle beruflicher und finanzieller Belastungen - eine Analyse der NAKO Gesundheitsstudie. *Dtsch Arztebl Int*, 119, 179-187. <https://www.aerzteblatt.de/archiv/223957/Zunahme-psychischer-Stoerungen-waehrend-der-COVID-19-Pandemie-die-Rolle-beruflicher-und-finanzieller-Belastungen>, online am 9. 3. 2022 (online first)
- [21] Hüfner, K. et al. (2022). Who Is at Risk of Poor Mental Health Following Coronavirus Disease-19 Outpatient Management? *Front. Med.*, 9, 792881. <https://doi.org/10.3389/fmed.2022.792881>
- [22] Mair, T. & David Bullock, D. (2022, March 15). Mentaler Stress ist Risikofaktor Nummer Eins für psychische Beschwerden nach COVID-19. *Medizinische Universität Innsbruck Aktuell*, Retrieved from <https://www.i-med.ac.at/mypoint/news/761234.html>
- [23] The Lancet (2021, October 8). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet*, 398(10312), 1700-1712. [https://doi.org/10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7)
- [24] AIHW Australian Institute of Health and Welfare (2021, June 25). Australia's youth: COVID-19 and the impact on young people. Retrieved from <https://www.aihw.gov.au/reports/children-youth/covid-19-and-young-people>
- [25] United Nations Children's Fund (2021, October). The State of the World's Children 2021: On My Mind – Promoting, protecting and caring for children's mental health. *UNICEF, New York*, Retrieved from <https://www.unicef.de/informieren/materialien/sowcr-2021/251192>
- [26] Bertram, H. (2021, April). Kinder – unsere Zukunft!! Der UNICEF-Bericht zur Lage der Kinder in Deutschland 2021. *Deutsches Komitee für UNICEF, Köln*, Retrieved from <https://www.unicef.de/informieren/materialien/kinder-unsere-zukunft-bericht-2021/239414>

- [27] Walker, P. (2021, August 8). Call centre used by UK firms accused of intrusively monitoring home workers. *The Guardian*, Retrieved from <https://www.theguardian.com/world/2021/aug/08/call-centre-uk-firms-accused-intrusively-monitoring-home-workers-teleperformance>
- [28] Premium automobile manufacturer (2019-2022), several documents
- [29] Quandt-Schubert, K. (2017, February 22). BMW: Mobilarbeit – Flexibel arbeiten, bewusst abschalten. In: Bartz, M., Gnesda, A., Schmutzer, T. (eds) Unternehmen der nächsten Generation. *Springer Gabler*, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-52819-8_11
- [30] Lipp, J., Quandt-Schubert, K., & Schwarz, M. (2016, November). BMW Group: Lebensphasenorientierte Arbeitszeiteinstrumente als personalwirtschaftlicher Erfolgsfaktor. *ifaa | Betriebspraxis & Arbeitsforschung*, Vol. 228, 25-28, Retrieved https://www.arbeitswissenschaft.net/fileadmin/Bilder/Angebote_und_Produkte/Publikationen/Zeitschriften_B_und_A/ifaa_B_A_228_web.pdf
- [31] Bund Verlag (2022, January 24). Gesetzesvorhaben: Hubertus Heil will dauerhaften Anspruch auf Homeoffice. *Bund Verlag*, Meldungen, Retrieved from <https://www.bund-verlag.de/aktuelles~Hubertus-Heil-will-dauerhaften-Anspruch-auf-Homeoffice~.html>
- [32] Bolzern-Konrad, B. (2021, November 8). VERTRAUEN | über die Meßbarkeit und strategische Implementierbarkeit des vermeintlich weichen Faktors. *Talks Seibert-Media.net*, Retrieved from https://talks.seibert-media.net/media/tools4agileteams2021/submissions/C33AJX/resources/Vortrag_Vertrauensmodell_Britta_Bolzer_prutfvd.pdf
- [33] Mayo Clinic Staff (2021, October 22). COVID-19 (coronavirus): Long-term effects. *Mayo Foundation for Medical Education and Research (MFMER)*, Retrieved from <https://www.mayoclinic.org/diseases-conditions/coronavirus/-in-depth/coronavirus-long-term-effects/art-20490351>
- [34] Corbyn, Z. (2022, April 27). Bossware is coming for almost every worker: the software you might not realize is watching you. *The Guardian*, Retrieved from <https://www.theguardian.com/technology/2022/apr/27/remote-work-software-home-surveillance-computer-monitoring-pandemic>
- [35] Digital.com Staff (2022, January 31). 6 in 10 employers require monitoring software for remote workers. *Digital.com*, Seattle, Retrieved from <https://digital.com/6-in-10-employers-require-monitoring-software-for-remote-workers/>

Authors' contacts:

Stefan Schmidt, Dipl.-Ing., Dipl.-Wirtschafting.
Technology Management and Management Training/VWI,
Fritzstrasse 41, D-82140 Olching, Germany
Stefan.Schmidt-1@t-online.de

Benjamin S. G. Schmidt, BA
Ludwigs Maximilian University, Munich,
Geschwister-Scholl-Platz 1,
80539 Munich, Germany

Article Title Only in English (Style: Arial Narrow, Bold, 14pt)

Ivan Horvat, Thomas Johnson, Marko Marić (Style: Arial Narrow, Normal, 10pt)

Abstract: Article abstract contains maximum of 150 words and is written in the language of the article. The abstract should reflect the content of the article as precisely as possible. TECHNICAL JOURNAL is a trade journal that publishes scientific and professional papers from the domain(s) of mechanical engineering, electrical engineering, civil engineering, multimedia, logistics, etc., and their boundary areas. This document must be used as the template for writing articles so that all the articles have the same layout. (Style: Arial Narrow, 8pt)

Keywords: keywords in alphabetical order (5-6 key words). Keywords are generally taken from the article title and/or from the abstract. (Style: Arial Narrow, 8pt)

1 INTRODUCTION (Article Design) (Style: Arial Narrow, Bold, 10pt)

(Tab 6 mm) The article is written in Latin script and Greek symbols can be used for labelling. The length of the article is limited to eight pages of international paper size of Letter (in accordance with the template with all the tables and figures included). When formatting the text the syllabification option is not to be used.

1.1 Subtitle 1 (Writing Instructions) (Style: Arial Narrow, 10pt, Bold, Align Left)

The document format is Letter with margins in accordance with the template. A two column layout is used with the column spacing of 10 mm. The running text is written in Times New Roman with single line spacing, font size 10 pt, alignment justified.

Article title must clearly reflect the issues covered by the article (it should not contain more than 15 words).

Body of the text is divided into chapters and the chapters are divided into subchapters, if needed. Chapters are numbered with Arabic numerals (followed by a period). Subchapters, as a part of a chapter, are marked with two Arabic numerals i.e. 1.1, 1.2, 1.3, etc. Subchapters can be divided into even smaller units that are marked with three Arabic numerals i.e. 1.1.1, 1.1.2, etc. Further divisions are not to be made.

Titles of chapters are written in capital letters (uppercase) and are aligned in the centre. The titles of subchapters (and smaller units) are written in small letters (lowercase) and are aligned left. If the text in the title of the subchapter is longer than one line, no hanging indents.

Typographical symbols (bullets), which are being used for marking an item in a list or for enumeration, are placed at a beginning of a line. There is a spacing of 10pt following the last item:

- Item 1
- Item 2
- Item 3

The same rule is valid when items are numbered in a list:

- 1) Item 1
- 2) Item 2
- 3) Item 3

1.2 Formatting of Pictures, Tables and Equations (Style: Arial Narrow, 10pt, Bold, Align Left)

Figures (drawings, diagrams, photographs) that are part of the content are embedded into the article and aligned in the centre. In order for the figure to always be in the same position in relation to the text, the following settings should be defined when importing it: text wrapping / in line with text.

Pictures must be formatted for graphic reproduction with minimal resolution of 300 dpi. Pictures downloaded from the internet in ratio 1:1 are not suitable for print reproduction because of unsatisfying quality.

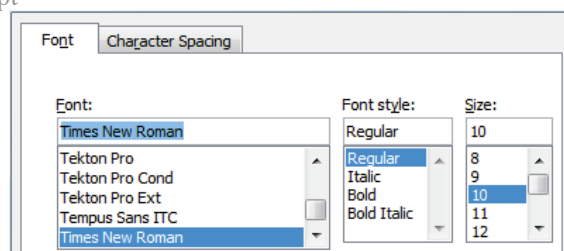


Figure 1 Text under the figure [1]
(Style: Arial Narrow, 8pt, Align Centre)

The journal is printed in black ink and the figures have to be prepared accordingly so that bright tones are printed in a satisfactory manner and are readable. Figures are to be in colour for the purpose of digital format publishing. Figures in the article are numbered with Arabic numerals (followed by a period).

Text and other data in tables are formatted - Times New Roman, 8pt, Normal, Align Center.

When describing figures and tables, physical units and their factors are written in italics with Latin or Greek letters,

while the measuring values and numbers are written upright.

10pt

Table 1 Table title aligned centre
(Style: Arial Narrow, 8pt, Align Centre)

	1	2	3	4	5	6
ABC	ab	ab	ab	ab	ab	ab
DEF	cd	cd	cd	cd	cd	cd
GHI	ef	ef	ef	ef	ef	ef

10 pt

Equations in the text are numbered with Arabic numerals inside the round brackets on the right side of the text. Inside the text they are referred to with equation number inside the round brackets i.e. "... from Eq. (5) follows" (Create equations with MathType Equation Editor - some examples are given below).

10pt

$$F_{\text{avg}}(t, t_0) = \frac{1}{t} \int_{t_0}^{t_0+t} F[q(\tau), p(\tau)] d\tau, \quad (1)$$

$$\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cdot \cos \frac{\alpha - \beta}{2}, \quad (2)$$

$$(AB)^T = B^T A^T. \quad (3)$$

10pt

Variables that are used in equations and also in the text or tables of the article are formatted as *italics* in the same font size as the text.

10pt

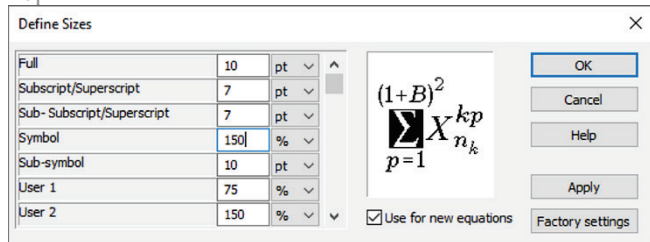


Figure 2 The texts under figures
(Style: Arial Narrow, 8pt, Align Centre)

10pt

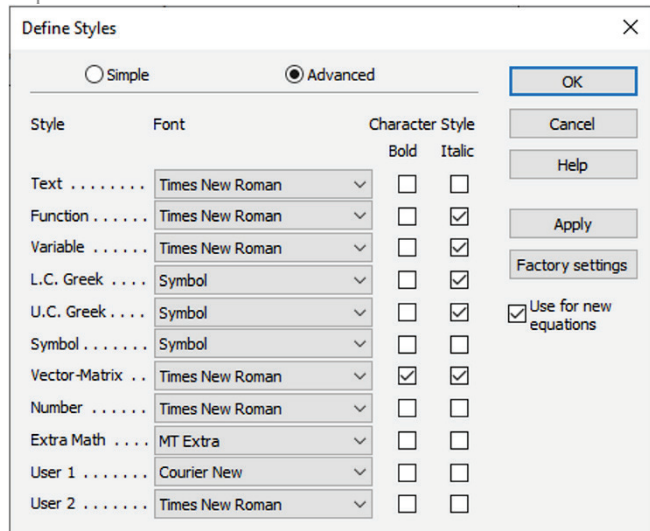


Figure 3 The texts under figures
(Style: Arial Narrow, 8pt, Align Centre)

Figures and tables that are a part of the article have to be mentioned inside the text and thus connected to the content i.e. "... as shown in Fig. 1..." or "data from Tab. 1..." and similar.

10pt

2 PRELIMINARY ANNOTATION

10pt

Article that is offered for publication cannot be published beforehand, be it in the same or similar form, and it cannot be offered at the same time to a different journal. Author or authors are solely responsible for the content of the article and the authenticity of information and statements written in the article.

Articles that are accepted for publishing are classified into four categories: original scientific papers, preliminary communications, subject reviews and professional papers.

Original scientific papers are articles that according to the reviewer and the editorial board contain original theoretical or practical results of research. These articles need to be written in such a way that based on the information given, the experiment can be repeated and the results described can be achieved together with the author's observations, theoretical statements or measurements.

Preliminary communication contains one or more pieces of new scientific information, but without details that allow recollection as in original scientific papers. Preliminary communication can give results of an experimental research, results of a shorter research or research in progress that is deemed useful for publishing.

Subject review contains a complete depiction of conditions and tendencies of a specific domain of theory, technology or application. Articles in this category have an overview character with a critical review and evaluation. Cited literature must be complete enough to allow a good insight and comprehension of the depicted domain.

Professional paper can contain a description of an original solution to a device, assembly or instrument, depiction of important practical solutions, and similar. The article need not be related to the original research, but it should contain a contribution to an application of known scientific results and their adaptation to practical needs, so it presents a contribution to spreading knowledge, etc.

Outside the mentioned categorization, the Editorial board of the journal will publish articles of interesting content in a special column. These articles provide descriptions of practical implementation and solutions from the area of production, experiences from device application, and similar.

10pt

3 WRITING AN ARTICLE

10pt

Article is written in the English language and the terminology and the measurement system should be adjusted to legal regulations, standards and the International System of Units (SI) (Quantities and Units: ISO 80 000 - from Part 1 to Part 14). The article should be written in third person.

Introduction contains the depiction of the problem and an account of important results that come from the articles that are listed in the cited literature.

Main section of the article can be divided into several parts or chapters. Mathematical statements that obstruct the reading of the article should be avoided. Mathematical statements that cannot be avoided can be written as one or more addendums, when needed. It is recommended to use an example when an experiment procedure, the use of the work in a concrete situation or an algorithm of the suggested method must be illustrated. In general, an analysis should be experimentally confirmed.

Conclusion is a part of the article where the results are being given and efficiency of the procedure used is emphasized. Possible procedure and domain constraints where the obtained results can be applied should be emphasized.

10pt

4 RECAPITULATION ANNOTATION

10pt

In order for the articles to be formatted in the same manner as in this template, this document is recommended for use when writing the article. Finished articles written in MS Word for Windows and formatted according to this template must be submitted using our The Paper Submission Tool (PST) (<https://tehnickiglasnik.unin.hr/authors.php>) or eventually sent to the Editorial board of the Technical Journal to the following e-mail address: tehnickiglasnik@unin.hr

The editorial board reserves the right to minor redaction corrections of the article within the framework of prepress procedures. Articles that in any way do not follow these authors' instructions will be returned to the author by the editorial board. Should any questions arise, the editorial board contacts only the first author and accepts only the reflections given by the first author.

10pt

5 REFERENCES (According to APA)

10pt

The literature is cited in the order it is used in the article. No more than 35 references are recommended. Individual references from the listed literature inside the text are addressed with the corresponding number inside square brackets i.e. "... in [7] is shown ...". If the literature references are web links, the hyperlink is to be removed as shown with the reference number 8. Also, the hyperlinks from the e-mail addresses of the authors are to be removed. In the literature list, each unit is marked with a number and listed according to the following examples (omit the subtitles over the references – they are here only to show possible types of references):

9pt

- [1] See <http://www.bibme.org/citation-guide/apa/>
- [2] See http://sites.umuc.edu/library/libhow/apa_examples.cfm
- [3] (Style: Times New Roman, 9pt, according to APA)
- [4] Amidzic, O., Riehle, H. J., & Elbert, T. (2006). Toward a psychophysiology of expertise: Focal magnetic gamma bursts as a signature of memory chunks and the aptitude of chess players. *Journal of Psychophysiology*, 20(4), 253-258.

<https://doi.org/10.1027/0269-8803.20.4.253>

- [5] Reitzes, D. C. & Mutran, E. J. (2004). The transition to retirement: Stages and factors that influence retirement adjustment. *International Journal of Aging and Human Development*, 59(1), 63-84. Retrieved from <http://www.baywood.com/journals/PreviewJournals.asp?Id=0091-4150>
- [6] Jans, N. (1993). *The last light breaking: Life among Alaska's Inupiat Eskimos*. Anchorage, AK: Alaska Northwest Books.
- [7] Miller, J. & Smith, T. (Eds.). (1996). *Cape Cod stories: Tales from Cape Cod, Nantucket, and Martha's Vineyard*. San Francisco, CA: Chronicle Books.
- [8] Chaffe-Stengel, P. & Stengel, D. (2012). *Working with sample data: Exploration and inference*. <https://doi.org/10.4128/9781606492147>
- [9] Freitas, N. (2015, January 6). People around the world are voluntarily submitting to China's Great Firewall. Why? Retrieved from http://www.slate.com/blogs/future_tense/2015/01/06/tencent_s_wechat_worldwide_internet_users_are_voluntarily_submitting_to.html
(Style: Times New Roman, 9pt, according to APA)

10pt

10pt

Authors' contacts:

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

Note: Gray text should be removed in the final version of the article because it is for guidance only.



14th International Conference on the Technology of Plasticity

September 24-29,
2023

BAY
OF CANNES
Mandelieu-La Napoule
Congress Center
France

**Abstract
submission
deadline:**
December 2nd, 2022

ictp2023.org





Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin



EURAM 20²³rd YEAR

Transforming Business for Good

14-16 June
Trinity Business School
Dublin, Ireland
www.euram.academy

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 16 – BROJ / ISSUE 4

PROSINAC 2022 / DECEMBER 2022 – STRANICA / PAGES 445-580



Sveučilište
Sjever

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TEHNICKIGLASNIK@UNIN.HR – [HTTP://TEHNICKIGLASNIK.UNIN.HR](http://tehnickiglasnik.unin.hr)