AGH

DECISION MAKING IN MANUFACTURING AND SERVICES

FACULTY OF MANAGEMENT

MMS

ANNUAL
VOL. 18
2024

AGH UNIVERSITY PRESS

KRAKOW 2024

The articles published in "Decision Making in Manufacturing and Services" have been given a favourable opinion by the reviewers designated by the Editorial Board.

Cover and title page design *Joanna Rokimi Marszewska*

Typesetting and Desktop Publishing by *Maciej Torz*

ISSN 1896-8325
ISSN 2300-7087 (on-line)
DOI: https://doi.org/10.7494/dmms

AGH University Press
al. A. Mickiewicza 30, 30-059 Krakow, Poland
tel. +48 12 617 32 28, +48 12 636 40 38
e-mail: redakcja@wydawnictwoagh.pl
www.wydawnictwo.agh.edu.pl

Number of copies: 45

Contents

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 5–18

Resilience of Robotic Solutions under Extreme Conditions

Dariusz Sala*, Pavlo Pikulin**, Valentyn Sobczuk***, Igor Kotsan****

Abstract. This study is devoted to the problems of the use of modern advanced technologies by logistics companies in their efforts to increase the speed of their technological operations and transform their business processes; this is aimed at reducing their financial costs, increasing the efficiency of their use of labor resources, and minimizing their risks. Today, this is a decisive factor in increasing a company's competitiveness in the market, increasing its profitability, and realizing its long-term leadership. Innovative logistics is an effective tool for streamlining flow processes through the introduction of high-tech innovations in the operational and strategic management of the market structures that are aimed at improving the quality of their customer service, increasing the efficiency of their flow processes, and reducing the total cost of their implementation in order to achieve key business objectives.

The paper examines approaches to the automation of business processes in the logistics sector in the context of the robotization of technological operations while taking those features that are due to the functioning of enterprises under conditions of constant exposure to extreme risks into account. The concept of the robotization of processes has been developed, which will increase the productivity and efficiency of businesses, help reduce their operating costs, reduce their likelihood of personnel errors, and contribute to improving their business security. The results are implemented in the practice of a number of logistics companies in the real sector of the economy.

Keywords: industrial robots, robotics, extreme risks, logistics enterprises, business process optimization

Mathematics Subject Classification: 91B38

JEL Classification: L23

Submitted: January 15, 2024

Revised: February 23, 2024

* AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: sala@agh.edu.pl
** AGH University of Krakow, AGH University Doctoral School, Krakow, Poland, e-mail: pavel@deusrobotics.com
*** Taras Shevchenko National University of Kyiv, Faculty of Mechanics and Mathematics, Kyiv, Ukraine, e-mail: sobchuk@knu.ua
**** AGH University of Krakow, Faculty of Materials Science and Ceramics, Krakow, Poland, e-mail: ikotsan@agh.edu.pl

DOI: https://doi.org/10.7494/dmms.2024.18.6054

1. INTRODUCTION

It is impossible to imagine the rapid development of the modern world economy without constant innovative solutions. Nowadays, there is a trend toward constant steady progress in both research and innovation in various industries, and the field of logistics is no exception. The use of modern advanced technologies by logistics companies guarantees the high speed of the executions of necessary operations and operations, thus reducing their financial costs and labor costs; this serves as a decisive factor in increasing the companies' competitiveness in the market, increasing their profitability, and realizing their long-term leadership. Innovative logistics is the most relevant component of logistics activities. This is an effective tool for streamlining flow processes through the introduction of high-tech innovations in the operational and strategic management of market topologies that are aimed at improving the quality of customer service, increasing the efficiency of flow processes, and reducing the total cost of their implementation in order to achieve key business objectives.

Moreover, constant work on implementing innovative solutions in the flow processes of logistics companies is the key to business stability in the face of constant pressure (risks of a very diverse nature). First of all, companies are aimed at administering the minimization of classic business risks: organizational, commercial, financial, legal, etc. However, events such as the SARS-CoV-2 pandemic, the full-scale war between Ukraine and Russia, and Israel's war with Hamas require a separate study regarding the problems of the sustainable functioning of business processes under conditions of extreme risks.

Actually, this study is devoted to the study of the problems of extreme risks and building a strategy for the sustainable functioning of business processes under the influence of destabilizing internal and external factors.

2. LITERATURE REVIEW AND PROBLEM STATEMENT

The problems of optimizing, minimizing risks, and improving the efficiency of logistics-channel management are urgent tasks that are in the constant focus of the attention of the world's leading researchers.

Ho et al. (2015) provided a fairly broad comprehensive overview of SCRM publications. The article presented a classification of studies, a detailed overview of risk types, risk factors, and risk-management strategies in supply chains. The application of artificial intelligence algorithms to the supply-chain- management system through the implementation, visualization, automation, and intelligent management of all links in the supply chain was explored in (Lin et al. 2022). In Altiparmak et al. (2006), the authors proposed a new procedure for finding the optimal solution for the functioning of the network of a supply chain based on genetic algorithms for finding the Pareto set-optimal solutions to the multipurpose design problem.

In Nezamoddini et al. (2020), the authors considered a supply chain to be a network of suppliers, manufacturing enterprises, distribution centers, and markets. The researchers proposed a model that, in the face of the uncertainty that is associated with demand, facility disruptions, turnaround times, and disruptions in supply, production, and distribution channels, aimed to achieve risk-based optimization by the processing strategic, tactical, and operational decisions of a functioning supply chain. For a representative European supply chain, a model for the design and planning of backflow supply chains was proposed in Cardoso et al. (2013) using mixed integer linear programming (MILP). An expert approach to risk assessment in logistics systems was investigated in Aqlan & Lam (2015).

For the stochastic model of achieving global optimality, Baghalian et al. (2013) presented a transformation based on the method of piecewise linearization. The authors illustrated the initial data of the model with the help of several numerical examples and practical applied research in the agrifood industry. The proposed model took the uncertainties on the supply and demand sides into account at the same time, which made it quite applicable when compared to the other similar models that have been described in the existing literature.

One of the key roles in minimizing risks is played by the problem of ensuring the functional stability of business processes and technological processes under the influence of internal and external destabilizing factors. Studies of this problem were carried out in detail in Barabash et al. (2023), Obidin et al. (2017), Pichkur & Sobchuk (2021), Sobchuk et al. (2021).

The resilience of critical infrastructure from unauthorized external intrusions was studied in Laptiev et al. (2023), Pichkur et al. (2022), Svynchuk et al. (2021), Yevseiev, et al. (2021), Yevseiev et al. (2023). In the data from the studies, the authors studied the approaches to ensuring the minimization of the risks of information loss in detail; these can lead to critical consequences for the functioning of the information systems of enterprises, the prohibited methods of minimizing such impacts, and the developed algorithms that were aimed at increasing the cybernetic stability of information systems.

In recent decades, industrial robots have been developing rapidly in the leading sectors of the global economy; these cover many new industries such as aerospace, military, medical, etc. The development trend of industrial robots in the future should focus mainly on the following areas of development: human-robot collaboration, artificial intelligence, new industrial users, digitization, and facilitation. Dzedzickis et al. (2022) described the current state of the development of new industrial robots and described the trends in their future development. The work aimed to create a theoretical basis for the development of companies that special in the development of industrial robots.

In Bernardo et al. (2022), the authors provided an overview of the advanced applications of robotic technologies in the real industrial sector. A review of survey publications and technical reports (classified according to the criteria for their application) was carried out. The results of the analysis revealed the prerequisites for the existing obstacles and problems in this innovation sector. In particular, the problems that were related to the spheres of psychology, human nature, the introduction of special artificial intelligence, and the paradigm of the robot-oriented design of objects were disclosed.

An overview of the prospects of existing robotic systems for the intralogistics of companies (which aimed to determine which research paths had been used to date and highlight current and future research directions) was made in Bernardo et al. (2022). The authors of the paper focused on the study of localization and route planning, task scheduling, optimization, and the representation of knowledge in robotic systems. Given the rapid growth in the amount of information that robotic agents must process, the application of strategies that are based on semantic knowledge is becoming increasingly important. Transforming domain knowledge and minimizing ambiguity will allow agents to reason and facilitate the exchange of knowledge between robotic agents and humans. In the near future, it will be increasingly important to rethink production and logistics systems from a human perspective. Business processes will facilitate the balanced use of automation and digital technologies in order to enhance the unique and irreplaceable capabilities of their operators, who will continue to play fundamental roles for the companies of the future (Cimini et al. 2022). The rapid development of robotics is impossible without effective collaboration. The authors in Atzeni et al. (2021) paid special attention to this issue in the context of the challenges that were faced in the Logistics 4.0 environment.

Despite the significant interest in the problems of automating the processes in logistics, the problems of minimizing risks in the face of extreme risks that are caused by hostilities and the operations of the logistics infrastructure under the constant risk of damage or complete destruction are extremely important and have been poorly studied. This work is specifically devoted to this problem.

3. FORMULATION OF PROBLEM

To explore approaches to the automation of business processes in the logistics sector in the context of the robotization of technological operations. Taking the peculiarities that are caused by the functioning of some enterprises when under the constant influence of extreme risks into account, developing a concept of robotic-process automation will increase the productivity and efficiency of businesses, help reduce their operating costs, reduce their likelihood of personnel errors, and contribute to their improving business security.

4. MAIN SECTION

It should be noted right away that robotization and automation are not panaceas in the processes of minimizing the risks of modern enterprises. Moreover, an extremely important place in this process is further given to the development and implementation of various measures that are aimed at minimizing risks such as clear safety procedures, the training of employees, the regular monitoring of equipment, etc. At the same time, robotics is a powerful tool that can help companies better protect themselves from risks under extreme conditions. This is especially true when it comes to enterprises that operate under constant exposure avoiding extreme risks – these companies include those enterprises in the chemical, nuclear, military, and other particularly vulnerable sectors of the economy. Even though ordinary logistics companies do not classify such risk groups at first glance, this is not the case. Numerous examples have confirmed that the processing of the most innocent postal messages can be accompanied by the risk of processing parcels that contain toxic substances, explosive devices, bacterial infections, etc. When developing strategies for the long-term development of companies in a wide variety of industries, it is therefore necessary to proceed by default from the reality of the most extreme risks.

It is useful for investors, visionaries, and business leaders to understand that robotization in our time is the key to the success of their enterprises in the future. Let us consider a number of specific examples of how robotics can be used to minimize risks under the extreme conditions of modern enterprises:

- *In industry*, robots are used to perform hazardous tasks such as working with toxic materials and working under the conditions of elevated temperatures, radiation, pressure, etc. They are used to monitor and manage critical infrastructure, e.g. power plants, water supply systems, etc.
- *In logistics*, robots can perform the tasks of transporting goods, sorting goods, maintaining warehouses, and delivering goods under extreme conditions (during natural disasters, pandemics, etc.).
- *In healthcare*, robots are used to provide first aid, care for patients, and disinfect facilities. They can also be used to develop new drugs and medical products that can help people survive under extreme conditions.

It is extremely important to note that robotization is designed to have a significant impact on minimizing risks under the conditions of a company's operation. First of all, a number of such advantages should be highlighted:

- *Increased productivity and efficiency.* Robots perform tasks faster and more accurately than humans can; they help increase the productivity and efficiency of the company. Among other things, large-scale robotization aims to help the company better respond to the effects of extreme conditions, as this will allow for faster recovery from them.
- *Reducing the likelihood of human error.* Robots are less-prone to human-like errors, resulting in reduced risks of accidents, and other problems.
- *Increased safety.* Robots can perform tasks in hazardous environments that are hazardous to humans. This can help companies protect their employees from the risks of injury and damage under extreme conditions.
- *Cost reduction.* Robots can help the company reduce its labor, training, and other personnel costs; therefore, this has the direct effect of increasing the company's profits. The company gains new opportunities in investment activities via the implementation of sustainable development strategies while minimizing risks, costs, etc.

5. CASE STUDY – NOVA POST

NOVA is a group of companies that provide a full range of logistics, financial, and IT-related services in Ukraine and around the world. The group includes Ukrainian and international companies: Nova Post in Ukraine, Nova Post Europe (with its own branches and offices in 11 European countries), SuperNova (its own cargo airline company), NovaPay (a financial company), Nova Digital (an IT company), and Nova Global (which provides cross-border services around the world).

Among other things, Nova Post uses robots from Deus Robotics in its warehouses in Ukraine for handling logistics tasks inside their sorting rooms and hubs. In the discussed case, transport robots were analyzed that made it possible to transport entire racks with postal parcels on them. Each robot can carry 500 kg and move along designated routes (Figure 1), optimizing each route at a maximum speed of 1.2 M/s thanks to AI. The robots are electrically powered, and their charging time is two hours; they can work without interruption for eight hours at full load.

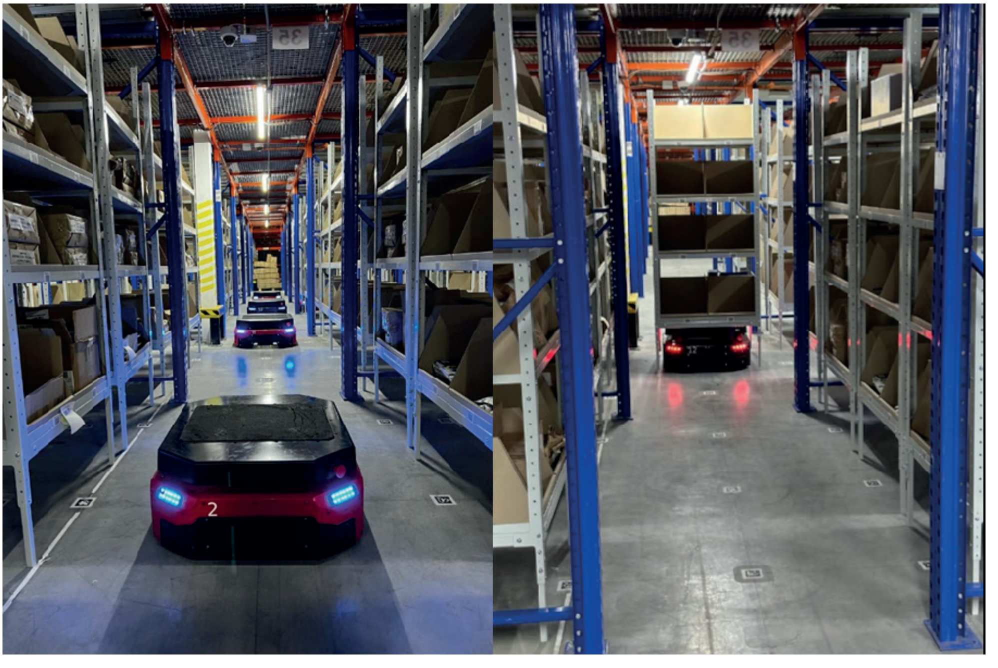


Fig. 1. *Robots at work in warehouses of Nova Post – photo by Deus Robotics*

Logistics robots such as those that are used by Nova Post enable the introduction of self-service sorting functions. Using advanced artificial intelligence (AI) and routing algorithms, the robots independently plan and implement optimal routes in warehouses; this eliminates the need for employees to manually sort packages and speeds up the entire process of preparing packages for shipment. Thanks to AI, the robots are able to dynamically optimize routes in real time. The system analyzes data on route load, parcel location, and current warehouse traffic, which allows the robots to adjust their routes depending on the current conditions. This not only speeds up the process but also minimizes the risks of collisions and disruptions.

Thanks to the use of advanced vision systems and sensors (as well as the use of QR codes), the robots are able to precisely identify, lift, and move packages. This eliminates the risk of manual handling errors, thus improving customer service and reducing losses. Automation also enables the constant monitoring of logistics processes and their continuous reporting. The systems collect data on efficiency, routes, numbers of shipments handled, etc.; this allows for the continuous improvement of the processes and quick responses to any possible problems.

Additionally, the data that is collected in this way is indispensable for AI, which learns to react in emergency situations for improving route-optimization algorithms based on previous events and increases the effectiveness and efficiency of the transport. In conclusion, the use of robots in the logistics of Nova Post not only accelerates the processes of sorting and moving parcels but also introduces flexibility and scalability to its logistics operations; this translates to improved efficiency and increased customer service quality.

The use of Deus Robotics robots in the logistics process of Nova Post seems to be a comprehensive approach for automating and optimizing operations; this translates to efficiency, precision, and sustainable development in the area of logistics. Apart from the logistic aspect e.g., the ecology-related ones, this shows the benefits of the solutions that have been proposed by Deus Robotics. The robots that are used by Nova Post in Ukraine are electrically powered, which has several key benefits. An electric power supply is in line with sustainable development trends, minimizing greenhouse gas emissions and other air pollutants in the workplace. Compared to traditional power sources such as internal-combustion engines, electric robots are more friendly to people and the environment – especially in closed facilities such as warehouses. Electric power can lead to savings in operating costs in the long run.

Electricity prices are often more stable than fossil fuel prices, and electric drives can be more energy-efficient when compared to traditional combustion drives. The short charging time (two hours) compared to the long working time (eight hours) means that the robots can work for longer periods of time without the need for long charging breaks. This increases the overall efficiency of the logistics system.

Electric robots generate less noise and do not emit harmful substances, which helps to keep internal working environments such as warehouses and sorting rooms clean. This may also affect the comfort of the staff and the general atmosphere in the workplace. Importantly, electrically powered systems easily integrate with renewable energy sources such as solar or wind energy. This opens the way for logistics companies toward more-sustainable energy models.

An example route diagram in a warehouse with postal parcels is presented in Figure 2. The green points represent the positions of key QR codes on the map (the robots moves between these points), while the red points represent additional QR codes for better navigation. Those points with four red wheels and one green one are the positions where shelves can be stored.

The statistics so far: as of August 2022, each deployed Deus robot has traveled more than 1000 KM and moved more than 5000 items with an accuracy above 99.99% and an uptime of 99.9%.

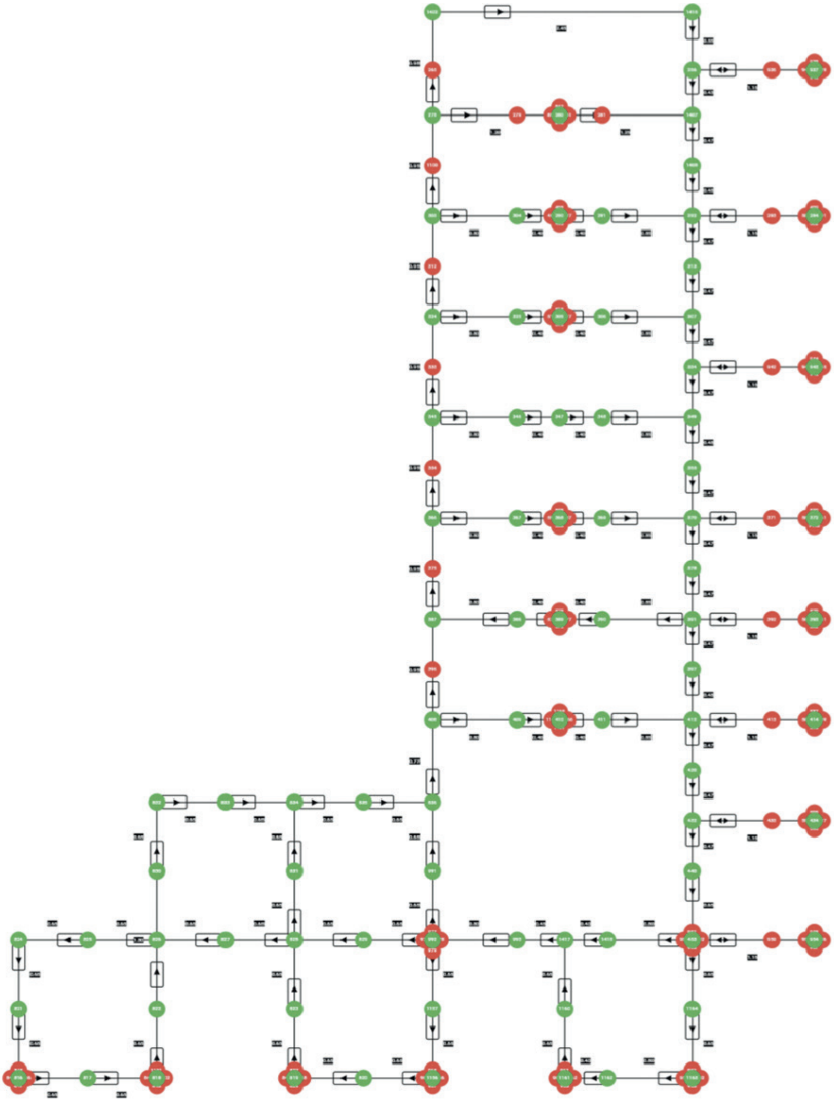


Fig. 2. *Robot-moving layout in Nova Post warehouses – layout by Deus Robotics*

6. DISCUSSION

Let us illustrate an analysis of real data (Figures 3 and 4) from a Tier 1 manufacturing company in the United States. As of now, the company's warehouse employs a small number of employees in one daily shift. At the same time, the company doubles its operational efficiency annually thanks to the robotization of its technological processes for processing commodity resources and raw materials.

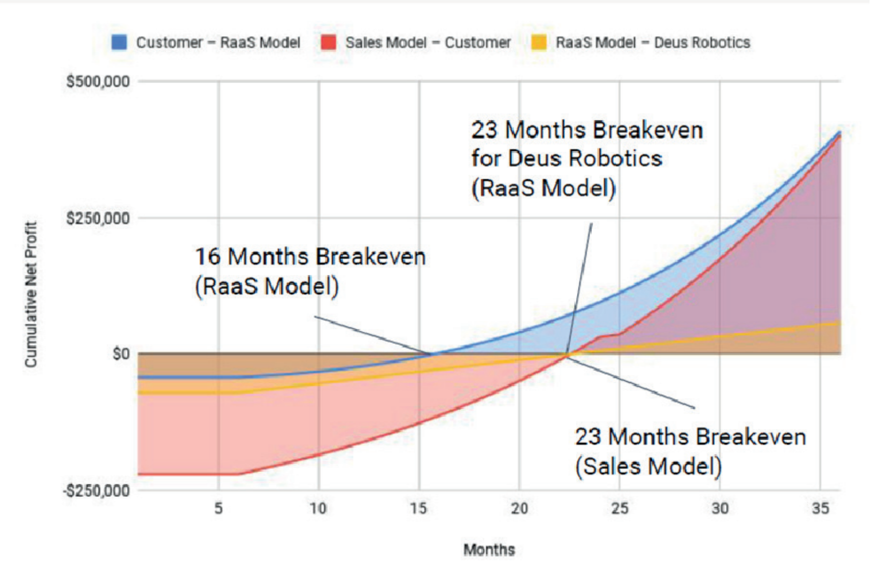


Fig. 3. *RaaS model & sales model dynamics graphs – Deus Robotics*

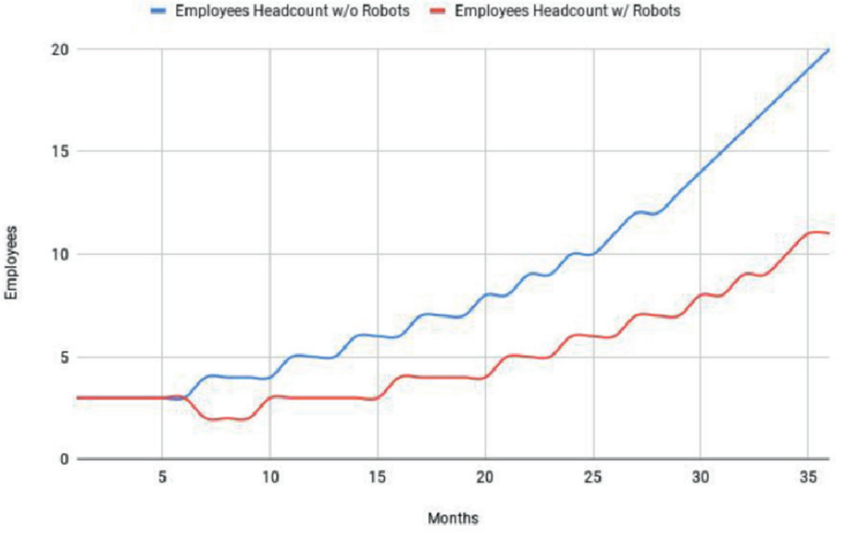


Fig. 4. *Graphs of dynamics of number of employees after taking robotization into account – graphs by Deus Robotics*

At the same time, the dynamics of the numbers of employees are shown in Figure 4 after taking the robotization of the technological processes into account (with and without them).

Let us illustrate the effect that occurs if we assume that the company increases its number of employees to 100 people working in two shifts instead of one (Figure 5). In this scenario, the return on investment (ROI) is shown in Figure 6.

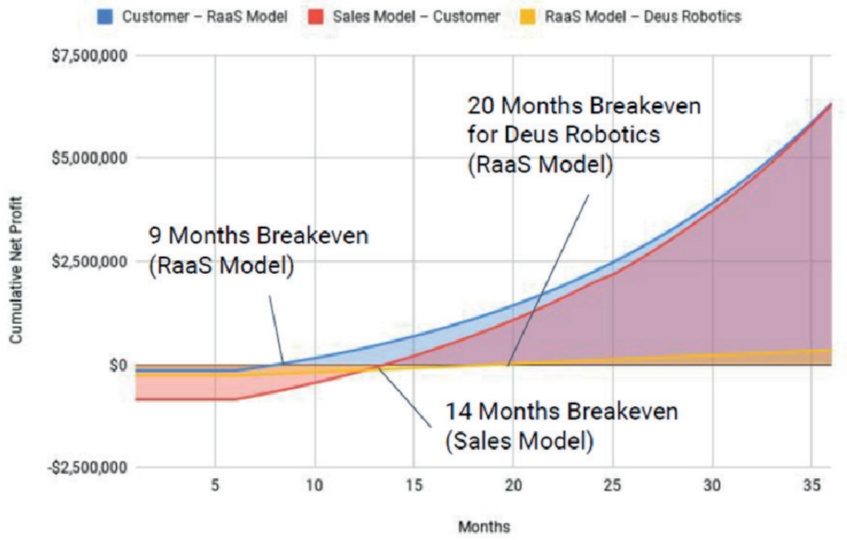


Fig. 5. *Graphs of dynamics of RaaS model & sales model with 100 employees – model by Deus Robotics*

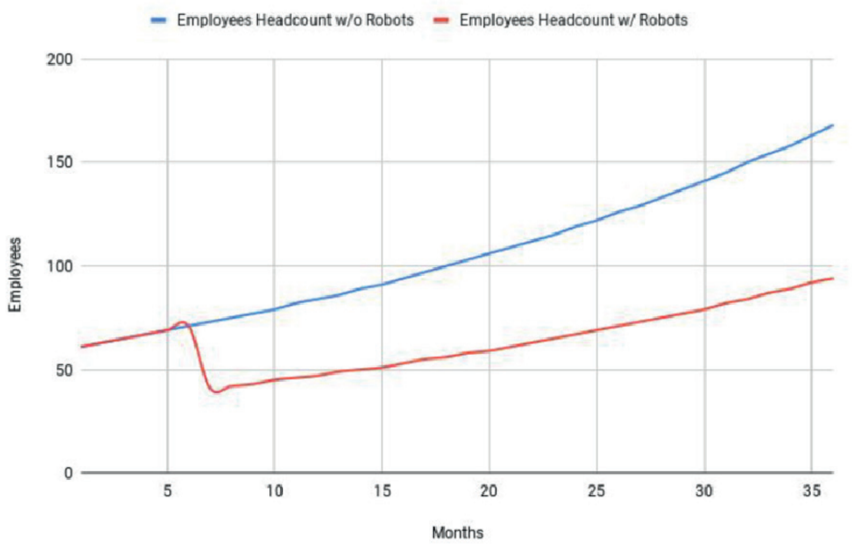


Fig. 6. *ROI dynamics charts – graph by Deus Robotics*

Similar modeling should be carried out for all localizations where large-scale robotization programs are being implemented. Such calculations are especially important for those industries where the robots operate in aggressive environments or where they operate under constant exposure to extreme risk factors. In particular, the management of Deus Robotics implemented a multi-component strategy after conducting relevant research in order to minimize risks under conditions of extreme risks, the focus of which was aimed at the following:

1. Improving employee safety:
   – relocation of employees and their families to safe places;
   – implementation of software for management of robotic sections of logistics complexes;
   – taking specifics of work during war period into account in current business processes.
2. Compensation for impact of labor shortages:
   – implementation of online training programs for clients;
   – minimization of queues of receiving parcels by means of warning system.

Therefore, a combination of measures that are aimed at organizing a safe workplace for employees through the large-scale implementation of cloud technologies for order administration and the robotization of cargo-handling processes became the key for successfully minimizing the extreme risks that arose as a result of the acute phase of the armed invasion of the aggressor country.

The carried-out full-scale modeling illustrates that the automation of the technological processes through the introduction of robotic solutions allowed them to both obtain an immediate effect of increasing productivity (use of working times) and guarantee a sustainable effect of the return on investment in the long term.

7. CONCLUSIONS

This paper examines the current state of the results that highlight the problems of using modern advanced technologies by logistics companies in attempts to increase the speed of their technological operations and transform their business processes; their ultimate goals were reducing their financial costs, increasing the efficiency of the use of their labor resources, and minimizing their risks. In today's economic realities, this is a decisive factor in increasing a company's competitiveness in the market, thus increasing its profitability and realizing its long-term leadership.

It is innovative logistics that is an effective tool for streamlining flow processes through the introduction of high-tech innovations in the operational and strategic management of market structures that are aimed at improving the quality of customer service, increasing the efficiency of flow processes, and reducing the total cost of their implementation in order to achieve key business objectives.

The paper examines approaches to the automation of business processes in the logistics sector in the context of the robotization of technological operations, taking into account the features of to the functioning of enterprises under the constant impacts of the extreme risks that were caused by the SARS-CoV-2 pandemic and are still being caused by the full-scale war between Ukraine and Russia. Actually, the experimental part describes the effects that illustrate the results of the study of the problems of extreme risks and the construction of a strategy for the sustainable functioning of business processes under the influence of destabilizing internal and external factors.

The proposed concept of robotic process automation will increase the productivity and efficiency of a business, helping it reduce its operating costs, reduce the likelihood of personnel errors, and contribute to improving its business security. The results have been implemented in the practices of a number of logistics companies in the real sector of the economy.

The real effect of the robotization of technological processes in clients warehouses has made it possible to accomplish the following:

– increase process productivity by 300%;
– reduce distance of movement of employees by 72%;
– increase usable warehouse space by 20%;
– reduce frequency of personnel errors five-fold;
– increase productive workload of each employee by +3.5 hours per employee per shift.

Similar effects can be obtained for other enterprises that use logistics complexes with auto-lubricated complexes for the handling and storage of goods and the robotization of technological processes. In the future, it is planned to improve the concept of the robotization of business processes, develop new route-management algorithms, and develop strategies for minimizing the impacts of extreme risks.

REFERENCES

Altiparmak F., Gen M., Lin L. & Paksoy T. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers & Industrial Engineering*, **51**(1)1, pp. 196–215. **DOI**: https://doi.org/10.1016/j.cie.2006.07.011.

Aqlan F. & Lam S.S. (2015). A fuzzy-based integrated framework for supply chain risk assessment. *International Journal of Production Economics*, **161**, pp. 54–63. **DOI**: https://doi.org/10.1016/j.ijpe.2014.11.013.

Atzeni G., Vignali G., Tebaldi L. & Bottani E. (2021). A bibliometric analysis on collaborative robots in Logistics 4.0 environments. *Procedia Computer Science*, **180**, pp. 686–695. **DOI**: https://doi.org/10.1016/j.procs.2021.01.291.

Baghalian A., Rezapour S. & Zanjirani Farahani R. (2013). Robust supply chain network design with service level against disruptions and demand uncertainties: A real-life case. *European Journal of Operational Research*, **227**(1), pp. 199–215. **DOI**: https://doi.org/10.1016/j.ejor.2012.12.017.

Barabash O., Sobchuk V., Musienko A., Laptiev O., Bohomia V. & Kopytko S. (2023). System analysis and method of ensuring functional sustainability of the information system of a critical infrastructure object. In: M. Zgurovsky & N. Pankratova (Eds.), *System Analysis and Artificial Intelligence*, Springer Nature, Cham, Switzerland, pp. 177–192. **DOI**: https://doi.org/10.1007/978-3-031-37450-0_11.

Bernardo R., Sousa J.M.C. & Gonçalves P.J.S. (2022). Survey on robotic systems for internal logistics. *Journal of Manufacturing Systems*, **66**, pp. 339–350. **DOI**: https://doi.org/10.1016/j.jmsy.2022.09.014.

Cardoso S.R., Barbosa-Póvoa A.P.F.D. & Relvas S. (2013). Design and planning of supply chains with integration of reverse logistics activities under demand uncertainty. *European Journal of Operational Research*, **226**(3,) pp. 436–451. **DOI**: https://doi.org/10.1016/j.ejor.2012.11.035.

Cimin C., Lagorio A., Cavalieri S., Riedel O., Pereira C.E. & Wang J. (2022). Human-technology integration in smart manufacturing and logistics: current trends and future research directions. *Computers & Industrial Engineering*, **169**, art. no. 108261. **DOI**: https://doi.org/10.1016/j.cie.2022.108261.

Dzedzickis A., Subačiūtė-Žemaitienė J., Šutinys E., Samukaitė-Bubnienė U. & Bučinskas V. (2022). Advanced applications of industrial robotics: New trends and possibilities. *Applied Sciences*, **12**(1), art. no. 135. **DOI**: https://doi.org/ 10.3390/app12010135.

Ho W., Zheng T., Yildiz H. & Talluri S. (2015). Supply chain risk management: a literature review. *International Journal of Production Research*, **53**(16), pp. 5031–5069. **DOI**: https://doi.org/10.1080/00207543.2015.1030467.

Laptiev O., Musienko A., Nakonechnyi V., Sobchuk A., Gakhov S. & Kopytko S. (2023). Algorithm for Recognition of Network Traffic Anomalies Based on Artificial Intelligence. In: *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. **DOI**: https://doi.org/10.1109/HORA58378.2023.10156702.

Lin H., Lin J. & Wang F. (2022). An innovative machine learning model for supply chain management. *Journal of Innovation & Knowledge*, **7**(4), art. no. 100276. **DOI**: https://doi.org/10.1016/j.jik.2022.100276.

Nezamoddini N., Gholami A. & Aqlan F. (2020). A risk-based optimization framework for integrated supply chains using genetic algorithm and artificial neural networks. *International Journal of Production Economics*, **225**, art. no. 107569. **DOI**: https://doi.org/10.1016/j.ijpe.2019.107569.

Obidin D., Ardelyan V., Lukova-Chuiko N. & Musienko A. (2017). Estimation of functional stability of special purpose networks located on vehicles. In: *2017 IEEE 4th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), October 17–19, 2017, Kyiv, Ukraine*, National Aviation University, Kyiv, pp. 167–170. **DOI**: https://doi.org/10.1109/APUAVD.2017.8308801.

Pichkur V., Laptiev O., Polovinkin I., Barabash A., Sobchuk A. & Salanda I. (2022). The method of managing man-generated risks of critical infrastructure systems based on ellipsoidal evaluation. In: *2022 IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT)*, pp. 133–137. **DOI**: https://doi.org/10.1109/ATIT58178.2022.10024244.

Pichkur V.V. & Sobchuk V.V. (2021). Mathematical model and control design of a functionally stable technological process. *Journal of Optmiization, Differential Eqations and Their Applications*, **29**(1), pp. 32–41. **DOI**: https://doi.org/10.15421/142102.

Sobchuk V., Olimpiyeva Y., Musienko A. & Sobchuk A. (2021). Ensuring the properties of functional stability of manufacturing processes based on the application of neural networks. In: V. Snytyuk, A. Anisimov, I. Krak, M. Nikitcheko, O. Marchenko, F. Mallet, V. Tsyganok, A. Chris, A. Pester, H. Tanaka, K. Henke, O. Chertov, S. Bozóki, V. Vovk (Eds.), *Proceedings of the 7th International Conferene* "*Information Technology and Interactions*" *(IT&I-2020). Workshops Proceedings, Kyiv, Ukraine, December 2–3, 2020*, pp. 106–116. **URL**: https://ceur-ws.org/Vol-2845/Paper_11.pdf.

Svynchuk O., Barabash A., Laptiev S. & Laptieva T. (2021). Modification of query processing methods in distributed databases using fractal trees. In: *International Scientific and Practical Conference* "*Information Security and Information Technologies*"*, 13–19 September 2021, Kharkiv – Odesa, Ukraine. Conference Proceedings*, Simon Kuznets Kharkiv National University of Economics, Kharkiv – Odesa, pp. 39–44. URL: https://drforum.science/wp-content/uploads/2021/12/proceedings_ibit-2021.pdf.

Yevseiev S., Khokhlachova Y., Ostapov S. & Laptiev O. (Eds.) (2023). *Models of Socio-cyber-physical Systems Security. Monograph.* PC TECHHNOLOGY CENTRE, Kharkiv.

Yevseiev S., Ponomarenko V., Laptiev O. & Milov O. (Eds.) (2021). *Synergy of Building Cybersecurity Systems. Monograph.* PC TECHNOLOGY CENTER, Kharkiv.

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 19–38

Six Sigma vs. Other Quality Improvement Tools: Comparative Analysis of Trends over Period of 1985–present

Marcin Nakielski*, Anna Ludwig**

Abstract. Six Sigma is a widely adopted method in various industries that is aimed at process improvement and quality management. Understanding the evolving interest and utilization of Six Sigma can provide valuable insights into its current significance and prospects. Using data from Google Trends, Google Books, Web of Science, and Scopus, this study examined the search volumes and interests in keywords and phrases that were related to Six Sigma over a specified period of time. The global analysis revealed the overall direction of interest in Six Sigma worldwide, highlighting periods of peak interest and potential significant shifts in the method's popularity. By identifying those times with the highest concentrations of interest, the article provides a deeper understanding of the adoption and perception of Six Sigma. On top of this, Six Sigma was compared in popularity (by trends) with other known methods such as Lean, Kaizen, PDCA, and TQM. This research contributes to the existing body of knowledge by shedding light on the current trends and future directions of Six Sigma globally. The findings offer valuable insights for practitioners, researchers, and organizations that seek to leverage Six Sigma for process improvements and quality management.

Keywords: Six Sigma, Lean Six Sigma, Lean, DMAIC, Kaizen, TQM, PDCA, trends, Google Trends, WoS, Google Books Ngram Viewer

Mathematics Subject Classification: 62P25

JEL Classification: C44

Submitted: April 08, 2024

Revised: June 11, 2024

* AGH University of Krakow, AGH Doctoral School, Krakow, Poland; Nexteer Automotive Sp. z o.o., Tychy, Poland, e-mail: nakielsk@agh.edu.pl.
** EMBS, Gliwice, Poland, e-mail: annaludwig01@wp.pl.

DOI: https://doi.org/10.7494/dmms.2024.18.6227

1. INTRODUCTION

Six Sigma is a quality-management method that focuses on minimizing errors and defects in production and service processes. It is based on data collection and analysis to identify the root causes of problems, eliminate unnecessary variabilities, and minimize variabilities in manufacturing and business processes. The main goal of Six Sigma is to reduce defects to a level that corresponds to fewer than 3.4 cases per million (Muralidharan, 2015; Sokovic et al., 2005).

The history of Six Sigma began in the 1980s, when Motorola (an American telecommunications giant) made exceptional efforts to improve its manufacturing processes. During this period, Bill Smith (an engineer and quality specialist at Motorola) initiated the concept of Six Sigma as a quality-management strategy that focused on minimizing variations in production. His work in this field led to the introduction of a revolutionary quality-control system based on advanced statistical tools.

Motorola's success in implementing Six Sigma as a method for quality control and process improvement quickly drew the attention of other corporations, resulting in its widespread adoption on a large scale. Over time, Motorola became a symbol of Six Sigma's success, thus contributing to the global popularization of this concept (Eckes, 2001; Watson, 1994).

Six Sigma is mainly applied in large organizations. According to industry consultants, companies with fewer than 500 employees are not suitable for implementing Six Sigma or need to adapt the standard approach in order to make it effective; however, Six Sigma encompasses many tools and techniques that work well in small and medium-sized enterprises. The infrastructure that is described as being necessary for supporting Six Sigma is a result of an organization's size rather than a requirement of Six Sigma itself (Dusharme, 2024).

Six Sigma projects are guided by a couple of project methods (which are presented below). The DMAIC methodology in Six Sigma is a five-step improvement cycle; i.e., define, measure, analyze, improve, and control (Palací-López et al., 2020). Essential components of a DMAIC project encompass team discipline, the structured utilization of metrics and tools, and the execution of a well-crafted project plan (with clearly defined goals and objectives). Lean Six Sigma (LSS) modifies the DMAIC approach by placing an emphasis on speed; this primarily focuses on streamlining a process by identifying and eliminating non-value-added steps. By implementing a lean production process, waste is eliminated. Target metrics include achieving zero wait times, zero inventories, scheduling based on customer demands, reducing batch sizes (in order to enhance their flows), line balancing, and decreasing overall process times. The ultimate objective of Lean Six Sigma is to manufacture high-quality products that fulfill customer requirements as efficiently and effectively as possible. If a process cannot be improved in its current design, another widely recognized problem-solving approach within Six Sigma can be employed. The DMADV process (define, measure, analyze, design, validate) is used for fundamentally redesigning such a process (De Feo & Barnard, 2005; Muralidharan, 2015). Six Sigma offers a quality-improvement and business-excellence roadmap that is inspired by statistical thinking and guided by data-driven techniques (Goh, 2020). In order to be able to benefit from its implementation, however, a considerable amount of time and resources should be devoted (Uluskan, 2022). There are still companies today that are trying to decide, among other things, whether it is wise or "safe" to embrace Six Sigma after being exposed to various forms of publicity on the subject (Goh, 2020). Among other reasons, this article was prepared in order to shine a fresh light on the current interest in Six Sigma.

This article is structured as follows. First, methods are presented that describe the approaches that the authors of scientific publications have taken to understand the problem. This starts with a keyword selection based on the contents of the bibliographical databases of available publications. Then, the individual data sources are described, with details about the search approach that has been taken by these authors.

Second, our results are presented in the forms of graphs and are typically constructed in the following manner: 1) focusing on Six Sigma-related terms and the method's behavior over time (i.e., terms/phrases such as "Six Sigma," "DMAIC," and "Six Sigma methodology"); and 2) presenting a comparative view on Six Sigma versus other similar methods (i.e., terms such as "Lean," "Kaizen," and "PDCA").

The results section is followed by a discussion where the authors interpret the data, analyze the outcomes, and explain the possible reasons for the given findings. Finally, the conclusion section provides a high-level summary and the outcomes of the study.

2. METHOD

The method that was used to conduct the study can be described as follows. The authors started with a keyword selection using the Web of Science database and 1000 recent publications that were sorted by relevance in order to determine the phrases to be compared in two groups – the first group was focused on quality management and process-improvement tools, while the second compared phrases that were related to the term "Six Sigma."

Once the keywords were determined, the authors conducted searches in four databases – Google Trends, Google Books Ngram Viewer, Web of Science, and Scopus. The first two databases were selected due to their wide reaches (not being limited to academic and scientific publications) – the intentions of the authors was to capture the interest among current or potential Six Sigma users. The final two databases (Web of Science, and Scopus) were reliable and widely used when seeking academic publications – the data that was gathered from these two sources was collected and cross-checked in order to ensure that the authors were correct in their driven conclusions.

The majority of the results are presented in the form of graphs that present the trends over a given time period. The graphs were analyzed, and the conclusions were drawn with regards to the peaks of interest, stability, and inclining/declining trends in popularity.

2.1. Keyword selections

The authors started the research by finding the most popular keywords among the Six Sigma-related publications that could be found in the Web of Science database. The Pareto graph that is presented in Figure 1 shows the frequency of the keywords that were triggered by the phrase "process improvement." To create the graph, 1000 publications that were dated during the period of 2020–present were selected and then sorted by relevance.

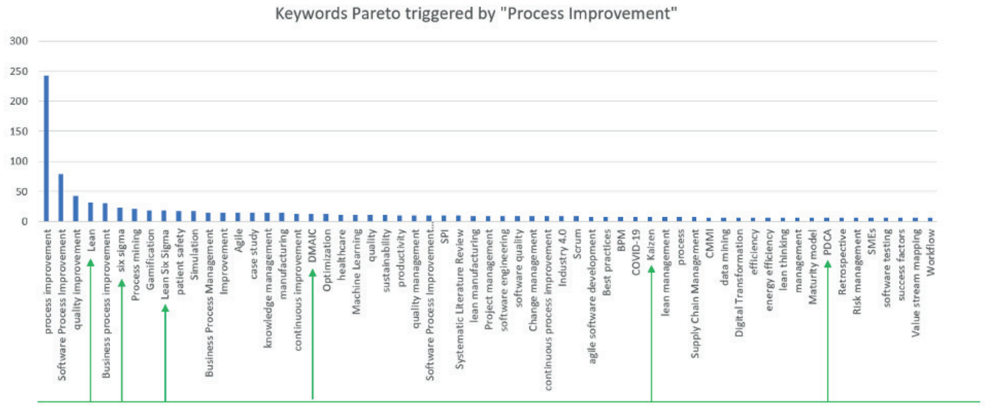


Fig. 1. *Frequency of keywords triggered by phrase "process improvement" (based on Web of Science search)*

Among the key phrases that were proposed by the authors of the publications, there were several phrases that were related to certain methods that were used for process improvement in quality management; those that are pointed out by the green arrows on the graph are Lean, Six Sigma, Lean Six Sigma, DMAIC, Kaizen, and PDCA. These terms are explained below.

The Lean (or lean manufacturing) philosophy focuses on reducing waste and non-value-added activities that do not create value for the customer. Organizations that adopt this philosophy have managed to improve their productivity, product quality, profitability, and competitiveness (Maware & Parsley, 2023).

Lean Six Sigma (LSS) combines Lean manufacturing principles and Six Sigma quality management in order to reduce waste and defects in business processes. LSS builds on the strengths of both approaches by emphasizing the importance of customer satisfaction, process improvement, and data-driven decision-making (Huang et al., 2023).

The DMAIC methodology in Six Sigma is a five-step improvement cycle: define, measure, analyze, improve, control (Palací-López et al., 2020).

Kaizen means "small, incremental, continuous improvement." Kaizen is a philosophy in the Lean system that focuses on both the process and the results; this is a process that, when done correctly, humanizes the workplace, eliminates unnecessarily hard work (both mental and physical), teaches people how to do rapid experiments using scientific methods, and eliminates waste in business processes (Prošić, 2011).

The PDCA cycle is a routinized and standardized way of working; this is a management process that is characterized by a spiraling cycle with large loops and small loops based on the principle of planning (P), execution (or do) (D), check (C), and action (A). The PDCA cycle continuously finds and solves problems in order to improve work efficiency (Zhong et al., 2023). The above phrases are going to be compared with the term "Six Sigma" according to their popularity and trends over the considered time period.

The second search was scoped to find trends among Six Sigma-related terms. In order to find proper keywords, those publications that could be found in the Web of Science database (triggered by the phrase "Six Sigma") were identified. Again, the first 1000 publications that were dated during the period of 2020–present were selected; then, they were sorted by relevance. The obtained results are presented in a Pareto graph (see Figure 2).

For further analysis, the following phrases were used: "Six Sigma," "Lean Six Sigma," "DMAIC," and "Six Sigma methodology." The phrase "total quality management" (TQM) was found among those phrases that appeared relatively frequently. TQM is a systematic management technique for developing a process-driven culture inside an organization in order to achieve quality as well as customer and employee satisfaction (Alawag et al., 2023). TQM was also used to compare Six Sigma with the other methods.

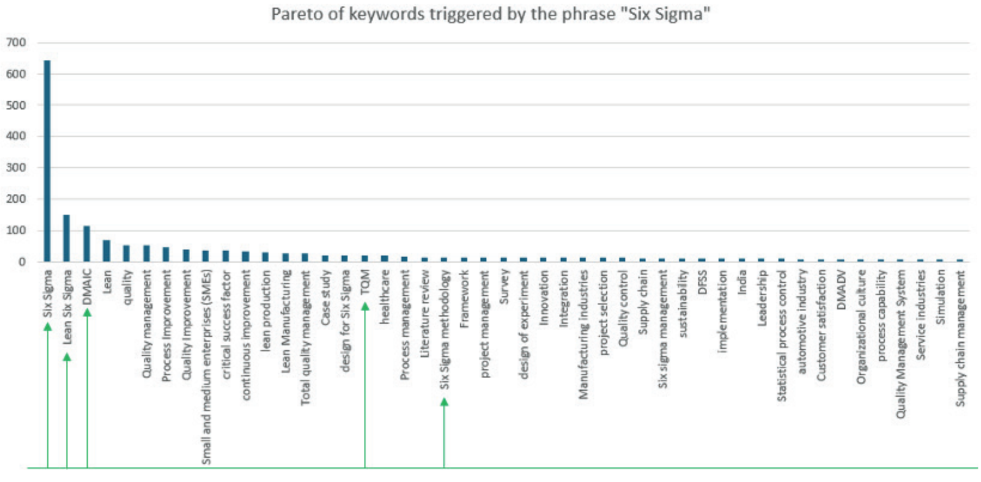


Fig. 2. *Frequency of keywords triggered by term "Six Sigma" (based on Web of Science search)*

2.2. Google Trends

Google Trends is a search-analysis tool that is offered by Google, Inc. (https://www.google.com/trends/); its main purpose is to analyze the behaviors of Google users while mapping human consciousness. This shows how often a specific search term or phrase is entered relative to the total search volume across various regions of the world as well as in different languages. It has been utilized to measure people's interest in a particular topic, displaying related queries and topics that are associated with the searched term/phrase. Google Trends has only been used to identify keyword search frequency since 2004. The keywords that were used to analyze the interest in Six Sigma itself were "Six Sigma," "Lean Six Sigma," "DMAIC," "Six Sigma methodology," and "Six Sigma certification." The keywords that were used to compare the interest between Six Sigma and the other methods were "Lean," "Six Sigma," "TQM," "Kaizen," and "PDCA."

2.3. Google Books Ngram Viewer

Google Books Ngram Viewer is a tool that was developed by Google that allows for an analysis of the frequency of the occurrences of the words and phrases in the language corpus that have been collected by Google Books. This corpus contains a vast number of books, journals, and other textual materials that have been scanned or entered into a database by Google.

Google Books Ngram Viewer enables users to search for words or phrases and presents the results in the form of a graph. The graph represents the frequency of the occurrences of a specific term/phrase over time. Users can customize the time range as well as the language in which the analysis is conducted.

Google Books Ngram Viewer has been used to demonstrate the frequency of the occurrences of words and phrases in the language corpus that has been collected by Google since 1985.

The keywords that were used to analyze the interest in Six Sigma itself were "Six Sigma," "Lean Six Sigma," "DMAIC," "Six Sigma methodology." and "Six Sigma certification." The keywords that were used to compare the interest between Six Sigma and the other methods were "Lean," "Six Sigma," "TQM," "Kaizen," and "PDCA."

2.4. Web of Science

Web of Science (WoS) is an information platform and research tool that provides access to a wide range of scientific publications, journals, and citation indexes. Web of Science enables the searching for and indexing of a significant number of scientific journals, conferences, patents, and other sources of scientific information.

The number of publications that have contained keywords on Web of Science since 1985 was analyzed. The keywords that were used to analyze the interest in Six Sigma itself were "Six Sigma," "Lean Six Sigma," "DMAIC," and "Six Sigma methodology." The keywords that were used to compare the interest between Six Sigma and the other methods were "Lean," "Six Sigma," "TQM," "Kaizen," and "PDCA."

2.5. Scopus

Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books, and conference proceedings. Delivering a comprehensive overview of the world's research output in the fields of science, technology, medicine, social sciences, and arts and humanities, Scopus features smart tools for tracking, analyzing, and visualizing research. Scopus is a database that is widely used by academics, business, and governments.

The number of publications that have contained keywords on Scopus since 1985 was analyzed. Just as with WoS, the keywords that were used to analyze the interest in Six Sigma itself were "Six Sigma," "Lean Six Sigma," "DMAIC," and "Six Sigma methodology." The keywords that were used to compare the interest between Six Sigma and the other methods were "Lean," "Six Sigma," "TQM," "Kaizen," and "PDCA."

3. RESULTS

3.1. Google Trends – Six Sigma and related terms

As explained in the method section of this paper, each analysis was split into two areas: 1) focusing on Six Sigma and Six Sigma synonyms to understand the trends over the years; and 2) focusing on comparisons between Six Sigma and other quality-improvement methods. Six Sigma and those terms/phrases that could be considered to be Six Sigma-related are presented on the following graphs.

The graph that is shown in Figure 3 presents the trends in the popularity of the term "Six Sigma" and the terms and phrases that were related to Six Sigma during the period of 2004 through 2024. Google Trends presented the data as time-based results with their relations to the highest achieved results over the given period. In the case of the graph above – the term "Six Sigma" significantly dominated other terms with a similar meaning. The time-based analysis clearly showed a decreasing trend, with a significant drop during the period of 2004–2012. After 2012, a decrease in popularity could also be observed; however, it was not as significant as during the preceding eight years. The data that was obtained for the last six years suggested a certain stability without major drops or increases in the method's popularity.

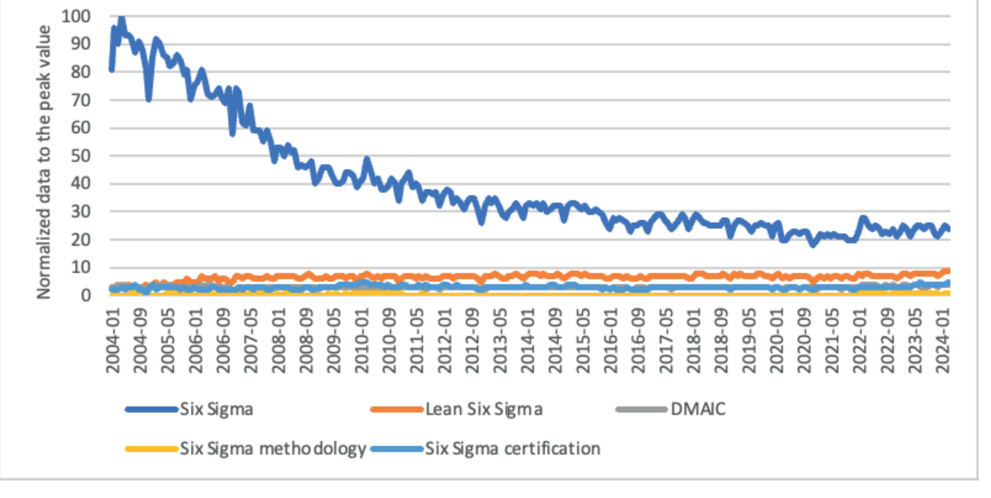


Fig. 3. *Google Trends – Six Sigma and Six Sigma-related terms over period of 2004–present*

Since the term "Six Sigma" significantly dominated the other terms/phrases that were used in the analysis (as was previously mentioned), the authors decided to remove the term from the comparison and focus on the four remaining terms/phrases (see Figure 4). Among the used terms/phrases, the greatest interest could be found with "Lean Six Sigma"; however, the differences between this and the other terms/phrases were not significant. Unlike in the previous graph, no major drop in interest could be observed; all of the terms/phrases seemed to maintain similar levels for nearly 20 years.

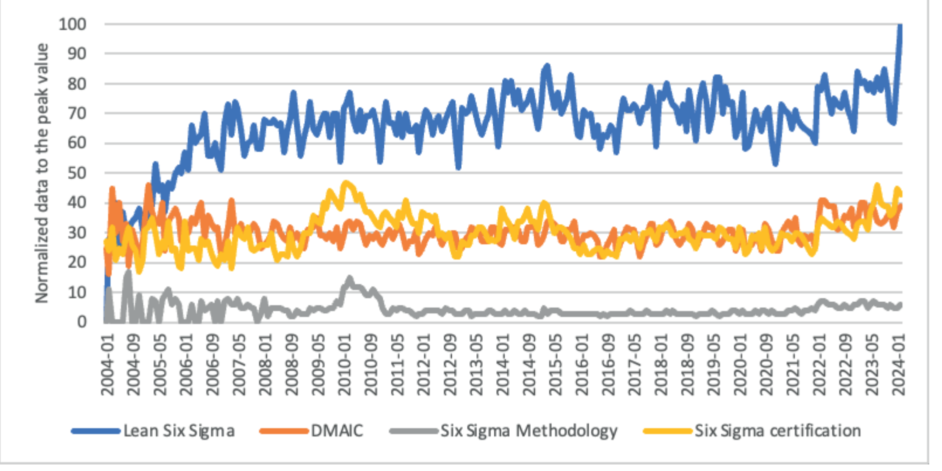


Fig. 4. *Google Trends – Six Sigma-related terms (without "Six Sigma") over period of 2004–present*

A observation was made by the authors while searching for the rationale of the drop of interest during the period of 2004–2012. Similar trends that showed drops of interest within the first eight years of the available data could also be found for other terms and phrases that were related to industry processes; e.g., production, quality, engineering, and manufacturing (see Figure 5). All of these terms/phrases reached their peaks of interest in 2004 (which was similar to the interest in the term "Six Sigma"). Like the Six Sigma trend, the following years featured waning interest for these terms until they finally reached a stable period, at last.

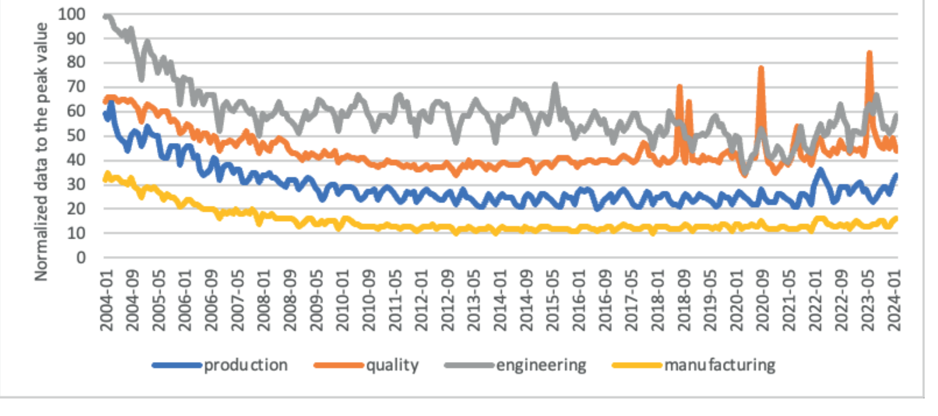


Fig. 5. *Google Trends – industry-related terms over period of 2004–present*

3.2. Google Trends – Six Sigma comparison with other methods

The second analysis focused on comparing Six Sigma with the other methods that were known in the process-improvement area. Among the terms/phrases that were selected in the analysis were "Lean," "Kaizen," "TQM," "PDCA," and "Six Sigma" (see Figure 6). The data from Google Trends is shown starting in 2004.

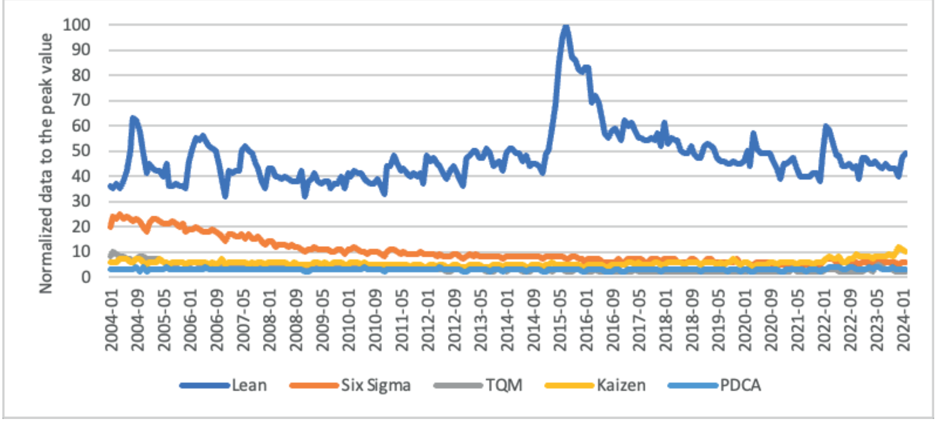


Fig. 6. *Google Trends – Six Sigma versus other process-improvement methods over period of 2004–present*

The dominant term in order of interest within this term selection was "lean." The term was significantly more popular than the others though the whole time range that is presented in this study. "Lean" had stable search results over the analyzed period – with one significant peak of interest around the years of 2015 and 2016. An important note to this search is that the term "Lean" does not have only one meaning; it is likely that the data was affected by searches that were not intended to find process-improvement methods. Since "Lean" significantly dominates on the graph, the authors decided to create another analysis without this term. The obtained results are presented in Figure 7.

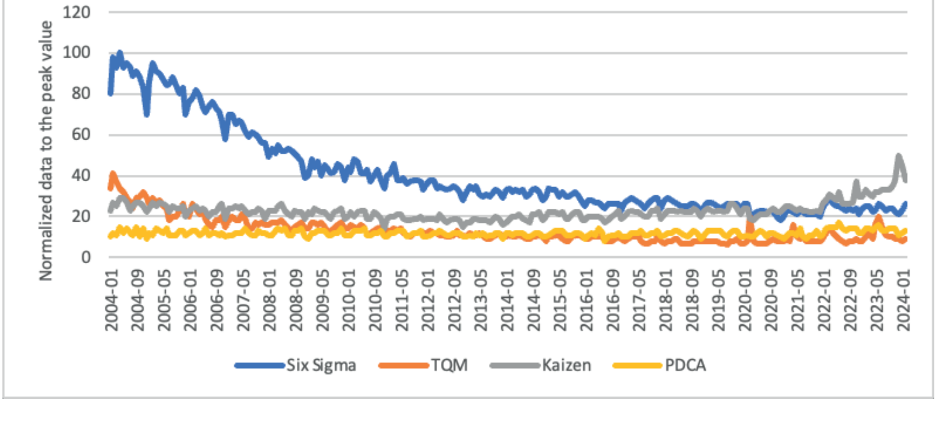


Fig. 7. *Google Trends – Six Sigma versus other process-improvement methods (without "Lean") over period of 2004–present*

The trends in the interest of the search of the term "Six Sigma" were described above. In terms of a comparison, the graph presents a dominating picture of Six Sigma during the first 12 years of the analysis. Since 2016, the terms "Six Sigma," "TQM," "Kaizen," and "PDCA" have been in near proximity to each other with regard to the numbers of registered searches. Over the last two years, the term "Six Sigma" has shown less popularity than term "Kaizen," while "TQM," "Kaizen," and "PDCA" have been stable over the analyzed period (with no major changes in their relative numbers of searches).

3.3. Google Books Ngram Viewer – Six Sigma and related terms

The same methods of research that were utilized in the Google Trends-based analysis were used to analyze the data in the Google Books Ngram Viewer. The first analysis focused on Six Sigma and Six Sigma-related terms over the period of 1985–2019. These results are presented in Figure 8.

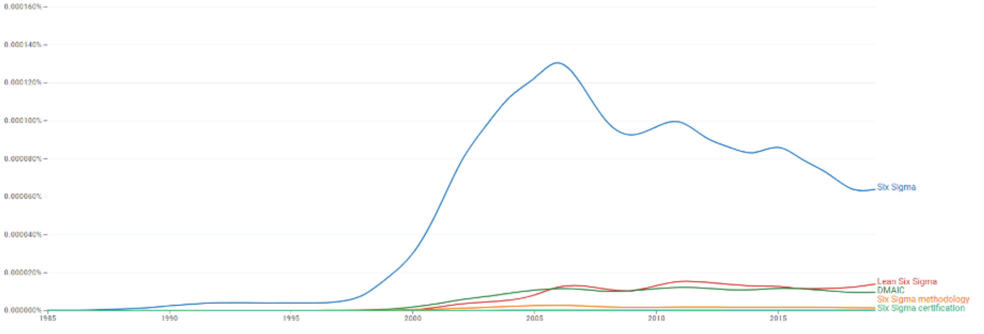


Fig. 8. *Google Books Ngram Viewer – Six Sigma and Six Sigma-related terms/phrases over period of 1985–2019*

Again, dominating on the graph is the term "Six Sigma," with major differences when compared to the terms "DMAIC," and "Lean Six Sigma". Both "Six Sigma methodology" and "Six Sigma certification" represented negligible numbers of results.

In terms of trends in the data, an increase in the number of occurrences of the term "Six Sigma" could be found in the literature between 1996 and 2006; this clearly indicated the rising popularity of the method. Since its peak interest in 2006, this trend has been continuously decreasing; this might suggest lower interest in the method. For the terms "Lean Six Sigma" and "DMAIC," the data indicated stability over the whole time period of 2005 through 2019.

3.4. Google Books Ngram Viewer – Six Sigma comparison with other methods

Similar to the analysis that was done with Google Trends, Google Books Ngram Viewer was used to compare Six Sigma with the other methods. The data from the years 1985 through 2019 is shown in Figure 9.

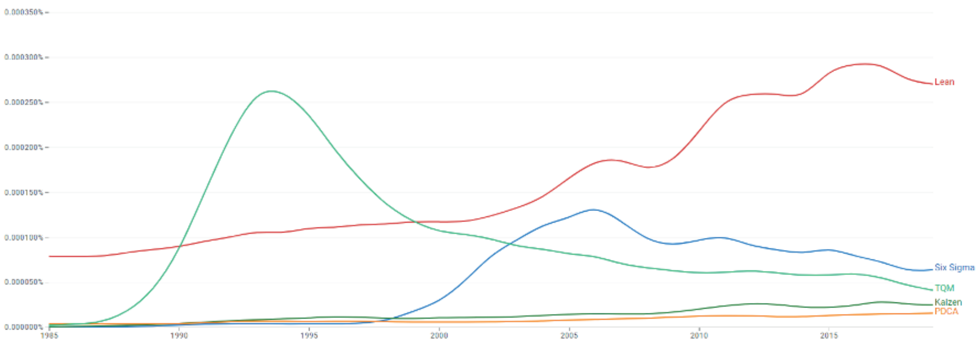


Fig. 9. *Google Books Ngram Viewer – Six Sigma versus other process-improvement methods over period of 1985–2019*

In recent years, the term "Lean" has led in terms of the number of occurrences in the literature that has been covered by the Google database; this term has been significantly more popular than the other terms. "Lean" has enjoyed a continuous rising trend since 1985. One should remember, however, that the word "lean" has more meanings than the one that is used for process improvements; this could have affected the data.

An interesting trend could be observed in the case of the term "TQM." After 1985, the data showed increasing occurrences of the term in publications (reaching its peak in 1994). After this, the data showed a significant decline through 2000; then, it endured a continuous but significant drop over the period of 2000–2019. The data that was related to "Six Sigma" seemed to reflect a similar trend. Data present apparent shift of ten years , however.

3.5. Web of Science – Six Sigma and related terms

The data that was collected from Web of Science was from the years 1985 through 2024. The graph in Figure 10 shows the total number of publications that were triggered by "Six Sigma" and Six Sigma-related terms as a percentage ratio to the total number of publications that were available on Web of Science.

Similar to the findings from the application of Google Books Ngram Viewer, "Six Sigma" showed an increasing number of publications starting in 1995 and growing year by year until its peak of interest in 2007. After 2007, the number of publications slowly decreased through 2023. The data for 2024 cannot be compared with the previous year's as of yet, since the year is not yet completed; this year's trend is similar in its shape to the results from Google Books Ngram Viewer.

In terms of the number of publications, the other Six Sigma-related terms were less popular than the term "Six Sigma." Nonetheless, both "Lean Six Sigma" and "DMAIC" have shown positive trends over the years since their first use (starting in around 2000).

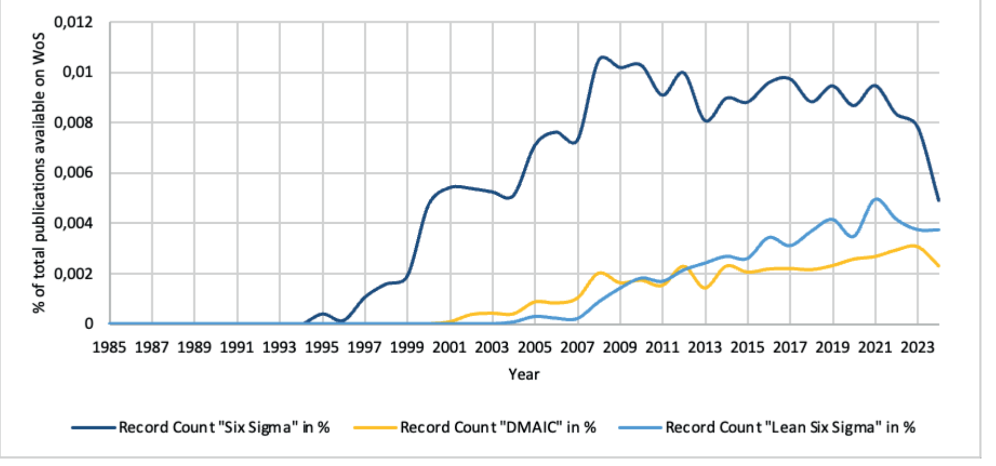


Fig. 10. *Web of Science – Six Sigma and Six Sigma-related terms over period of 1985–present*

3.6. Web of Science – Six Sigma comparison with other methods

The data from 1985 onward was analyzed in order to understand Six Sigma's popularity versus the other process-improvement methods; these results are shown in Figure 11.

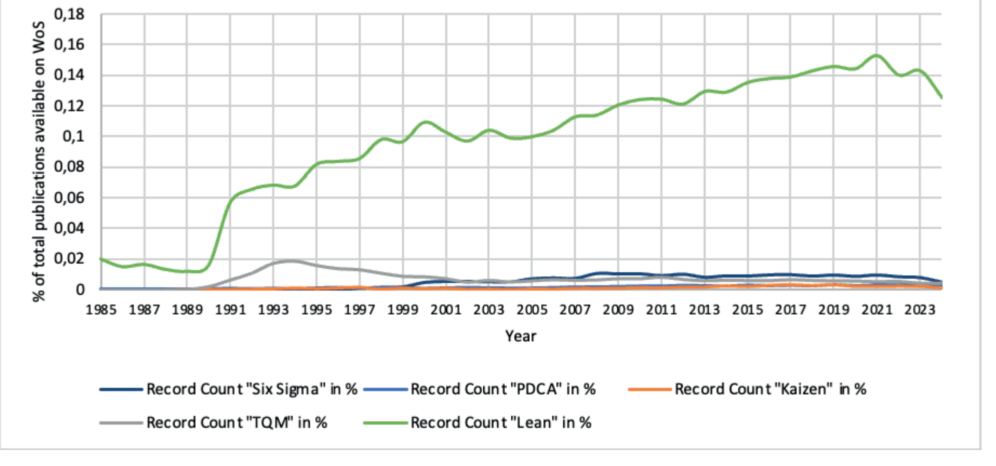


Fig. 11. *Web of Science – Six Sigma versus other process-improvement methods over period of 1985–present*

Like in the other graphs that were previously presented, the data that was available on Web of Science again showed the domination of the term "Lean" among the available publications. The term "Lean" has shown a continuously increasing trend since 1990. Similar to Google Trends and Google Books Ngram Viewer, the data that was triggered by the word "lean" could have referred to not only the process-improvement method, as the word is frequently used in the other fields. To understand Six Sigma's popularity versus the other methods, a similar graph was created without the term "Lean" (see Figure 12).

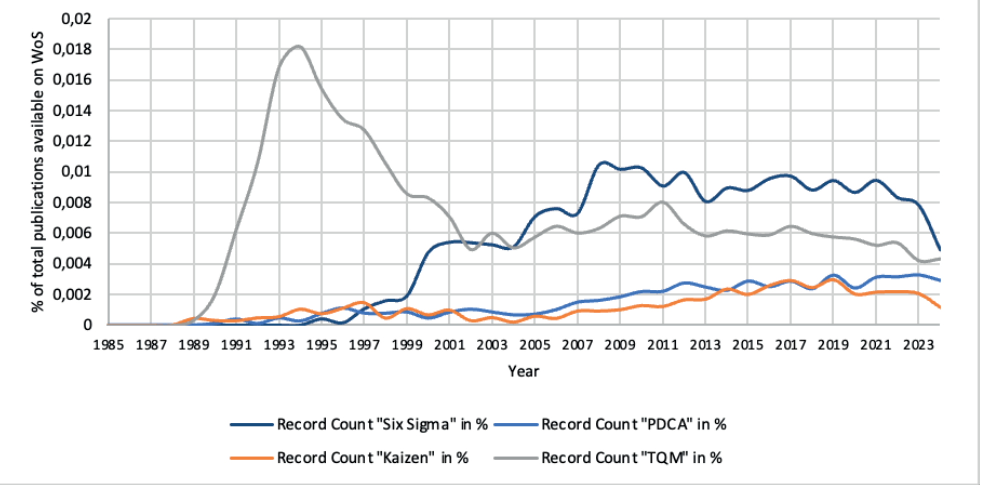


Fig. 12. *Web of Science – Six Sigma versus other process-improvement methods over period of 1985–present (without term "Lean")*

Over the last 20 years, Six Sigma has shown the greatest popularity as compared to the other methods (PDCA, Kaizen, and TQM). Similar to Google Books Ngram Viewer, the term "TQM" showed significant increase over the period of 1989–1994. After 1994, TQM showed drops in its frequency in publications; this trend finally stabilized in 2003.

3.7. Scopus – Six Sigma and related terms

A similar approach was taken to analyze the Scopus database. First, the term "Six Sigma" and its related terms were analyzed, taking those publications that were available during the period of 1985–present into consideration. Figure 13 shows the numbers of publications that were related to the terms "Six Sigma," "Lean Six Sigma", and "DMAIC," as percentages of the total numbers of publications that were available on Scopus.

The early years of Six Sigma's presence among scientific publications was stable and at a low in terms of the total percentage of the publications that were available on Scopus. Starting in 2000, the number of publications rapidly grew to reach its peak in 2006. After 2008, a declining trend could be observed until reaching a stable area in or around 2013. Since then, the trend has been constant up until now.

The other terms that were taken into consideration in this analysis ("Lean Six Sigma," and "DMAIC") showed similar trends, with continuous increases in the numbers of publications from 2000 through 2020. Throughout the whole analyzed time period, there were fewer publications that were related to these terms than those that were related to the term "Six Sigma."

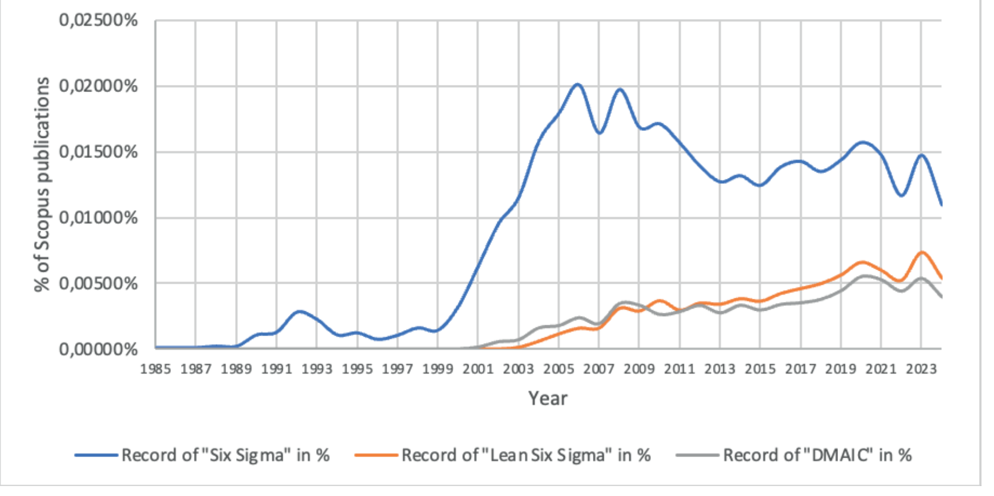


Fig. 13. *Scopus – Six Sigma and Six Sigma-related terms over period of 1985–present*

Besides the numbers of publications, the Scopus database allows us to understand the types of publications that are related to a given term. In terms of Six Sigma, most of the published works were in the forms of articles or scientific papers during the period of 1985–present. The results of the numbers of publications and the percentages of the total numbers of publications are shown in the table below (see Table 1 and Figure 14).

Table 1. *Scopus – Six Sigma during period of 1985–present by publication type*

Document type	Number of publications	Share [%]
Article	4626	48.9
Conference paper	3095	32.7
Review	623	6.6
Book chapter	408	4.3
Conference review	221	2.4
Book	182	1.9
Short survey	85	0.9
Other	219	2.3
Total	9459	100

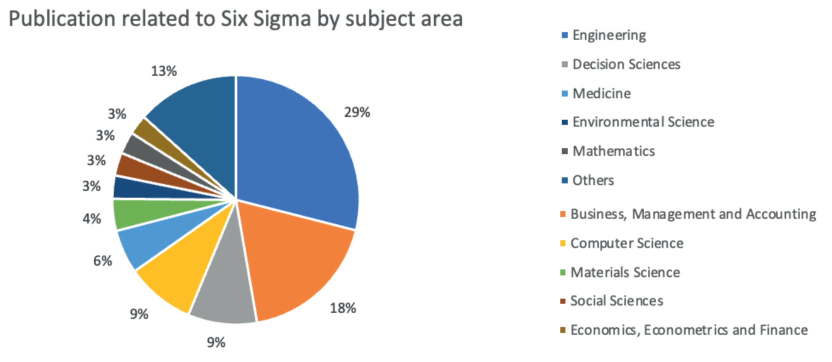


Fig. 14. *Scopus – Six Sigma in publications by scientific discipline during period of 1985–present*

Scopus allowed us to also understand the scientific disciplines that were related to the term "Six Sigma" during the period of 1985–present. The most common discipline to reach for the Six Sigma method has been engineering, followed by business, management, and accounting as well as decision science. A summary of these results are presented in the graph presented in Figure 14.

Scopus – Six Sigma comparison with other methods

Similar to the WoS analysis, the data from Scopus was collected in order to compare Six Sigma with other process-improvement approaches like Lean, Kaizen, PDCA, and TQM. The data from the time period of 1985–present was analyzed; this is presented in Figure 15 as the year-by-year percentages of the total numbers of publications that have been available on Scopus.

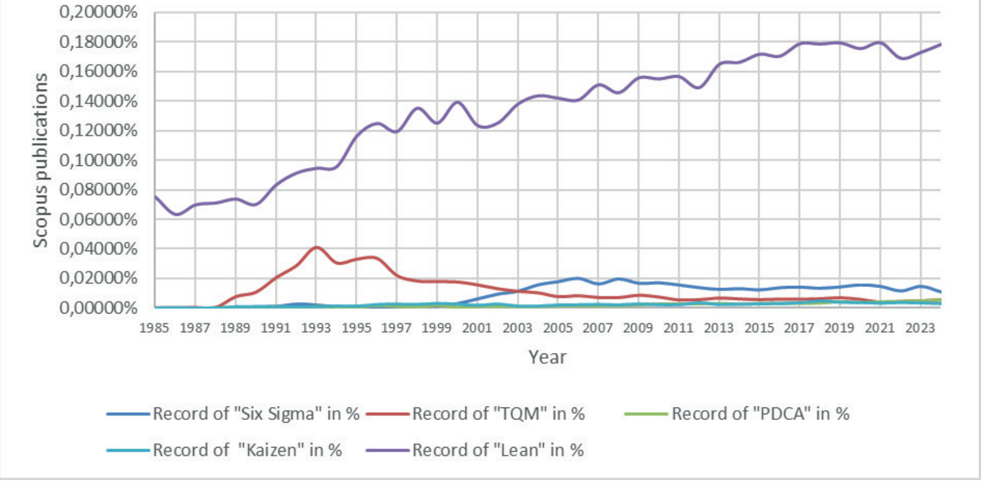


Fig. 15. *Scopus – Six Sigma versus other process-improvement methods over period of 1985–present*

Similar to the WoS results, the dominant position of the term "Lean" can be observed. Since the beginning of the analyzed time period, the trend for this term has increased. Since the word "lean" is used in many different scientific areas (not only those that are related to process improvement), the data was prepared without the term "Lean" in order to better understand Six Sigma's popularity versus the other methods. These results are shown in Figure 16.

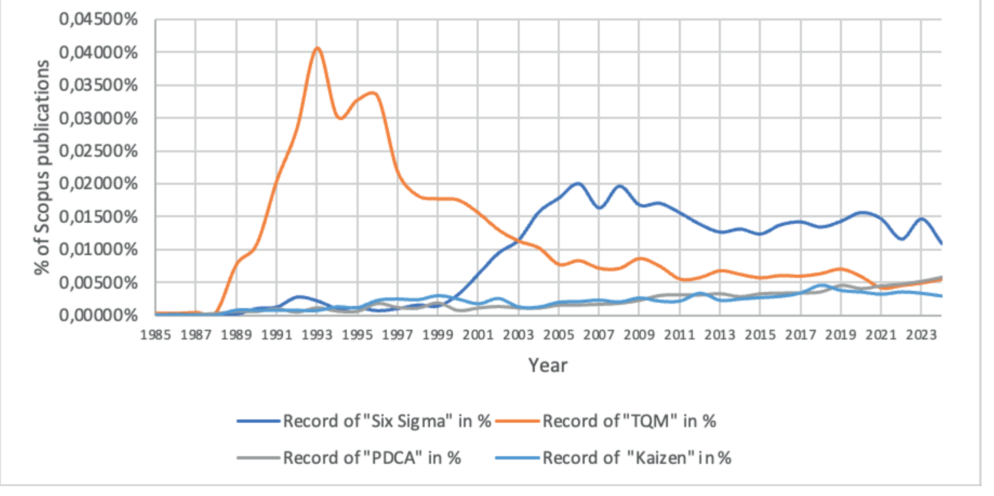


Fig. 16. *Scopus – Six Sigma versus other process-improvement methods over years of 1985–present (without term "Lean")*

The left part of the graph (related to the years before 2000) shows a significant interest in TQM. Starting in 1988 (and reaching its peak in 1993), the term "TQM" was more frequently used than the other terms that are analyzed in this study. After 1993, a decrease of interest could be observed in this term until 2005 (when the stable period began). In 2003, "Six Sigma" surpassed "TQM" in terms of the numbers of publications where these terms were utilized. This trend is similar to the Six Sigma results, where Six Sigma-related publications rapidly grew during the period of 2000–2006 and declined during the period of 2008–2013, finally attaining stability after 2013.

There have been significantly fewer publications that have been related to the terms "PDCA" and "Kaizen" than have been related to "Six Sigma" and "TQM." The trends for PDCA and Kaizen have been growing continuously since 1987.

4. DISCUSSION

4.1. Six Sigma and Six Sigma-related terms

During the study, the authors have analyzed the results of the interest in Six Sigma and Six Sigma-related terms over the years – reaching back as far as 1985. The conclusion that came from all three of the analyzed sources was that the method was either on a declining trend in terms of its popularity or in the area of stable interest.

Figure 17 illustrates all of the analyzed trends that came from Google Trends, Google Books Ngram Viewer, and Web of Science (normalized in one graph); the reader should focus on the time scale and the shapes of the curves. The graph was created in order to show the interest in the method over the given time period and check the consistency of the data that came from the different sources.

All of the sources represented similar trends – the popularity of Six Sigma continuously increased starting in 1985 until it reached its peak during the period of 2000–2010. After this, all three of the sources showed decreasing interest, finally reaching a period of stability or slight continuous decrease over the past five years.

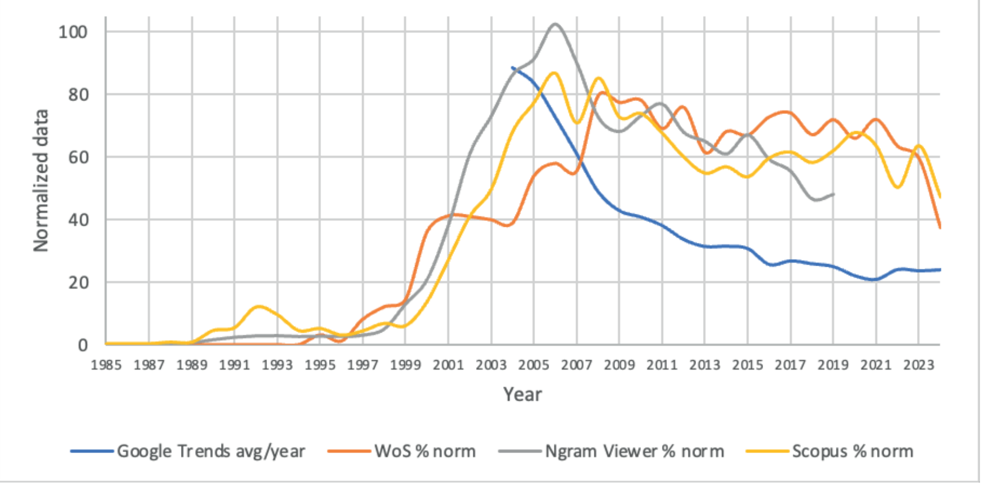


Fig. 17. *Normalized results – data from four sources (Google Trends, Google Books Ngram Viewer, Web of Science, and Scopus) presented on time-based scale during period of 1985–present*

In the shape of the graph that was described above, this analysis is like the typical S-curve of a product's life cycle (see Figure 18). Typically, an S-curve represents a product's life cycle in the following stages: introduction, growth, maturity, and decline; eventually, the product is retired from the market.

Combining the theory that was presented above with the previously shown data about Six Sigma, the following conclusions could be made:

1) *Introduction* phase of Six Sigma (1985–1995) – represented by slow increase in number of publications;
2) *Growth* phase of Six Sigma (1995–2005) – represented by significant increase in number of publications observed on year-by-year basis;
3) *Maturity* phase of Six Sigma (2005–2015) – when data shows peak of interest and slow decrease in numbers of publications;
4) *Decline* phase of Six Sigma (2015–present) – where significant decreases in popularity and numbers of related publications can be observed.

Fig. 18. *S-curve for typical product life cycle (Sanders, 2024)*

Analyzing the S-curve concept of a product's life cycle, it is typical that one product or concept is followed by the next invention while in its decline phase. Among others things, Six Sigma brought about special attention regarding data to industry. Given the common data-driven approaches, it is possible that, following Six Sigma, the concepts that are based on machine-learning or big data are those inventions.

4.2. Six Sigma and other process-improvement methods

Among all of the data that was analyzed from the three sources, the term "Lean" significantly dominated. As mentioned in the previous sections, the word has multiple meanings; therefore, only a fraction of the available publications that were triggered by the word were related to process or quality improvements. Without a deeper analysis (and based on the presented data), a direct comparison between "Lean" and "Six Sigma" cannot be conclusive.

Without the term "Lean" in the picture, Six Sigma has tended to be more popular and featured in more publications than the other methods; however, the differences were not significant. Following the product life-cycle analysis, a conclusion can be made that all of the presented methods reached periods of maturity.

The observation that was made above was especially clear when analyzing the data that was related to the numbers of publications that were triggered by the term "TQM." Here, all four phases of a product's life cycle can be observed: introduction (1985–1990), growth (1990–1993), maturity (1993–1996), and decline (1996–present). The decline phase can be characterized by a long stability period over the past 20 years.

Since the early years of the TQM method occurred before the development of Six Sigma, it was highly likely that Six Sigma would show a similar trend in popularity. This means that the method has already entered its long period of stability and its interest will stay at its current level for the future years.

5. CONCLUSIONS

In this article, the authors studied the trend of Six Sigma's popularity over the years based on the numbers of searches in Google Trends and the numbers of publications or mentions that were available on the Google Books Ngram Viewer, Web of Science, and Scopus databases. The data was analyzed for the period of 1985–present and was analyzed in two ways: looking at the trends over the years for the term "Six Sigma" and other directly related terms in a comparable way, and looking at Six Sigma versus other process-improvement methods.

Based on the conducted research, a conclusion can be made that the interest in the topic of Six Sigma has reached the stable phase at levels that are definitely lower than its peak after years of decline. Given the S-Curve theory of a product's life cycle, this would indicate that Six Sigma is in its decline phase and that industry is already using new methods in process and quality improvements more frequently.

On the other hand, the stable interest over the last few years could indicate that the concept of Six Sigma was (and is still) a strong foundation for process improvements among industries. Following this thought, it can be assumed that Six Sigma's DMAIC data-driven approach and all of the tools that are utilized by the method have become a core asset of knowledge in the field of problem-solving and process improvements.

What is more, emerging trends like Big Data and Industry 4.0 can open new pathways for Six Sigma practitioners, as the integration of Six Sigma with Big Data can provide superior results for many organizations (Antony et al., 2022). Process-mining can serve as an important support technology for process-improvement frameworks such as Six Sigma (Graafmans et al., 2021). In this case, it would be feasible for Six Sigma to show positive inclining trends among searches and publications in the future.

REFERENCES

Alawag A.M., Alaloul M.S., Liew W.S., Baarimah A.O., Musarat M.A. & Al-Mekhlafi A.-B.A. (2023). The role of the Total-Quality-Management (TQM) drivers in overcoming the challenges of implementing TQM in Industrialized-Building-System (IBS) Projects in Malaysia: Experts' Perspectives. *Sustainability*, **15**(8), 6607. **DOI**: https://doi.org/10.3390/su15086607.

Antony J., Sony M. & Gutierrez L. (2022). An empirical study into the limitations and emerging trends of Six Sigma: Findings from a global survey. *IEEE Transactions on Engineering Management*, **69**(5), pp. 2088–2101. **DOI**: https://doi.org/10.1109/TEM.2020.2995168.

De Feo J.A. & Barnard W. (2005). *Juran Institute's Six Sigma Breakthrough and Beyond: Quality Performance Breakthrough Methods*. New York: McGraw-Hill Professional.

Dusharme D. (2024). *Six Sigma Survey: Breaking Through the Six Sigma Hype*. **URL**: https://www.qualitydigest.com/static/magazine/nov01/html/sixsigmaarticle.html [24.03.2024].

Eckes G. (2001). *The Six Sigma Revolution: How General Electric and Others Turned Process into Profits*. New York: Wiley.

Goh T.N. (2020). Six triumphs and six tragedies of Six Sigma. *Quality Engineering*, **22**(4), pp. 299–305. **DOI**: https://doi.org/10.1080/08982112.2010.495102.

Graafmans T., Turetken O., Poppelaars H. & Fahland D. (2021). Process mining for Six Sigma. *Business & Information Systems Engineering*, **63**(12), pp. 277–300. **DOI**: https://link.springer.com/article/10.1007/s12599-020-00649-w.

Huang J., Irfan M., Fatima S.S. & Shahid R.M. (2023). The role of lean six sigma in driving sustainable manufacturing practices: an analysis of the relationship between lean six sigma principles, data-driven decision making, and environmental performance. *Frontiers in Environmental Science*, **11**, 1184488. **DOI**: https://doi.org/10.3389/fenvs.2023.1184488.

Maware C. & Parsley D.M. (2023). Can Industry 4.0 assist lean manufacturing in attaining sustainability over time? Evidence from the US organizations. *Sustainability*, **15**(3), 1962. **DOI**: https://doi.org/10.3390/su15031962.

Muralidharan K. (2015). *Six Sigma for organizational excellence. A statistical approach*. Springer (India).

Palací-López D., Borrŕs-Ferrís J., da Silva de Oliveria L.T. & Ferrer A. (2020). Multivariate Six Sigma: A case study in Industry 4.0. *Processes*, **8**(9), 1119. **DOI**: https://doi.org/10.3390/pr8091119.

Prošić S. (2011). Kaizen management philosophy. In: *I International Symposium Engineering Management And Competitiveness 2011 (EMC 2011)*. Zrenjanin, Serbia, June 24–25, 2011, pp. 173–178.

Sanders A. (2024). *What Is the Product Life Cycle? Stages and Examples*. **URL**: https://www.thestreet.com/markets/commodities/productlife-cycle-14882534 [24.03.2024].

Sokovic M., Pavletic D. & Fakin S. (2005). Application of Six Sigma methodology for process design. *Journal of Processing Materials Technology*, **162–163**(3), pp. 777–783. **DOI**: https://doi.org/10.1016/j.jmatprotec.2005.02.231.

Uluskan M. (2022). Enhancing Six Sigma Understanding: Insights into Various Dimensions and Aspects of Six Sigma, *Engineering Management Journal*, **34**(1), pp. 99–128. **DOI**: https://doi.org/10.1080/10429247.2020.1852806.

Watson G.H. (1994). *The Motorola Six Sigma System Breakthroughs*. New York: McGraw-Hill.

Zhong X., Wu X. & Xie X. (2023). A descriptive study on clinical department managers' cognition of the Plan-Do-Check-Act cycle and factors influencing their cognition. *BMC Medical Education* **23**(1), 294. **DOI**: https://doi.org/10.1186/s12909-023-04293-2.

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 39–64

Predictions and Application of Queueing Analysis: Case of Regional Hospital Limbe, Cameroon

Daphne T. Machangara*, Habiboulaye Amadou Boubacar**, Giovanni Andreatta***, Antony Ndolo****

Abstract. In this work, we applied queue analysis and the predictions of waiting times at Regional Hospital Limbe (RHL) in Cameroon. The main purpose of the work was to be able to make mathematical sense of a real-life scenario that concerned queues (waiting lines) and try to come up with models for performance measurements and improvements; this could be achieved by using queueing theory concepts that were composed of queueing models that provided some operational insights because of their analytical nature. The observations included studying patient arrival and waiting times, along with doctor service times; the results showed busy departments in the hospital, busy days, and busy times. Long waiting times were mainly found to exist in general practitioner (GP) and specialist consultations. The queueing concept was applied to only one service segment – GP consultation. Although strong scientific conclusions cannot be made on the queuing models that were obtained due to inefficient data, the value of this work lies mainly in the methodologies and proposals of different operating systems that could be adopted. Furthermore, some predictions were made using machine learning to see how long a patient could wait in a queue for service; the model predictions had an average of 10 minutes and 53 seconds of error.

Keywords: queueing theory, queueing model, queueing system, waiting time, prediction

Mathematics Subject Classification: 60K25

JEL Classification: C53

Submitted: July 15, 2024

Revised: October 21, 2024

* African Institute for Mathematical Sciences (AIMS), Limbe, Cameroon, e-mail: daphne.machangara@aims-cameroon.org
** Universite de Lille, Lille, France
*** Università degli Studi di Padova: Padova, Veneto, Italy
**** Karadeniz Technical University, Trabzon, Türkiye; Deep Learning Indaba, Africa, e-mail: ndolo@deeplearningindaba.com

DOI: https://doi.org/10.7494/dmms.2024.18.6442

1. INTRODUCTION

Waiting times are generally a problem, and many have tried to tackle this issue. Over the years, there has been growing evidence of long waiting times in healthcare organizations, thus leading to different waiting times for patients (Garcia-Corchero & Jimenez-Rubio, 2022); therefore, long queues and waiting times negatively impact healthcare service deliveries through patient outcomes and satisfaction in terms of increased risks of delays when attending to time-sensitive issues and the outcomes of patient frustrations. As a result, this work sought to make mathematical sense of queues, adopt models for performance measurements and improvements, and provide some waiting-time predictions for patients; this was achieved by understanding the arrival and waiting/service times as well as the busy times, busy departments, and busy days at our hospital. Moreover, at least 90% of patients should be seen within 30 minutes of their scheduled appointment times according to the recommendation by the Institute of Medicine (IOM) (O'Malley et al., 1983). Whether this is being implemented remains a challenge that is faced by many. We studied this problem in our thesis (Machangara, 2018), and this paper is a continuation of this thesis.

Queueing theory was pioneered in the 20$^{\text{th}}$ century by Agner Krarup Erlang – a Danish mathematician, statistician, and engineer (Brockmeyer et al., 1948; Cooper, 1981; Erlang, 1909; Lakatos et al., 2019; Saaty, 1957). It continues to be a vital interdisciplinary field of study that explores how queues (or waiting lines) form and examines how they can be efficiently managed. The fundamental principles of queueing theory revolve around reviewing waiting-line processes. The main principle of this theory by Erlang (1909) paved the way by developing models to analyze a Copenhagen telephone exchange. An improvement to the theory was done by Palm (1943), thus extending Erlang's work, introducing the concepts of random arrivals and intensity fluctuations in a queueing system. This was particularly pertinent in all of the service disciplines that covered various sectors, like telecommunications (network routing, load balancing, and packet-switching systems (Dshalalow, 1995) and manufacturing/production (determining the optimal service rates for machinery or labor) (Boukas et al., 1995; Liberopoulos et al., 2006; Yao, 1994). In the energy sector, queueing theory is in smart-grid management and power-plant scheduling (as was commented on in (Nair et al., 2021; Zavanella et al., 2015). At the same time, this theory is also used in financial service, retail, and service industries to optimize the operations of trading systems, ATM networks, retail outlets, restaurants, banks, and call centers (Koole, 2013). More recently, its application in the optimizations of cloud-computing resources, data-center operations, and task scheduling in distributed systems has increased rapidly (Harchol-Balter, 2013; Newell, 1982). There will undoubtedly be other possible future applications; the possible fusion of artificial intelligence and deep learning with queueing theory will lead to intelligent queuing systems. The application of distributed queueing systems integrated with blockchain technology is also suggested (Xiong, 2023). In healthcare, this can be applied in emergency-room-service, patient-triage, and appointment systems (Green, 2013; Harki, 2024); thus, the formations of queues exist in day-to-day life. The various applications of queueing theory are summarized in Figure 1 (as speculated by Elalouf & Wachtel, 2021 as well as by Shortle et al., 2018).

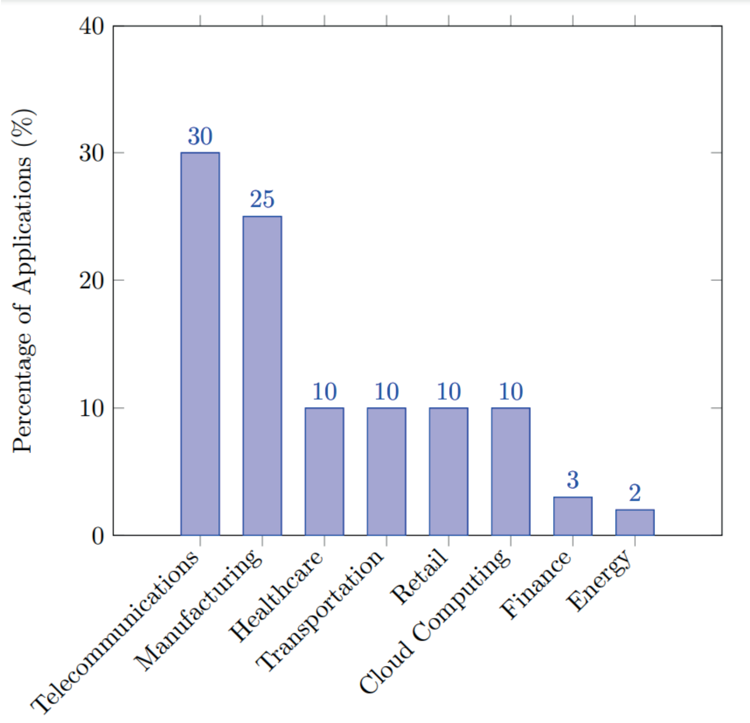


Fig. 1. *Percentages of applications of queueing theory in different fields*

There are factors that affect waiting times; for example, Biya et al. (2022) found the distance that was traveled, the day, and the time of a hospital visit to affect patients' waiting times. Thus, those patients who travel from afar, visit a hospital on a Monday, and arrive in the early morning were found to spend more waiting time than others.

Queueing theory has been extensively utilized in computer science for managing processes efficiently. Ding (2023) provided a comprehensive analysis of computer queueing theory – from its historical background and fundamental concepts to its mathematical models. The paper highlighted key contributions of queueing theory in computer systems; it also examined the characteristics of queues, like arrival and service rates, queue lengths, and waiting times, using different queueing models like single-server queues, multi-server queues, and network queues as well as their relevance and applicability in computer systems. Thus, some of the practical applications of queueing theory in computer systems include performance analyses. The authors of Chakka et al. (2009) introduced and proposed a generalized Markovian queueing model for performance analyses of telecommunications networks. The authors indicated that the model was a potential viable performance predictor, with its applications in the performance analysis of high-speed downlink packet access and optical burst switching nodes. Research by Li (2024) focused on analyzing and optimizing queueing systems in the context of cloud computing, with attention being paid to energy conservation. *M&M* queues tend to be widely used in cloud computing in its queueing technology, in physical queueing scenarios like grocery stores, fast food establishments, or the queues at airport terminals before boarding airplanes. The study examined the performance of $M\,|M\,|$ queues for developing efficient and expedient queues for cost-effectiveness. The $M\,|M\,|C$ queueing models were discovered to be of benefit when handling tasks of varying sizes and greater income levels as compared to the $M\,|M\,|1$ models.

In the health sector, queueing theory has been utilized in various departments in helping to identify and analyze existing queues by providing insights on queue management and resource allocation for serving patients with minimal waiting times (Saastamoinen et al., 2023). Elalouf & Wachtel (2021) demonstrated the use of both scientific methodologies like queueing models and simulations as well as managerial approaches like bed management to manage queueing-related problems in hospital emergency departments. Although queueing models have been used in healthcare systems, Peter & Sivasamy (2021) noted the limitations in its applications. For this reason, the authors investigated the applicability of queueing techniques to understand queueing modeling and solution techniques that are useful in applications. Ji (2023) described queueing theory as a mathematical approach that concerned the dynamics of waiting lines (or queues); thus, its applications were pivotal in analyzing and optimizing systems where the timing of customer arrivals and service was of importance in various sectors. Furthermore, the theory provided insights into patient arrival, patient flow, and the ways of optimizing both processes and people to meet demand according to (Johnston et al., 2022).

Waiting times not only differ in countries but even in the centers within these countries. This is a major problem for both developed and developing countries, and some reports on long waiting times have been made (although most date back to as long as 20 years ago).

Some of the studies that have been carried out have noted the management of waiting times as a challenge in hospitals (Mbwogge et al., 2022), where wait times and patient satisfaction were assessed in an eye hospital in Cameroon. The study revealed the lack of a significant association between wait times and satisfaction, with mean pre-intervention waiting, service, and idling times of 449.6, 111.9, and 337.7 minutes, respectively. Through the application of the Plan-Do-Study-Act (PDSA) quality improvement method, a 14.5% reduction in waiting time was achieved. Furthermore, Almusawi et al. (2023) did a study in Saudi Arabia at primary healthcare centers in Riyadh. Excluding emergency cases, the median total waiting time was found to be 23 minutes, whereas the median waiting times before and during the service were 6 and 6.78 minutes, respectively. The total waiting times at urban primary healthcare centers were longer than at rural primary healthcare centers, with a significant difference between both groups ($t = -15.5$, $P < 0.001$). Among the significant factors affecting waiting times on weekdays, they identified a patient's age, marital status, educational level, and occupation as influential ($P < 0.05$). In Nigeria's Ahmadu Bello University, Zaria, Kaduna State Teaching Hospital, Adeniran et al. (2022) considered a multi-server single-channel queueing model and obtained a utilization factor of 13%. From the questionnaires that were distributed, they also discovered dissatisfaction with the service quality in the hospital, where the number of patients outnumbered the hospital staff. Such pressure on a hospital staff forces them to dispose patients without thorough treatment. Still, a multi-server exponential queueing system was adopted by Suleiman et al. (2022) in northwestern hospitals in Nigeria; it showed the busiest and least-busy hospitals having utilization factors of 89 and 10%, respectively. In terms of waiting times, the average times that were spent in queues by patients at the busiest and least-busy hospitals were 0.35 and 0.29 minutes, respectively. On the other hand, Rema and Sikdar (2021) used a Monte Carlo simulation technique to analyze queueing patterns and study patient flow in order to manage queues and minimize delays. Le et al. (2021) sought to improve patient wait times due to overcrowding in the emergency departments (EDs) at public hospitals in Vietnam. Using a lean approach, the reductions in delay and waiting times were 33% for patients who required operations (from 134.4 to 89.4 minutes).

With attempts to reduce patient wait times, Fun et al. (2022) used discrete event simulation (DES) in a Malaysian public hospital. The model was used to evaluate the effects of changing consultation start times and patient arrival times. As per the findings, matching consultation start times and patient arrival times indicated the potential of reducing both waiting times and crowding by 40%, with the number of patients waiting per hour possibly reduced by 10–21% during peak hours. Mohammadi et al. (2022) planned the optimal use of resources and the improvement of service quality to estimate the average length of a stay (LOS), bed occupancy rate (BOR), bed blocking probability (BBP), and the throughput of patients in a cardiac surgery department (CSD) through simulation models at Farshchian Hospital, Hamadan, Iran; they used post-operative ward (POW) and intensive care unit (ICU) patients and beds as servers. With a combination of a Monte Carlo simulation and the use of Python software, the queueing simulation results indicated that, for a fixed number of beds in an ICU, the BOR in a POW decreased as the number of beds in the POW increased, and the LOS in an ICU increased with a decrease in the number of beds in a POW. According to the simulation, the results indicated the problem to be poor queuing-system management rather than insufficient resources; the possibility of reducing the overall average waiting times in the department during business hours was shown (from 37.24 to 29.22 minutes).

On the other hand, Proudlove (2022) applied queueing theory in Greater Manchester, United Kingdom (UK), at a National Health Service (NHS) acute hospital trust for resizing its pediatric inpatient department. The hospital sought to reduce the number of bed occupancies from 54 to 85% using a Monte Carlo simulation. Through the basic application of queueing theory, the recommendation was not to reduce the bed occupancy level, as it was revealed that using a bed occupancy target of 85% would result in a risk of 33% that all beds would be full and that using a very low risk of all beds being full of 0.1% would result in an average bed occupancy of 55%. Other researchers like Palomo et al. (2023) provided insights on queueing theoretic methods when analyzing the time evolution of patients who were hospitalized due to the coronavirus (COVID). The number of COVID patients was modeled as a dynamical system based on the theory of infinite server queues with time inhomogeneous Poisson arrival rates. One of the key findings was the existence of a lag between the time of the peak arrival rates of infected patients not coinciding with the times of the peak numbers of hospitalized patients (or deaths). In addition, Kalwar et al. (2021) analyzed the contribution of queueing theory and discrete event simulation in the improvement of healthcare and noted the queueing system mismanagement of resources and the queuing system to be among the main reason for the low quality of the healthcare service delivery in public sector hospitals of Pakistan. Medical service reports indicated the problems that were faced by patients to be delayed service, long waiting times, and less departmental capacity (at emergency, out-patient departments [OPDs], and laboratories) as well as the inadequate number of doctors. The authors were therefore convinced that applications of queueing theory could largely contribute in the healthcare system.

Concerning predictions, Joseph et al. (2022) ascertained the importance and usefulness of predicting patient waiting times in reducing uncertainty regarding wait times. Their study used machine learning to predict patient waiting times before consultations and throughput times in the outpatient clinics. The predictions made use of four models (random forest, XGBoost, random forest with SMOTE, and XGBoost with SMOTE) regarding gender, the day of a visit, the month of a visit, the time of a visit, a consultation's start time, vital examinations, laboratory visits, pharmacy visits, repeated arrivals, consultation sessions, and weather conditions as some of the input variables. In terms of feature importance, the time of a visit was considered to be the major predictor in determining the throughput time for all of the models, with high feature importance scores of 0.396, 0.321, and 0.266 for random forest, RF with SMOTE, and XGB, respectively. The waiting time before consultation was predicted with an accuracy of 0.86, and the throughput time accuracy was 0.93. The areas under the curve (AUC) for the best models that predicted waiting times before consultations were 0.85 and 0.82 for XGBoost and XGBoost with SMOTE, respectively. The AUC that was obtained for the highest performing XGBoost model for predicting the throughput time was 0.89, and the random forest model received an AUC of 0.87.

In a recent study, Benevento et al. (2023) tested various machine-learning techniques using predictive analytics that were applied to two large data sets from actual emergency departments (EDs). They evaluated the predictive ability of Lasso, random forest, support vector regression, an artificial neural network, and the ensemble method using different error metrics and computational times. For prediction accuracy improvement, new queue-based variables that captured the current states of the EDs were defined as additional predictors. The results indicated the ensemble method to be the most effective at predicting waiting times. Concerning the accuracy and computational efficiency, random forest was a reasonable trade-off. In addition, Walker et al. (2022) sought to validate machine-learning models to predict patient waiting times in various emergency departments. They discovered the best-performing models to be random forest, and linear regression models performed the best in waiting-time predictions. The important variables were the triage category, last$-k$ patient average wait times, and arrival times.

2. METHODS

The study was conducted in Limbe (in the Southwest Region of Cameroon) at one of the local hospitals (Regional Hospital Limbe [RHL], which also serves as the principal referral hospital in the region). For this reason, the hospital is busy; this results in high patient traffic. It is a public government hospital with a capacity of 200 beds and offers a wide range of services, like an out-patient department (OPD), a laboratory, and a pharmacy; these units include general and specialist units (like cardiologists, among others).

The instruments that were used in the data-collection process were questionnaires and observations; in addition, registers were used with the required information from some selected departments. Questionnaires (in French and English) were distributed to both patients and medical personnel in order to obtain opinions and general information on the functioning of the hospital. The survey was carried out over 28 days using the convenience sampling method, with a patient response rate of 84% (a total of 170 respondents and 32 refusal cases). The medical personnel had a 100% response rate, with 26 respondents and no refusal cases.

To simplify the work, the out-patient department (OPD) that was mentioned in the questionnaire was composed of four service segments:

1) almoner (cash payments);
2) triage (parameter observations);
3) screening room (registration of patient details, complaints, and taking of vital parameters);
4) consultation (both general practitioner [GP] and specialist doctors).

To account for different perceptions and to avoid exaggeration and biases, the study had to be complemented by observations. These were performed upon the arrival of the patients in the respective departments for their various services. A stopwatch was used to time and record important variables such as service times, waiting times, and patient arrival rates. In total, the number of days for the observations amounted to 43. Initially, three departments were to be observed (namely, OPD, laboratory, and imaging center); however, the imaging center was omitted due to its complexity. The OPD had four service segments, but only two were selected: the almoner, and the consultation (GPs and specialists). Of the specialists, only five were observed: the internist (only one server), the ear nose throat doctors (ENT) (two servers alternating on different days), and the cardiologist (two servers working in parallel). The reason for not considering the other specialists was that they either had a finite population that was not part of the study or they rarely experienced queues.

For the data-analysis process, the questionnaire and observation data was entered into IBM SPSS software (Statistical Package for Social Sciences – Version 20) and analyzed using Python. The analysis was in terms of general descriptions and visualizations as well as the performances of the correlation tests on some of the observed variables.

The research implemented two methodologies: machine-learning concepts (for the predictions), and queueing theory. From the different machine-learning algorithms that were used for the predictions, this study focused on random forest only because of its flexibility and ease of use; it is widely used among the other related algorithms because of its simplicity and use for both regression and classification. Random forest is a supervised machine-learning algorithm that creates a forest and makes it random; it works by building and merging several decision trees in order to obtain more-accurate and stable forecasts (Abdulkareem & Abdulazeez, 2021). It adds randomness to a model while the trees grow, and it searches for the best feature among a random subset of features instead of searching for the most important feature while splitting a node; as a result, there is a wide variety that usually leads to a better model. Therefore, the algorithm only takes a random subset of characteristics into account to split a node in the random forest; one can even randomize trees by using random thresholds for each function instead of looking for the best possible thresholds (like a normal decision tree). Although random forest has its advantages and disadvantages, one of the important points is its ability to provide a fairly good indicator of the importance that it assigns to features.

The predictions for both the waiting and the total time spent were made using the Jupyter Notebook in Python. The main objective was to check the possibility of predicting the waiting and total times that a patient spends in the hospital on any other day and to determine how accurate this model was. The problem represented a supervised regression machine-learning problem, as both the hospital data and the times to be predicted were available (and also real values). This type of machine learning requires a lot of data, and the model can be trained as well.

The observed data was divided into two-thirds training data and one-third testing data while fitting the model using the random forest regression. Initially, the model had 1000 decision trees, which were further trimmed for simplicity to only 2 levels with 10 decision trees. The study used only two measures to assess prediction accuracy – the mean absolute error (MAE) and the mean absolute percentage error (MAPE). Of these, MAE was considered the more favorable metric for describing the average error and the expected magnitude of errors in the forecast.

Another important step was to use the importance of the characteristics to find the most important or relevant variables to predict these times. In terms of the queueing theory application, we illustrated the queueing system and its various components. The formation of the queues had two important properties: the maximum size (the capacity of a system)/queueing capacity, and the queueing discipline. The size was the population that could either be limited (finite) or unlimited (infinite). The queueing rules for selecting the patients for the services were called the queue discipline, and they were classified as First In/First Out (FIFO), Last In/First Out (LIFO), Priority, and so on. For simple queues, some queueing theory formulas exist, while for complex situations (like this study), computer simulations are needed.

From the two basic approaches that are available for analyzing queue systems, analytical and simulation approaches exist. The former attempts to find formulas (some algorithms) for calculating the steady-state performance measures of a system, while the latter has two types: simulating a known distribution, and simulating a non-specified distribution (bootstrapping). Thus, the simulation method has more flexibility when compared to the analytical approach, which simulates random elements of a system at the same time as keeping track of events as they occur over time. It is based on a simulation model, which is a computer model that mimics a situation in real life. This model is similar to other mathematical models, but it incorporates uncertainty in one or more input variables. The benefit of a computer simulation is the ability to answer what-if questions without actually changing the physical system.

In queueing theory, a queueing concept model known as 'Kendall's notation' exists in its simplest form (A/B/C), which was formalized by Kendall (1953); he was an English mathematician and statistician who was known for contributing to areas like probability, statistical shape analysis, and queueing theory. His notation is used in the description of his queue system; an explanation is given in Figure 2.

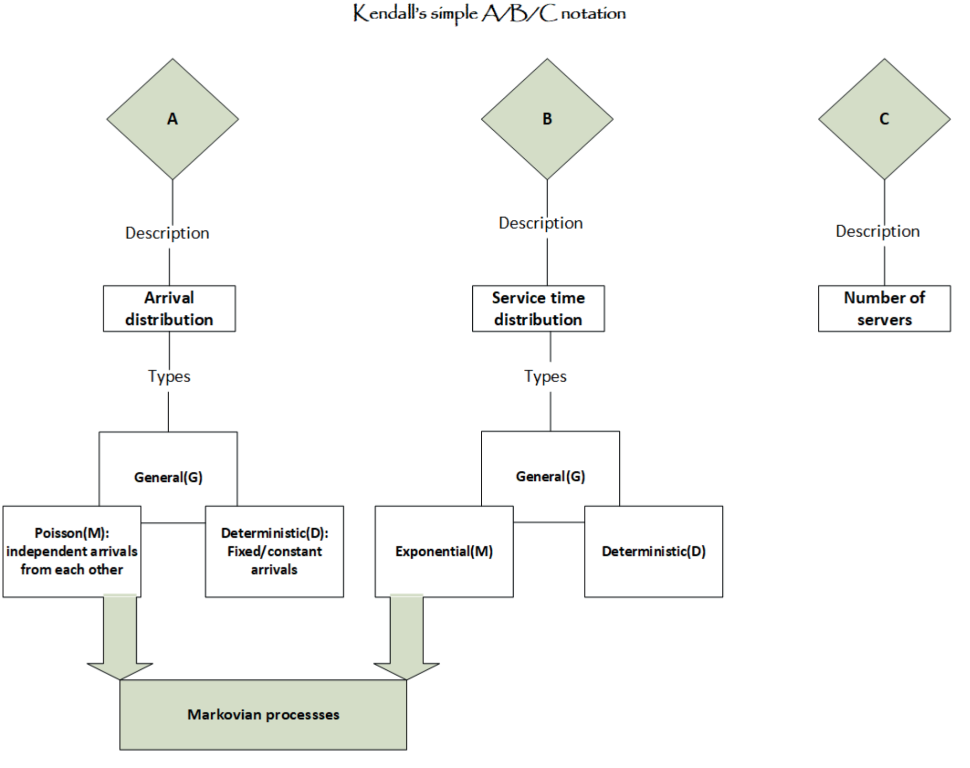


Fig. 2. *Kendall notation*

An extension of 'Kendall's notation' is given by A/B/C/D/E, where D is the queueing capacity and E is the queueing discipline. By the basic assumptions of queueing theory, the following apply using 'Kendall's notation':

A: It is assumed to follow a Poisson distribution and is a Markovian process.

B: It is assumed to follow an exponential distribution and is a Markovian process.

C: It is a positive integer.

D: Is assumed to be ∞, implying that no one is turned away when they come for service.

E: It is assumed to be FIFO/FCFS.

The basic model notation is, thus, given as M/M/C/∞/FIFO/FCFS, where 'M' is the notation for the Markovian processes. The Markov process is a random process in which, given the present, the future is independent of the past and assumes an arrival or service rate. An exponential distribution exhibits an important property of being memoryless; that is, the time for the next arrival is independent of when the last arrival occurred. Its characteristics include an equal mean and standard deviation.

If these assumptions are met, then the analytical approach can be used that makes use of Little's formulas (which are important formulas in queueing theory).

The notation for a Poisson distribution is given as $X \sim \mathrm{Po}(\lambda)$, where λ is the only parameter and signifies the average arrival rate of a set of patients. On the other hand, the notation for an exponential distribution is given as $X \sim \mathrm{Exp}(\mu)$, where μ is the only parameter (and is the average service rate of the patients).

For many queueing situations, arrivals occur randomly, so the occurrence of the next arrival cannot be predicted. Among the distributions that represent the times between successive arrivals, the most important is exponential distribution.

In terms of hospital settings, most arrivals are modeled by a Poisson distribution (where the patients arrive independently one after another) and an exponential service distribution. An example of deterministic arrivals in a hospital setup is when the consultations are by appointment and the patients must arrive at given fixed times. A general distribution is nonspecific and could be any kind of distribution.

We begin by illustrating a simple queue (shown in Figure 3), where queuing behaviors such as reneging and baulking are exhibited.

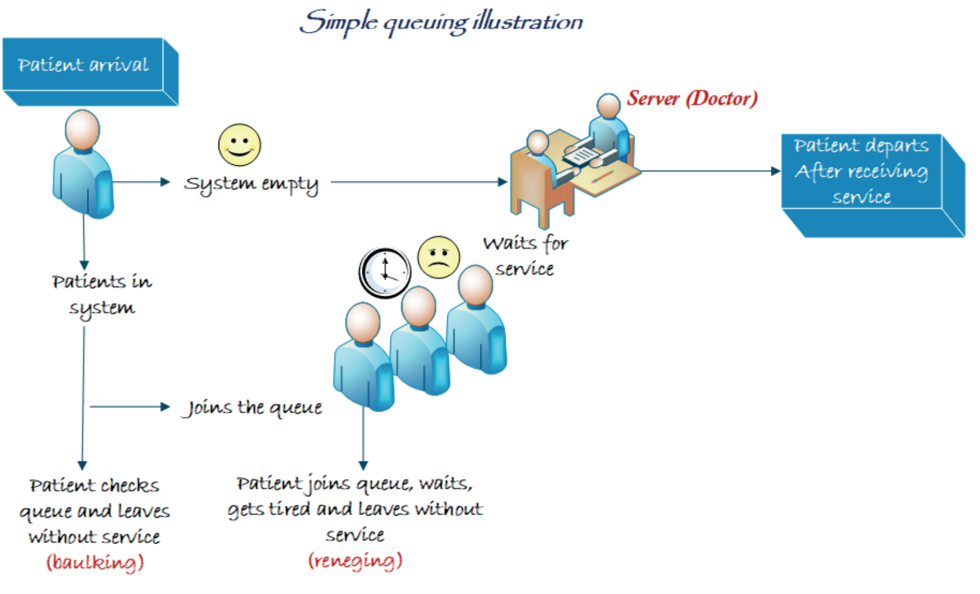


Fig. 3. *Illustration of simple queue*

The illustration simply shows the level of patience from when a patient enters a system up to when he/she is either served or decides to give up and leave. Thus, the queues may be organized in different ways, and they have various types (shown in Figure 4), where only three common types have been illustrated; in this case, the term server refers to a doctor.

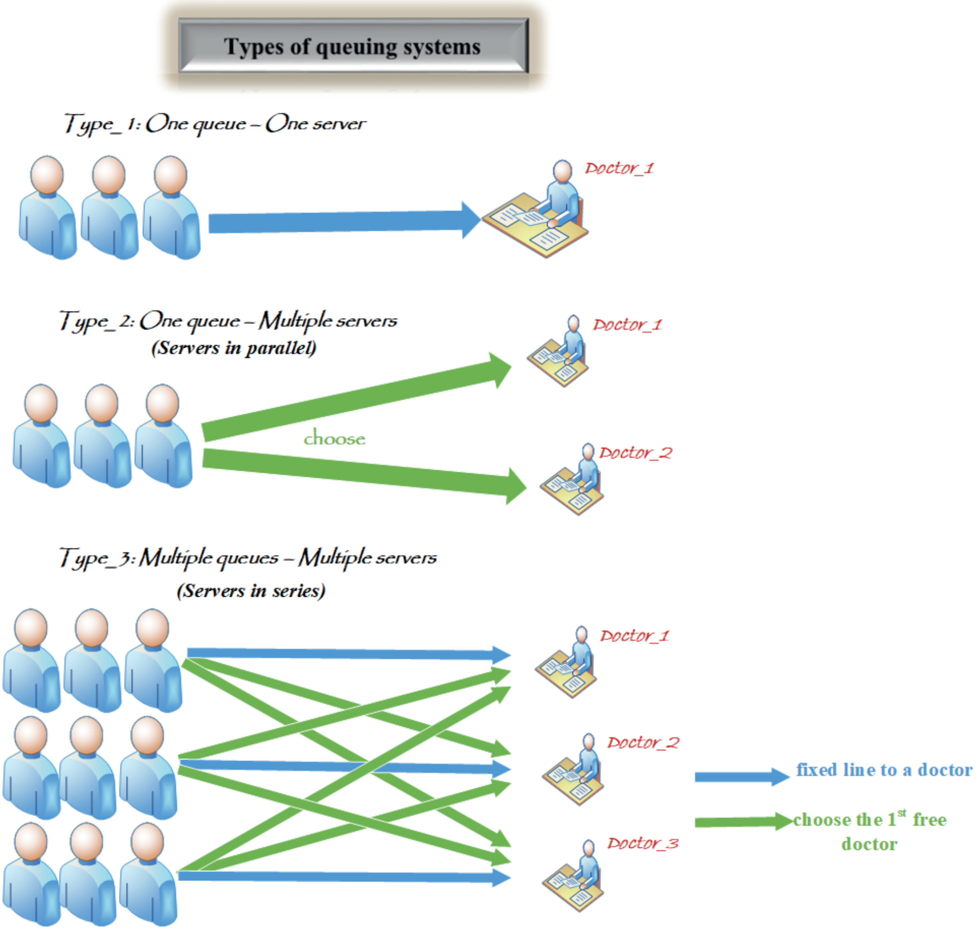


Fig. 4. *Queue types*

In some systems, patients have the option of seeing the first free doctor or simply maintaining the queue to the fixed doctor to whom they have been assigned. The patients are then selected for service by the various disciplines in the queue. In our case, the only department that was chosen was the consultation by a GP; this was mainly because of the time constraint and a lack of personnel. The system belongs to *Type_3* of Figure 4, with three consultation rooms (implying three servers) with patients that maintain a fixed queue. At times, a free doctor can select patients from another queue (which makes the situation complex). Moreover, the patients were from an infinite population, meaning that the system was able to receive all of the patients who came for consultation. The queueing discipline was FIFO, although priority and emergency cases were treated first at times, making the system preemptive.

During the observations, the system was noted to be more complex than was anticipated, and the queueing theory assumptions were violated. For example, the system service time was supposed to exhibit exponential distribution properties by the assumption of healthcare waiting lines; however, it did not since there was no continuity in serving the patients. The arrival process was not constant over time either, and it could not fit the Poisson process; hence, the arrivals and service times were instead categorized as a general distribution (meaning that it could be any kind of a distribution). As a result, the assumption of modeling it as M | M | 1 and using an analytical approach was, therefore, ruled out. The system was observed to be a $3 \times$ G| G | 1 | ∞ | FIFO (or priority), implying a complex system that required a simulation approach.

From the observations, the three observed doctors (named '*Doctor_ 1*,' '*Doctor_ 2*,' and '*Doctor_ 3*') observed 19, 12, and 13 patients, respectively; the patients arrived during the period of 7:30–12:40. The most important parameters that were taken note of were the arrival times, waiting times, service times per hour, and numbers of servers.

Since the system was a bit complex, only the simulation approach on this system (to find the averages) was possible (instead of using the standard queueing theory, with its associated mathematical formulas). The type of simulation that was used was bootstrap simulation, as the service time was not from a specified distribution. The tool that was used for this analysis was Excel, where the observations were used to perform a simulation. The initial step for this approach involved reforming the probability distribution based on random-number generation in reference to the real data, where the random numbers were generated by the *RAND*( ) function. One simulation was performed per doctor for each of the three using the *VLOOKUP*( ) function to simulate 1000 replications, where each replication was an independent replay of the occurring events. The replications were generated using a data table; to do this, the observed data in the spreadsheet was used to construct a typical "prototype" of the simulation. Also, a 95% confidence interval (CI) was estimated. Part of the replication process table that was created for '*Doctor_ 1*' is shown in Figure 5.

The simulation analysis was also useful for investigating what might have happened if a different policy or strategy had been used. After the observations, the simulation model was used to compare two situations: the first was an observed system "with disruptions" (where the doctors had other things to attend to during their consultations), and the second was termed a system "without disruptions" (for comparison, where the doctors only consulted without any other disturbances).

One of the virtues of the simulation was that it allowed us to experiment with alternatives. Although not used in this study, other alternatives could include simulating a system in which the number of doctors is increased to see the difference that can be made.

No. Replications	Average time in system	Average time in queue	% time server is idle
	22.4	13.72	0.1033058
1	15.32	8.72	0.5657895
2	8.8	2.52	0.6884921
3	8.68	2.08	0.4728435
4	9.52	2.16	0.5523114
5	12.8	5.92	0.4723926
995	14.24	6.76	0.5630841
996	11.2	3.84	0.54
997	18.12	10.4	0.4373178
998	11.12	4.72	0.638009
999	8.84	2.08	0.5517241
1000	9.88	2.16	0.6553571

Fig. 5. *Replication process for 'Doctor_ 1' (supporting software tool output)*

After meticulous data training and simulations, the results are stated in the next section.

3. RESULTS

The data collected for the analysis was both quantitative and qualitative; that is, arrival, waiting, service, and total waiting times in addition to feedback on busy departments and busy days, respectively.

After gathering the opinions from the respondents, making observations, and using the hospital records (registers), the results indicate the waiting times, busy departments, busy days and times, and possible causes of any long queues (where the term 'busy' was associated with a high number of patient arrivals and long queues). The perceptions that were provided by both the patients and the medical personnel helped confirm the findings. All analyses of the patients' movements were performed on an hourly basis, with the times being recorded in minutes.

The arrival rate in the selected departments is presented in Figure 6. Observing the trend, most patients arrived in the morning, and the numbers decreased as the day went by (although the emergency cases tended to increase after hours). Department-wise, the payment section experienced the highest number of arrivals – most likely because it was the starting point. These patients were then distributed to the various segments of the consultation service. Among the specialists, the internist seemed to be burdened the most.

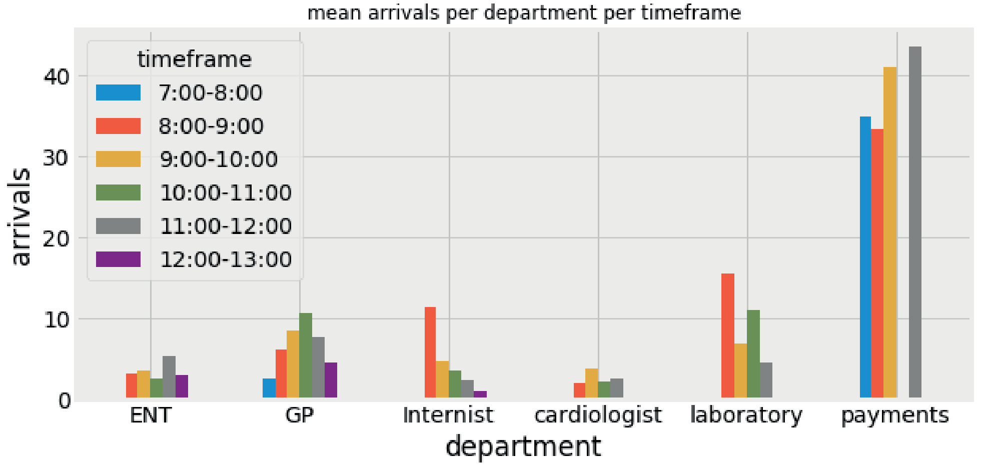


Fig. 6. *Department hourly arrivals*

Figure 7 shows the amount of time that a patient waited for service in the selected departments, given that they found zero patients (no one in the system) up to more than six people. This factor alone triggered waiting times, as other incoming patients were also affected. Normally, if no one is in the system upon arrival, the patient who is first in the queue did not incur any waiting time, with expectations of being served immediately; therefore, this was a concern.

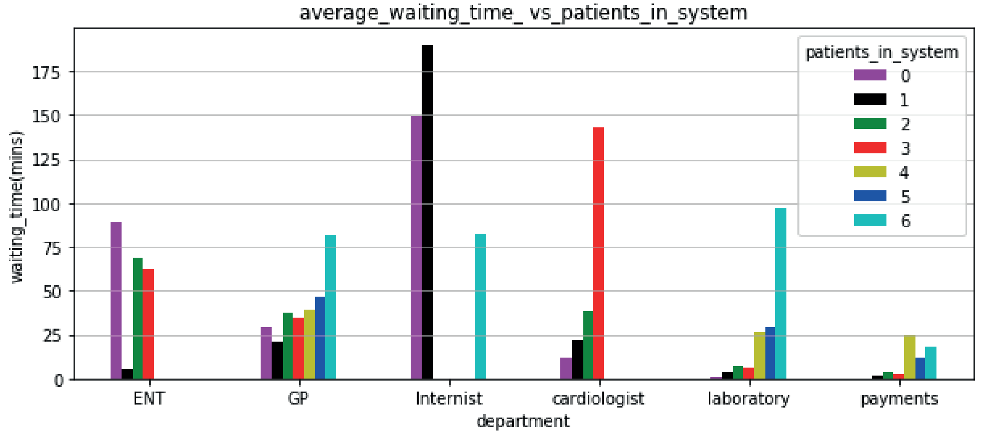


Fig. 7. *Department waiting time when patients are already in system*

The averages for the waiting, service, and total times that were spent in the hospital by the patients in the selected service segments were calculated; of all of the services, the internist had the highest times (illustrated in Figure 8).

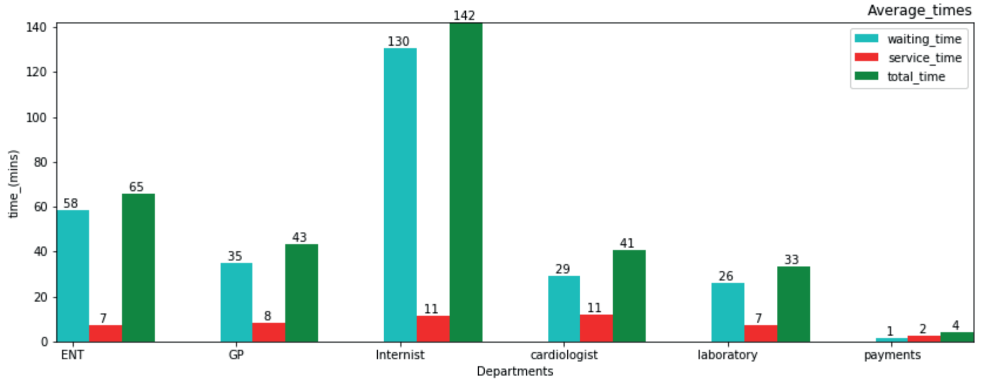


Fig. 8. *Observed times in hospital*

In terms of the laboratory department, the waiting times were in terms of turn-around times (TAT). The averages were calculated by using a sample of five patients for each laboratory test. TAT is defined as the total time that it takes from when a sample is submitted to when the results come out. For the research, only 6 common tests were selected (shown in Table 1); it can be seen that none of the tests were below 56, likely explaining why the patients waited so long.

Table 1. *TATs for RHL lab results*

Laboratory test	TAT (time)
Mp (malaria parasite)	0:56
Hb (hemoglobin)	1:09
FBC (full blood count)	1:45
MS	2:19
FBS (glycaemia)	2:34
Stool analysis	3:50

For correlation clarity, Figure 9 was constructed to check the existence of a relationship among the variables. Of all of the variables, only waiting times and total time spent show a significant relationship (with a strong positive correlation of 0.99).

Concerning the predictions on the waiting and total times that were to be spent by a patient in the hospital tomorrow (referring to the future), a random forest tree was used (where the output of the tree was random each time the notebook was run). Also, the variables that were important for predicting both parameters were calculated and constructed according to their importance. For the prediction of waiting times, the mean absolute error (MAE) was 10.53 minutes, with a standard deviation of 3.54 minutes. The MAE value implied that the model predicted an average of 10.53 minutes of error. To show how accurately the trained model fit the data set, the R-squared value was calculated and found to be 0.64; this showed that the model prediction was almost close to the actual data set.

Fig. 9. *Correlation of variables*

Figure 10 shows that, to predict the waiting times, the most important variables were the times of arrival and the numbers of patients that one would find in the system.

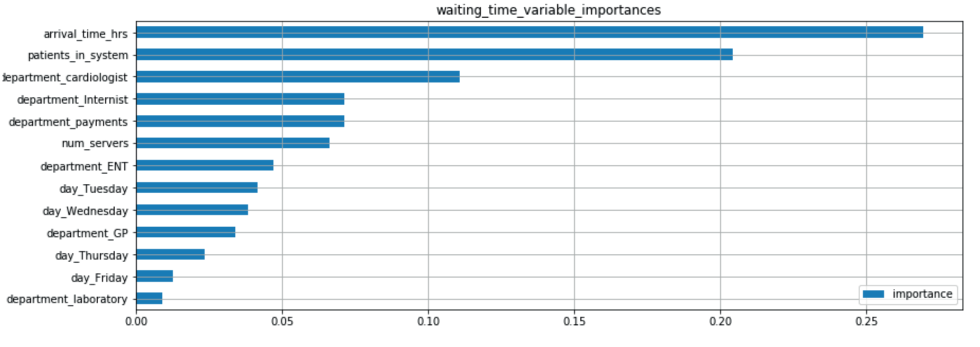


Fig. 10. *Important variables for predicting waiting times*

The trimmed waiting-time tree for prediction is given in Figure 11. Given the variables on the tree, a patient was able to predict how long they could expect to wait in the hospital before receiving service.

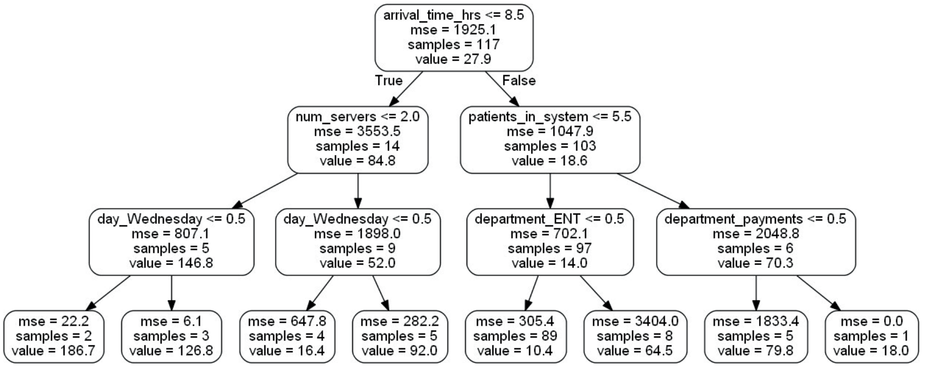


Fig. 11. *Waiting time predictions*

To interpret this random forest in particular, if a patient arrived before 8:50 a.m., they chose the indicated true arrow from the first node; otherwise, they chose the false arrow. If the first node was true, the next question encountered was the number of servers. We moved down the forest answering questions up to the last row, whose nodes indicated the predicted waiting times (which were given as 'value' in minutes). Consider the following example: if a patient arrived before 8:50 a.m. and the service had fewer than two servers on a Wednesday, the waiting time prediction was 186.7 minutes. The 'samples' that were shown on each node were indications of the numbers of samples that were randomly taken for sampling data points, and 'MSE' was the mean squared error.

For the total time prediction, predicting total time that is spent by a patient is somewhat infeasible in reality, as the outcome cannot be predicted without a patient going through the process of waiting and receiving the service first. Therefore, this prediction served as a verifier of the methods rather than a predictor of information beforehand. After training the model, however, the MAE was found to be 1.75 minutes; this implied that the model predicted 1.75 minutes of error on average. The median absolute error (MedAE) was 0.77 minutes, and the mean absolute percentage error (MAPE) was 7.45%.

As highlighted in Figure 12, the most important variables for the total time prediction were the waiting and service times.

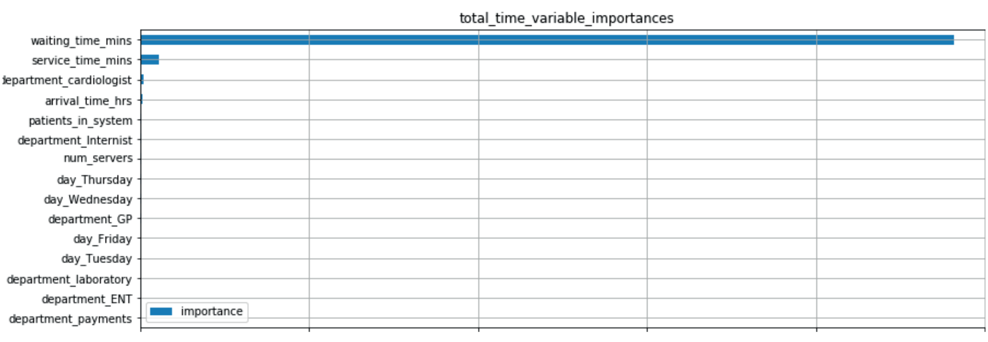


Fig. 12. *Important variables for predicting total time*

Removing all of the other variables and leaving the two important ones, MAE was reduced to 1.4 minutes, while MedAE decreased to 0.39 minutes and MAPE decreased to 5.23%. This implied that all of the other features were not important for the total time prediction after all. Scoring the model on the training data, the R-squared value was 0.99 minutes; this showed that the model prediction was very close to the actual data set. An illustration tree for the total time is shown in Figure 13.

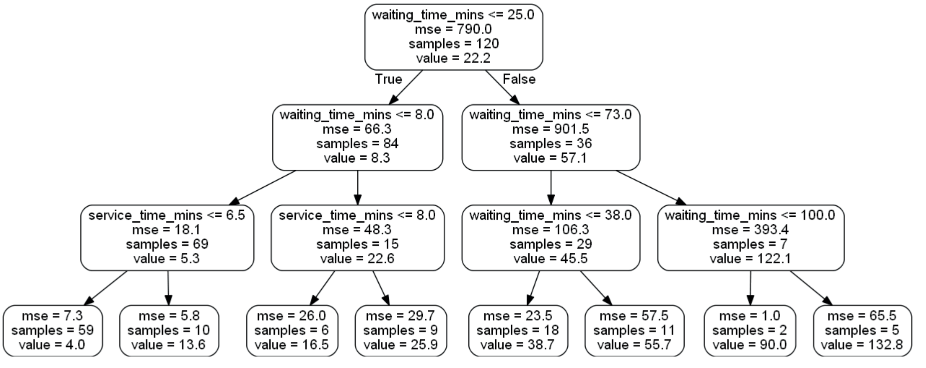


Fig. 13. *Total time predictions*

In order to see the samples from the validation data set and the range within which their errors lay, error plots were constructed for both the waiting and total times. It can be noted from the plots that most of the samples had small minute errors (as seen in Figure 14).

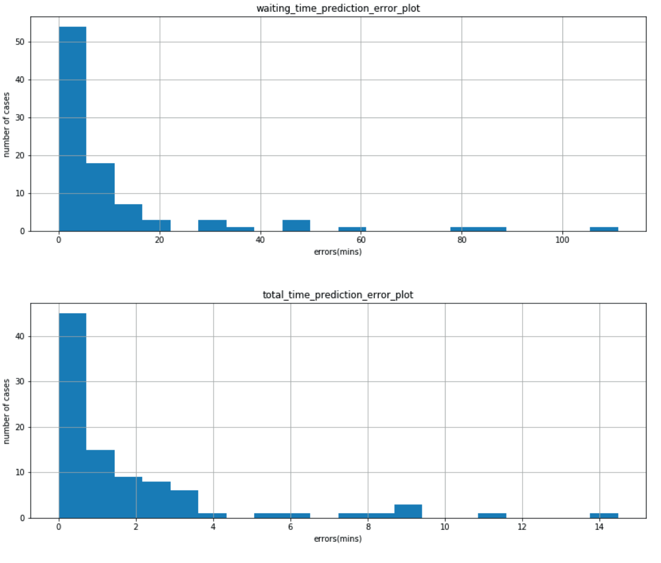


Fig. 14. *Prediction error plots*

For the queueing theory part, each of the three doctors had a summary of the results after a simulation of 1000 replications. The summary measures of performance for the simulations that were calculated included the average time in the system, the average time in a queue, and the percentage of time that a server was idle. The time in the system was the total time that a patient spent from his/her arrival waiting in a queue up to the time that he/she received service (which was simply the waiting time before the service). The time that a server was idle was when a doctor was not busy.

From the results of Figures 15 and 16, 'queueing results with disruptions' was the observed system where the doctors' arrival times were inconsistent and they engaged in other work other than serving the patients in queues. On the other hand, 'queueing results with no disruptions' was the model that it was compared to if the operating strategy was different, with doctors only serving patients without attending to other commitments.

	Average time in system (mins)	Average time in queue (mins)	% time server is idle	Sample Standard deviation for average time in system	Sample Standard deviation for average time in queue	95% Confidence interval for time in system		95% Confidence interval for time in queue	
				54.37	50.64	-40.14	177.34	0.00	156.89
MIN	7.48	1.92	0%						
AVERAGE	68.60	55.61	22.76%	**Doctor_1 queueing results with disruptions**					
MAXIMUM	378.08	350.08	71.38%						

Fig. 15. *Model for Doctor 1 with disruptions (supporting software tool output)*

	Average time in system (mins)	Average time in queue (mins)	% time server is idle	Sample Standard deviation for average time in system	Sample Standard deviation for average time in queue	95% Confidence interval for time in system		95% Confidence interval for time in queue	
				3.35	2.98	4.28	17.67	0.00	10.00
MIN	5.76	0.16	9.69%						
AVERAGE	10.98	4.03	54.97%	**Doctor_1 queueing results with no disruptions**					
MAXIMUM	37.76	29.28	78.89%						

Fig. 16. *Model for Doctor 1 with no disruptions (supporting software tool output)*

When comparing the 2 Systems for each doctor, it could be observed that there was a huge difference – especially in the times that a patient spent in both a queue and the system. For example, a patient spent 55 minutes on average queueing in Figure 15, whereas they only spent 4 minutes in Figure 16 because the doctor had no other duties other than consulting. The same concept was applied to the other two doctors, with any differences noted in the systems.

CI play a role in the modeling and simulation, as they were used in the model validation. The wide CI that was outputted in Figure 15 indicated the small sample size that was used; in our case, standard deviation (σ) told us how the collected time was spread out from the average.

Since random numbers were generated, the output was also random; this meant that running the spread-sheet again would have given slightly different results from what could be seen. It is to be noted, however, that a simulation is an approximate method and might not give exact answers; hence, the reason why the 'with no disruptions' comparison system might have seemed to be close to perfect. In the following section, we give an in-depth discussion of the results as well as our conclusions.

4. DISCUSSION AND CONCLUSIONS

After performing the analysis, we drew conclusions on the findings and gave recommendations by providing possible solutions. The results and findings of this research made use of registers where Mondays were normal days; therefore, the study assumed the nonexistence of a ghost town. Given the case that a ghost town happened and patient arrivals declined on Mondays, however, then these findings were not applicable and were subject to change only in terms of the busy days and the number of arrivals.

From the hospital registers, it was generally noted that the number of patient arrivals had been reducing over the past years; this was likely because of the crisis or standards being lowering without being noticed.

After observing and analyzing the possible causes of the queues that result in long waiting times, we found that patient waiting times were sensitive to doctors' arrival times and the time that the doctors spent on other activities. This could be supported by the waiting times that were noted for all of the patients who arrived and found zero people in the system. The primary and secondary data played important roles in our attempt to answer the research questions.

From the combined patient and medical responses, the three main departments were OPD, the laboratory, and the imaging center. As much as these results were obtained from the perceptions of 196 people, we concluded them to be true without further investigations. All of the other observations that were necessary for the research were then carried out in two of these departments.

According to the perception of the questionnaire, the top-three busy days were found to be Monday, Wednesday, and Tuesday; at the same time, our observations indicated that Monday, Wednesday, and Tuesday had the highest numbers of patient arrivals. In combination, this is evidence enough to conclude that Monday, Wednesday, and Tuesday are the top-three busiest days of the week for the hospital. As was stated from the other studies in the literature review, we can also compare our results with Biya et al. (2022) (who found Monday to be the busiest day of the week). Again, when analyzing the times that were spent in the hospital, Wednesday and Tuesday had the highest amounts of time that were spent by the patients. Here, Monday could not be observed because of the crisis; from our secondary data findings, however, we could assume that it featured the highest total time that was spent in the hospital.

We also wanted to find out if there could be a correlation that existed between the waiting lines and the day of the week (or even the time of day). From our assessment, there was an implication that waiting lines and the time that was spent were related to the day of the week. In simple form, we could put it in an equation as follows:

$$\uparrow \text{in arrivals} \Longrightarrow \text{long queues} = \text{busy day} \Longrightarrow \text{more time spent in hospital} \tag{1}$$

In terms of busy times of the day, most of the medical personnel responded to mornings; again, this matched with our observations, where the hospital was discovered to be more congested and busy during the morning hours.

In response to the predictions that were made (and using the developed models), a patient can predict the amount of time that they can expect to wait the next day in the hospital. Figure 12 complements Figure 9 regarding the fact that the waiting time is the most important variable for total time prediction; to support this, the two were portrayed to have a strong correlation.

In queueing theory, we based our conclusions on the data analysis (including the simulation and assumptions), as the data was not enough to draw strong scientific conclusions. This implied that the value of the essay was not in its actual conclusions but rather in the methodology that was used. Using a what-if analysis, an insight can be gained into the models in understanding and implementing waiting-time strategies. We can only explain what can be done to get a more reliable appreciation of the current situation vs. different possible working policies (which are indicated in the recommendation section). The summarized results are presented in Table 2.

Table 2. *The findings*

Research questions	Subjective findings
Main causes of long waiting times	Doctors' arrival times and commitments to other activities
Departments with long waiting times	OPD, laboratory, and imaging center
Busy days	Monday, Wednesday, and Tuesday
Busy times	Morning

Based on our work, an attempt to improve hospital performance, and the general operation, we suggest the following for each department:

1. OPD Department
   - Change the queueing system: use different queueing rules, like adopting "*Type* 2: *One queue- Multiple servers*" queueing from Figure 4 (or any other combination), where a patient is assigned to the first free/available doctor.
   - Increase the number of doctors per shift: the queuing system setup could be changed by having an increase of one or two doctors consulting (especially on the highlighted busy days) to reduce waiting times.
   - Morning meetings: the doctors on duty could either send a representative to any meetings, or they could be shifted to take place at a less-busy times (e.g., in the afternoon).
   - Ward rounds: the doctors on duty at OPD could be exempted from ward rounds while non-consulting doctors do rounds and attend to emergency cases.
   - Classify services: a class of patients that require shorter service times (like those coming for medical certificates) could be dedicated to a different server. Waiting an hour for a signature and a stamp that takes two minutes is not ideal.

  - For specialist consultations, appointment times could be introduced:
    - *Triage*: the process from triage to the screening room is more or less repetitive; there is a need to either combine or cut off some unnecessary steps and procedures to reduce the waiting times.
    - *Almoner (Payment section)*: the system is not computerized, and a lot of time is spent invoicing by hand; if this system were changed, many differences could be noted.
2. Laboratory Department
   From the findings concerning the long waiting times in the laboratory, this was due to the fact that results take a long time to come out (and/or handed to the patient) than the time it takes to collect samples. As a result, the solution can only lie within reviewing the way the results are grouped or timed. The results should be produced in two shifts (i.e., 11 a.m. and 2 p.m.).
3. Imaging center
   This was among the departments that were mentioned to have long waiting times; however, no observations could be made because of the complex setup and the sensitivity of the process. We cannot make any conclusions but to recommend a study of this for future work.

In general, the hospital can take the following points into consideration:

1. Staff allocation: in general, the staff allocations could be more concentrated for all of the service segments to suit the busier days and times.
2. Entertainment: providing patients with educational health talks, books, magazines, etc. could also help keep them occupied while they are in queues. This can also help avoid reneging and baulking, as most patients become impatient and bored and leave without receiving service.

For future research and use, the type of operational data that is needed as input for a queueing model could be introduced, as it is often unavailable in the registers. This is necessary for reviewing system performance as well.

Acknowledgements

Funding for the project was provided by The African Institute for Mathematical Sciences (AIMS)-Cameroon and Mastercard Foundation.

REFERENCES

Abdulkareem N.M. & Abdulazeez A.M. (2021). Machine learning classification based on random forest algorithm: a review. *International Journal of Science and Business*, **5**(2), pp. 128–142.

Adeniran D.A., Burodo M.S. & Suleiman S. (2022). Application of qeuing theory and management of waiting time using multiple server model: Empirical evidence from Ahmadu Bello University Teaching Hospital, Zaria, Kaduna State, Nigeria. *International Journal of Scientific and Management Research*, **5**(4), pp. 159–174. **DOI**: https://doi.org/10.37502/IJSMR.2022.5412.

Almusawi M.A., Radwan N.M., Mahmoud N.E., Haji A.M. & Alabdulkareem K.I. (2023). Patient waiting time at primary healthcare centers in Riyadh Region, Saudi Arabia: A cross-sectional study. *Dr. Sulaiman Al Habib Medical Journal*, **5**, pp. 65–69. **DOI**: https://doi.org/10.1007/s44229-023-00032-3.

Benevento E., Aloini D. & Squicciarini N. (2023). Towards a real-time prediction of waiting times in emergency departments: A comparative analysis of machine learning techniques. *International Journal of Forecasting*, **39**(1), pp. 192–208. **DOI**: https://doi.org/10.1016/j.ijforecast.2021.10.006.

Biya M., Gezahagn M., Birhanu B., Yitbarek K., Getachew N. & Beyene W. (2022). Waiting time and its associated factors in patients presenting to outpatient departments at Public Hospitals of Jimma Zone, Southwest Ethiopia. *BMC Health Services Research*, **22**, art. no. 107. **DOI**: https://doi.org/10.1186/s12913-022-07502-8.

Boukas E.-K., Zhang Q. & Yin G. (1995). Robust production and maintenance planning in stochastic manufacturing systems. *IEEE Transactions on Automatic Control*, **40**(6), pp. 1098–1102. **DOI**: https://doi.org/10.1109/9.388692.

Brockmeyer E., Halstrøm H. & Jensen A. (1948). *The life and works of A.K. Erlang*. Copenhagen: Copenhagen Telephone Comp.

Chakka R., Van Do T. & Pandi Z. (2009). A generalized Markovian queue and its applications to performance analysis in telecommunications networks. In: D.D. Kouvatsos (Ed.), *Performance Modelling and Analysis of Heterogeneous Networks*. River Publishers, pp. 371–387. **DOI**: https://doi.org/10.1201/9781003339052-23.

Cooper R.B. (1981). Queueing theory. In: *Proceedings of the ACM'81 conference* (1981), pp. 119–122.

Ding S. (2023). Analysis and application of computer queueing theory. In: P. Kar, J. Li & Y. Qiu (Eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*. Amsterdam: Atlantis Press, pp. 447–453. **DOI**: https://doi.org/10.2991/978-94-6463-300-9_45.

Dshalalow J.H. (Ed.) (1995). *Advances in Queueing Theory, Methods, and Open Problems*. Boca Raton: CRC Press. **DOI**: https://doi.org/10.1201/9781003418283.

Elalouf A. & Wachtel G. (2021). Queueing problems in emergency departments: a review of practical approaches and research methodologies. *Operations Research Forum*, **3**, art. no. 2. **DOI**: https://doi.org/10.1007/s43069-021-00114-8.

Erlang A.K. (1909). The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, **20**, pp. 33–39.

Fun W.H., Tan E.H., Khalid R., Sararaks S., Tang K.F., Ab Rahim I., Md. Sharif S., Jawahir S., Sibert R.M.Y. & Nawawi M.K.M. (2022). Applying discrete event simulation to reduce patient wait times and crowding: The case of a specialist outpatient clinic with dual practice system. *Healthcare*, **10**(2), art. no. 189. **DOI**: https://doi.org/10.3390/healthcare10020189.

Garcia-Corchero J.D. & Jimenez-Rubio D. (2022). Waiting times in healthcare: equal treatment for equal need? *International Journal for Equity in Health*, **21**(1), art. no. 184. **DOI**: https://doi.org/10.1186/s12939-022-01799-x.

Green L. (2013). Queueing analysis in health care. In: R. Hall (Ed.) *Patient Flow: Reducing Delay in Healthcare Delivery*, Second Edition, Springer, pp. 361–384. **DOI**: https://doi.org/10.1007/978-1-4614-9512-3.

Harchol-Balter M. (2013). *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge: Cambridge University Press. **DOI**: https://doi.org/10.1017/CBO9781139226424.

Harki N.A. (2024). Multi-level feedback queue scheduling technique. *Cihan University-Erbil Scientific Journal*, **8**(2), pp. 36–42. **DOI**: https://doi.org/10.24086/cuesj.v8n2y2024.pp36-42.

Ji J. (2023). Application of queuing theory in public service counters. In: P. Kar J., Li & Y. Qiu (Eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Amsterdam: Atlantis Press, pp. 509–517. **DOI**: https://doi.org/10.2991/978-94-6463-300-9_52.

Johnston A., Lang E. & Innes G. (2022). The waiting game: managing flow by applying queuing theory in Canadian emergency departments. *Canadian Journal of Emergency Medicine*, **24**(4), pp. 355–356. **DOI**: https://doi.org/10.1007/s43678-022-00318-9.

Joseph J., Senith S., Kirubaraj A.A. & Ramson J.S. (2022). Machine learning for prediction of wait times in outpatient clinic. *Procedia Computer Science*, **215**, pp. 230–239. **DOI**: https://doi.org/10.1016/j.procs.2022.12.026.

Kalwar M.A., Marri H.B., Khan M.A. & Khaskheli S.A. (2021). Applications of queuing theory and discrete event simulation in health care units of Pakistan. *International Journal of Science and Engineering Investigations*, **10**(109), pp. 6–18. **URL**: https://www.researchgate.net/publication/356980604.

Kendall D.G. (1953). Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain. *The Annals of Mathematical Statistics*, **24**(3), pp. 338–354.

Koole G. (2013). *Call center optimization*. MG Books.

Lakatos L., Szeidl L. & Telek M. (2019). Queuing networks, In: L. Lakatos, L. Szeidl, M. Telek, *Introduction to Queueing Systems with Telecommunication Applications*, Second Edition, Springer, pp. 421–444. **DOI**: https://doi.org/10.1007/978-3-030-15142-3_10.

Le D.X., Do H.T., Bui K.T., Hoang T.Q., Nguyen G.H., Nguyen A.V., Nguyen Q.T., Gorgui-Naguib H. & Naguib RG.N. (2021). Lean management for improving hospital waiting times – Case study of a Vietnamese public/general hospital emergency department. *The International Journal of Health Planning and Management*, **37**(1), pp. 156–170. **DOI**: https://doi.org/10.1002/hpm.3310.

Li X. (2024). Performance evaluation of M | M | 1 queueing models in cloud computing environments. *Highlights in Science, Engineering and Technology*, **85**, pp. 841–848. **DOI**: https://doi.org/10.54097/1xcw8y74.

Liberopoulos G., Papadopoulos C.T., Tan B., Smith J.M. & Gershwin S.B. (Eds.) (2006). *Stochastic modeling of manufacturing systems*. Springer. **DOI**: https://doi.org/10.1007/3-540-29057-5.

Machangara D.T. (2018) *Predictions and Application of Queuing Analysis at Regional Hospital Limbe* [MSc thesis], African Institute for Mathematical Sciences (AIMS), Cameroone, December 2018. **URL**: https://library.nexteinstein.org/wp-content/uploads/2023/01/Daphne-Essay-Template-en-2017-2018.tex_.pdf [1.10.2024].

Mbwogge M., Astbury N., Nkumbe H.E., Bunce C. & Bascaran C. (2022). Waiting time and patient satisfaction in a subspecialty eye hospital using a mobile data collection kit: pre-post quality improvement intervention. *JMIRx Med* **3**(3), e34263. **DOI**: https://doi.org/10.2196/34263.

Mohammadi T., Roshanaei G., Faradmal J., Sadeghifar M., Manafi B. & Mahjub H. (2022). Improving service efficiency and throughput of cardiac surgery patients using Monte Carlo simulation: a queueing setting. *Scientific Reports*, **12**(1), art. no. 21217. **DOI**: https://doi.org/10.1038/s41598-022-25689-y.

Nair A.M., Sreelatha K. & Ushakumari P. (2021). Application of queuing theory to a railway ticket window. In: *2021 International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Noida, India, pp. 154–158. **DOI**: https://doi.org/10.1109/ICIPTM52218.2021.9388368.

Newell G.F. (1995). *Applications of Queueing Theory.* Second Edition. London – New York: Chapman an Hall. **DOI**: https://doi.org/10.1007/978-94-009-5970-5.

O'Malley M.S., Fletcher S.W., Fletcher R.H. & Earp J.A. (1983). Measuring patient waiting time in a practice setting: A comparison of methods. *The Journal of Ambulatory Care Management*, **6**(3), pp. 20–27. **DOI**: https://doi.org/10.1097/00004479-198308000-00006.

Palm C. (1943). Intensity fluctuations in telephone traffic. *Ericsson Technics*, vol. **44**.

Palomo S., Pender J.J., Massey W.A. & Hampshire R.C. (2023). Flattening the curve: Insights from queueing theory. *PLoS ONE*, **18**(6), e0286501. **DOI**: https://doi.org/10.1371/journal.pone.0286501.

Peter P.O. & Sivasamy R. (2021). Queueing theory techniques and its real applications to health care systems – outpatient visits. *International Journal of Healthcare Management*, **14**(1), pp. 114–122. **DOI**: https://doi.org/10.1080/20479700.2019.1616890.

Proudlove N. (2022). Use and misuse of queueing theory for hospital capacity decisions. In: J. Vissers, S. Elkhuizen, N. Proudlove (Eds.) *Operations Management for Healthcare*. Second Edition, London: Routledge, pp. 197–216. **DOI**: https://doi.org/10.4324/9781003020011.

Rema V. & Sikdar K. (2021). Optimizing patient waiting time in the outpatient department of a multi-specialty Indian hospital: The role of technology adoption and queue-based Monte Carlo simulation. *SN Computer Science*, **2**, art. no. 200. **DOI**: https://doi.org/10.1007/s42979-021-00595-8.

Saastamoinen K., Rissanen A., Suni J., Hyttinen J., Paakkunainen P. & Liakka A. (2023). Simulation of the queuing situation of patients at a health center. *Engineering Proceedings*, **39**(1), art. no. 72. **DOI**: https://doi.org/10.3390/engproc2023039072.

Saaty T.L. (1957). Résumé of useful formulas in queuing theory*. Operations Research*, **5**(2), pp. 161–200.

Shortle J.F., Thompson J.M., Gross D. & Harris C.M. (2018). *Fundamentals of Queueing Theory*, Fifth Edition. Hoboken: John Wiley & Sons.

Suleiman S., Burodo M.S. & Ahmed Z. (2022). An application of single and multi--server exponential queuing model in some selected hospitals of the North--Western Nigeria. *Asian Journal of Probability and Statistics*, **16**(2), no. AJPAS.82945. **DOI**: https://doi.org/10.9734/ajpas/2022/v16i230396.

Walker K., Jiarpakdee J., Loupis A., Tantithamthavorn C., Joe K., Ben-Meir M., Akhlaghi H., Hutton J., Wang W., Stephenson M., Blecher G., Buntine P., Sweeny A. & Burak T. (2022). Emergency medicine patient wait time multivariable prediction models: a multicentre derivation and validation study. *Emergency Medicine Journal*, **39**(5), pp. 386–393. **DOI**: https://doi.org/10.1136/emermed-2020-211000.

Xiong Y. (2023). Research and application analysis of the basic theory of queuing theory. In: P. Kar, J. Li & Y. Qiu (Eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Amsterdam: Atlantis Press, pp. 454–463. **DOI**: https://doi.org/10.2991/978-94-6463-300-9_46.

Yao D.D. (Ed.) (1994). *Stochastic modeling and analysis of manufacturing systems.* Springer Science & Business Media. **DOI**: https://doi.org/10.1007/978-1-4612-2670-3.

Zavanella L., Zanoni S., Ferretti I. & Mazzoldi L. (2015). Energy demand in production systems: A queuing theory perspective. *International Journal of Production Economics* **170**(B), pp. 393–400. **DOI**: https://doi.org/10.1016/j.ijpe.2015.06.019.

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 65–81

Analyzing Activities of Mobile App Users Who are Preparing for Driving Tests as Sources of Knowledge about Consumer Behavior

Anna Zapiór*

Abstract. This article presents the partial results from ongoing research that uses mobile applications that help individuals prepare for their driving license exams. The aim of the presented research is to analyze the activity of the users of these applications (including their daily activities, any tasks that are performed, and the lengths of times that are spent on sample exams and tests). The theoretical implication of the article is to draw attention to the time of the highest consumer activity, while the practical implication is to emphasize the importance of using ICT (particularly, mobile applications) in knowledge and information management, marketing decision-making, and education.

Keywords: mobile applications, knowledge management, information management, user activity

Mathematics Subject Classification: 91E40

JEL Classification: O32

Submitted: May 16, 2024

Revised: November 17, 2024

1. INTRODUCTION

Mobile applications have become very popular tools nowadays, being widely used in many areas of life (Böhmer et al., 2011). Interest in this source of information is growing year by year, as can be seen in Figure 1 (which shows the numbers of mobile app downloads worldwide during the years of 2019–2022).

* AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: anna.zapior@gmail.com

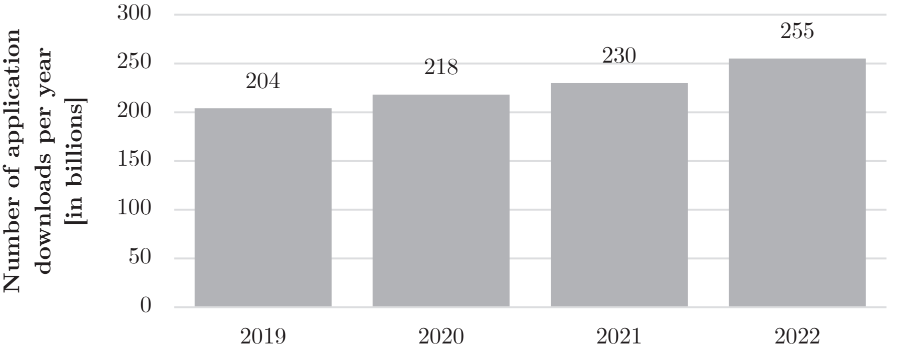


Fig. 1. *Number of mobile app downloads worldwide during years of 2019–2022 (Statista.com)*

Applications that are often sources of information and knowledge are gaining popularity in the field of education in a broad sense. Among others things, these offer learning through games and various forms of entertainment (gamification), the reviewing of materials (flashcards), and the testing of knowledge (for example, through quizzes). In 2020, nearly 167,000 educational applications were published during only the second and third quarters (Liftoff, 2020). Users who are preparing for various specialized exams are increasingly turning to various applications that feature sample exams. Following this trend, training companies are also expanding their products with proprietary tests (and even exams with questions from official publicly available databases). Examples of such applications can be found in many areas that requiring qualifications, such as driving licenses, medical exams, boat captain's/boat master's licenses, or pilot licenses.

Thanks to mobile technologies, users have easy and fast access to data that is of interest to them anywhere (Craik et al., 2019), with the ability to review and verify any necessary information (Zhang, 2022). This allows for learning new content as well as reviewing known content (which aids in the memorization process). Additionally, an analysis of user activity in the application allows for the precise tailoring of its content to the needs of its audience, thereby improving the quality of the training. Questions that frequently receive incorrect answers can be more thoroughly discussed during training sessions, and, various mnemonic techniques can be applied in the app itself (which allows for the better retention of any learned material). In a study that was conducted by Purnama et al. (2024, p. 323) on mobile learning (M-learning), the authors demonstrated, among other things, that the use of mobile technologies in education "will make it easier for teachers and students, (...) the learning process will be more effective and efficient, (...) students will be more active in the learning process, and the learning process will be more enjoyable." Other studies have shown that, compared to traditional methods, mobile learning can enhance students' enthusiasm for learning (Zhang, 2022). It has also been proven that M-learning is most effective when led by experienced teachers and integrated with clearly defined learning objectives (Naveed et al., 2023).

Considering that there is relatively little time between a launch and the peak activity time of an application (Chung et al., 2022), it is important to optimize its use for both consumers and producers. The present study took various variables into account to formulate a certain characteristic of consumer behavior; in the first part of the article, the research method is briefly presented, followed by a description of the concept of learning through testing as well as the research results. It concludes with possible future directions of the work.

2. PURPOSE OF PUBLICATION AND RESEARCH METHOD

The aim of the study is to attempt to identify the differences in the activities of the users of a mobile application over the course of days and determine the extent of their engagements at various times of the day (including the amount of time that is spent on preparation tests or exams).

The study is intended to identify certain trends in the use of applications, which will allow for modifying current applications and designing future ones while accounting for these factors in order to improve their functionality. The collected data can also be used for marketing analysis purposes.

As a research tool, a mobile application driving license test (Category B) was used, which is available for free in the Google Play store. No marketing campaign that promoted the product was used. The application is available for download under the brand names of three different Krakow driving school centers as well as under its own brand.

The application offers Category B driving license tests and is organized into four modules:

1) Learning Module – divided into subcategories such as "warning signs," "traffic signals," "signals given by the traffic controller," etc. Additionally, the questions in each subcategory were divided into groups, with each group containing up to 32 questions. A progress indicator was displayed for each group. While answering the questions, the users received real-time feedback on whether their responses were correct or incorrect. After completing a group of questions, a summary was displayed that showed the numbers of correct and incorrect answers (along with detailed information about the results).
2) Exam Module – at the start, the exam rules were presented (such as the number of questions, the time that was allowed for answering, and the minimum number of points that were required to pass the exam). The users were then shown a series of questions that needed to be answered. Each question displayed the points that were awarded for a correct answer as well as a time counter. At the end, the exam result was displayed (along with details on the correct and incorrect answers).
3) Statistics Module – this showed statistics such as the number of exams that were taken (including how many were passed, failed, or abandoned), the number of correct answers in each subcategory (with percentage indicators), and general information (such as the app's usage time and the time that was spent studying).
4) Settings Module – this allowed users to enable/disable app notifications and reset the statistics.

Upon the first launch of the app, a notice was additionally displayed that informed the users that statistical data would be collected for the proper functioning of the app, for research purposes, and to support the app's further development. The app also assured the users of their anonymity.

The research sample was selected randomly from those who downloaded the application from the Google Play store. Both the choice of the tool and the research sample were driven by the easily accessible and large research group – the Category B driving license exam is common and popular in society regardless of age group, economic status, or education level.

The time that was indicated on the figures as well as for the data analysis was measured in seconds. The data was collected using the so-called Unix Timestamp (POSIX time) and then converted to the Polish time zone. Seasonal time changes were taken into account in the studies. The data was grouped on the server based on UUIDs and collectively analyzed based on the activity. In the event of occurrences whose durations spanned across two days (e.g., from Thursday night to Friday morning), the data was counted toward the day on which the event ended.

Incomplete or damaged data was omitted from the analysis. Real-time reporting was not applied, as the application that served as the research tool was designed to operate offline.

The data that was collected via the application was divided into two categories: actions, and events:

1) actions referred to activities performed by user (such as clicking on selected option);
2) events recorded information about performed action (for example, start or end of learning session or exam).

"Actions" could generate an event (e.g., selecting the "exam" module) or not (e.g., selecting the settings module [which was not statistically significant from the perspective of the research]).

"Events" contained information such as the start and/or end times of an event (if this could be determined; i.e., if no error occurred during the data recording), the type of activity that was undertaken, and the duration of the event (if both the start and end times were successfully recorded).

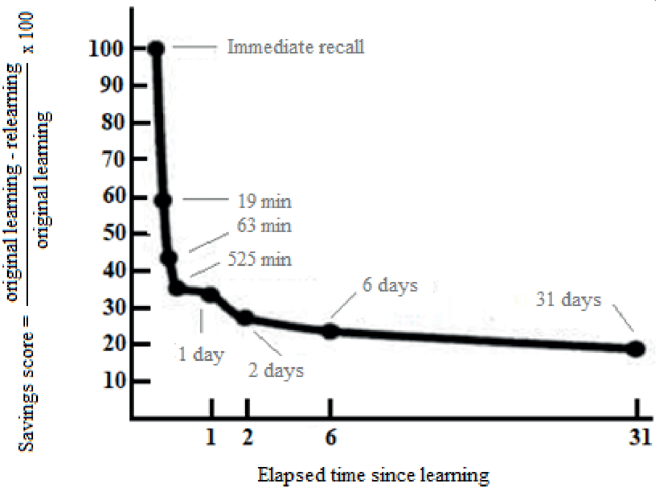
The study proceeded via the four following stages:

1) A market analysis of the mobile applications that were used in those sectors where a final exam was mandatory for granting licenses. This stage included an analysis of the literature on the subject that encompassed user profiles, previous research on the practical application of the apps and recommendations, and an analysis of the market for mobile applications (particularly, their popularity, product characteristics, and utility).
2) The design of a research tool (in the form of a learning-support application) that was compliant with the current standards and market requirements.
3) The deployment of the application and making it available to users, and conducting relevant research on a group of 356 users from July 23, 2020, through January 1, 2021.
4) An analysis of the collected data, the development of the results, the formulation of the conclusions, and directions for further research.

3. EBBINGHAUS CURVE AND TESTING EFFECT

The mobile applications in the current study were used to assess the level of knowledge through single-choice tests that covered the scope of road traffic regulations that were required for a Category B driving license exam. As numerous studies have shown, testing can not only be useful for assessing one's knowledge but also for enhancing one's memorization process (Bangert-Drowns et al., 1991; Roediger & Karpicke, 2006a; 2006b; Spitzer, 1939).

Numerous studies have demonstrated that testing can be used for assessing knowledge, and self-monitoring can enhance memorization and facilitate learning (Butler & Roediger, 2007; Larsen et al., 2013; McDaniel et al., 2007); these occur because the learning process involves the repetition of acquired content. While the timing and length of both the repetition process and the breaks between repetitions remain debatable, it is certain that the most significant knowledge deficits occur shortly after learning material (Fig. 2). It has also been shown that spaced repetitions lead to better results than single continuous-learning sessions or even learning with a single attempt at retrieving learned material (Fig. 3).



Elapsed time since learning	Retention [%]
Immediately	100
20 minutes	58
1hour	44
9 hours	36
1 day	33
2 days	28
6 days	25
31 days	21

Fig. 2. *Ebbinghaus forgetting curve (Yue, 2017)*

Furthermore, it has been shown that repeated testing with feedback improves learning (Roediger & Butler, 2011; Wiklund-Hörnqvist et al., 2014), and the testing effect itself enhances memory retention and brings benefits – even for individuals with cognitive impairments (Yang et al., 2021).

Taking the benefits of using testing applications into account, the activities of the users of such applications was analyzed, thus allowing for the adaptation of the contents and exam durations to consumer needs.

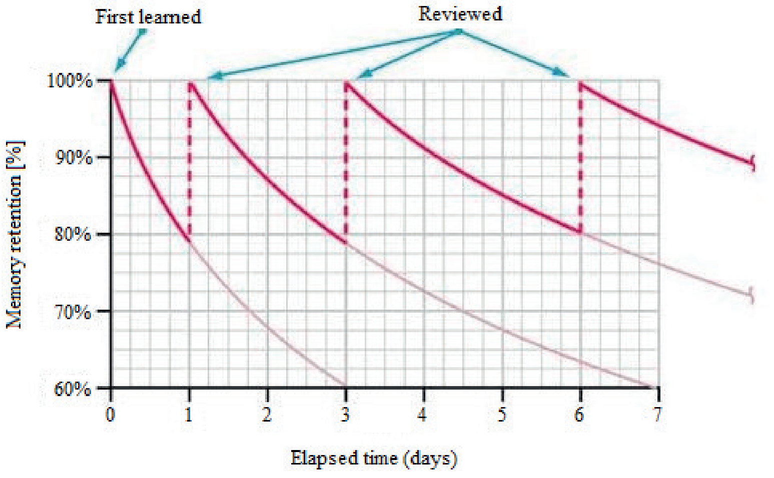


Fig. 3. *Ebbinghaus forgetting curve with marked repetition intervals (Bakhtiyari et al., 2014)*

4. ANALYSIS OF OBTAINED RESULTS

For a better understanding of the data that is presented in the article, the basic concepts that are used in its further parts are presented as follows:

- Session – the time that was counted from the beginning of an exam or test until its end.
- User – a single installation of the application (one installation of the application = one user). Each installation generated a unique UUID (universal unique identifier) – a unique user ID that was used to uniquely identify an object or entity on the Internet.
- Single attempt – each time a test, exam, or "retry" option was initiated (in the cases of questions that were marked with incorrect answers).
- Test (learning module) – consisted of a maximum of 32 questions from the selected module (i.e., "warning signs," "overtaking," "using external lights and vehicle signals," etc.). The time for completing the test was unlimited, and the selected answers provided immediate feedback on the correct and incorrect choices.
- Exam – consisted of 32 questions (20 questions from basic knowledge, and 12 questions from specialized knowledge) in accordance with the guidelines of the Ministry of Infrastructure. The examination procedure was consistent with that which was conducted in the examination centers. The time to complete the entire exam was 25 minutes, of which:
  - In the basic section, there were 20 seconds to read the question and 15 seconds to provide an answer. Among the questions that were drawn were ten questions that were worth three points each, six questions that were worth two points each, and 4 questions that were worth 1 point each.

- In the specialist section, the time that was allocated for familiarizing oneself with a question and providing an answer was 50 seconds, and the pool of randomly selected questions consisted of six questions that were worth three points each, four questions that were worth two points each, and two questions that were worth one point each. Feedback regarding the numbers of incorrect and correct answers was provided to the user only after an entire exam was completed.

– Reported data (i.e., days, activities, time spent on specific actions) – data that was collected during the use of the application and correctly sent to the collecting system. The data was sent in packets no more frequently than once per day. As a result, the system did not transmit partial (current) data; for example, no report was sent in a situation where a user installed the application and uninstalled it within 24 hours. This meant that, despite 669 users installing the application on the first day, reports from only 356 of them were sent after 24 hours. Considering the market-characteristic analysis, it was likely that the application was uninstalled due to its size and, consequently, its occupying an above-average amount of memory on a device. Studies by Ickin et al. (2017) showed that applications that occupy too much memory on a device is the third-most-common reason why users uninstall an application (the second and first places in the survey were errors/bugs in the application [leading to its crashing] and the application's uselessness, respectively). It is worth noting that the application's size was approximately 790 MB, while the statistically average size of applications with similar characteristics on the market was around 60 MB. This size was associated with additional features such as video playback and offline content viewing, which allowed users to access the materials within it even when they were offline. It is also significant that, due to the fact that full reports with data were received from 313 users, only this number of consumers were included in the presented study. Furthermore, it was assumed that the report was incomplete and not subject to analysis when a user started a session (test/exam); however, there were no records of its closure. This situation occurred when access to the application was interrupted during a session – the user closed the application, turned off the phone, or performed another event that led to its closure.

An analysis of the collected data (Fig. 4) indicated that, on average, the greatest numbers of active users occurred between 8:00 p.m. and 10:00 p.m. Large numbers of users could also be observed between 10:00 a.m. and 12:00 midnight (with a peak being observed at 3:00 p.m.). The application was used around-the-clock by at least one person each hour.

The significance of the users' daily activities (Fig. 4) can be understood when considering the number of days in which the reports contained activities from a given hour (Fig. 5). For example, reports were collected from 100 different days for the 3:00 p.m. hour, while the activities at 5:00 a.m. were recorded in the statistics from only 11 different days (note that the days with recorded hourly user activities were not necessarily consecutive; the figure shows the total number of such days in the application).

Therefore, the greatest user activity occurred between noon and 10:00 p.m. after adjusting the above information with the data that is presented in Figure 4. This was the period when the average numbers of active application users were high.

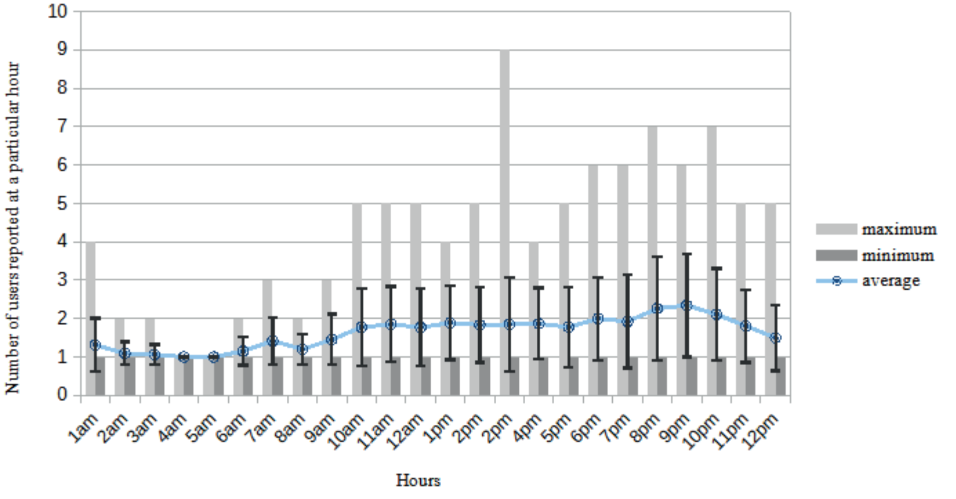


Fig. 4. *Users' daily activities, N = 313*

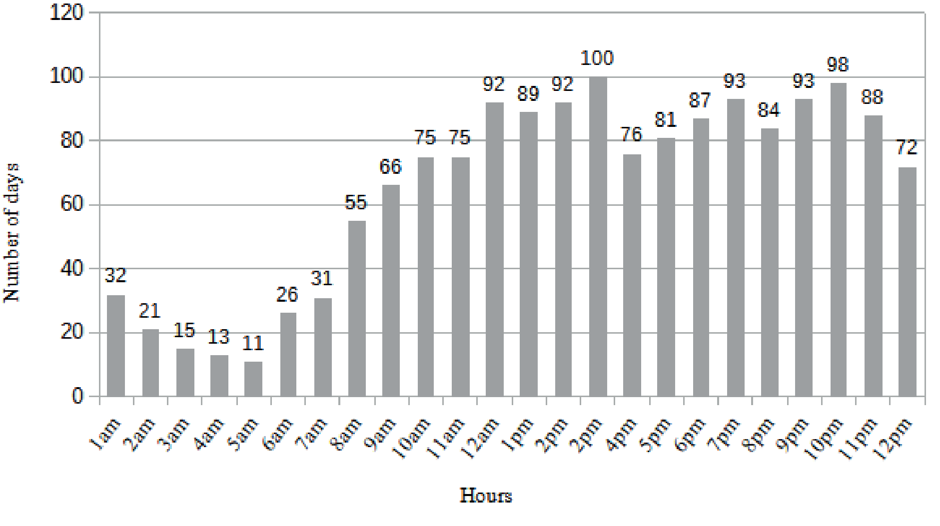


Fig. 5. *Numbers of days with reported activities at particular times, N = 313*

The above data regarding the users' activities was also confirmed by the data that was collected on the number of people who ended sessions (regardless of its type: the learning, exam, or question review mode with incorrect answers); the curve in Figure 6 largely overlaps with the curve in Figure 5. It is also worth noting that the increases in user activities during a specific period (e.g., at 3:00 p.m. or 8:00–10:00 p.m.) were correlated with the numbers of people who had completed sessions (which is shown for clarity in Figure 7). At the same time, it should be noted that starting a session did not require that the user would complete it – they could have interrupted it at any time by exiting the application or the session itself, resulting in a registration of activity without the completion of a session.

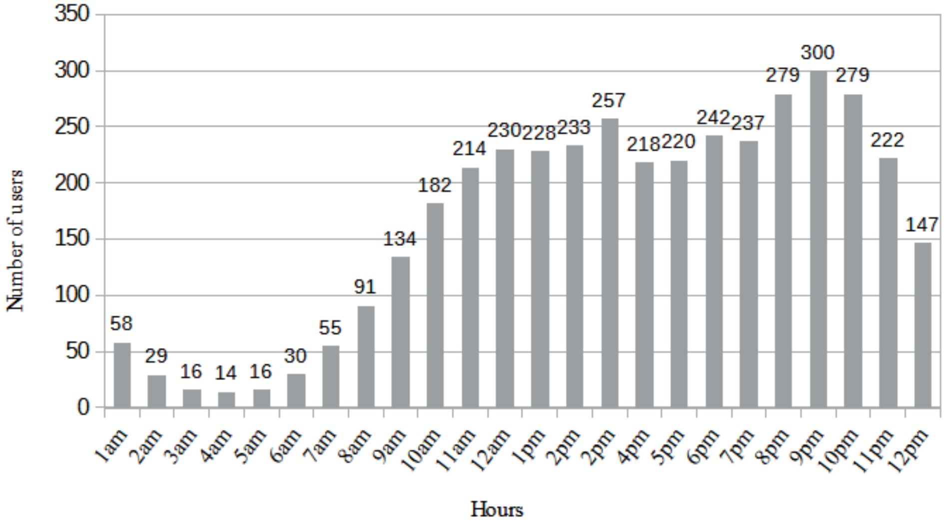


Fig. 6. *Number of people with completed session during particular hour of day, $N = 313$*

To construct the scatter plot below, the Pearson linear correlation coefficient was calculated based on the following formula:

$$r_p = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}.$$

Pearson correlation coefficient amounted to $r_p = 0.97$, where $\overline{x}$ and $\overline{y}$ are average parameter values, $N = 24$ (number of hours in a time interval), significance coefficient $p < 0.001$.

Regression line equation:

$$\hat{Y} = -38.1582 + 3.097X$$

Linear regression report:

$$R^2 = 0.94,\ F(1.22) = 351.01,\ p < 0.001$$

X predicted Y,

$$\beta = 3.1,\ p < 0.001,\ \alpha = -38.16,\ p = 0.004$$

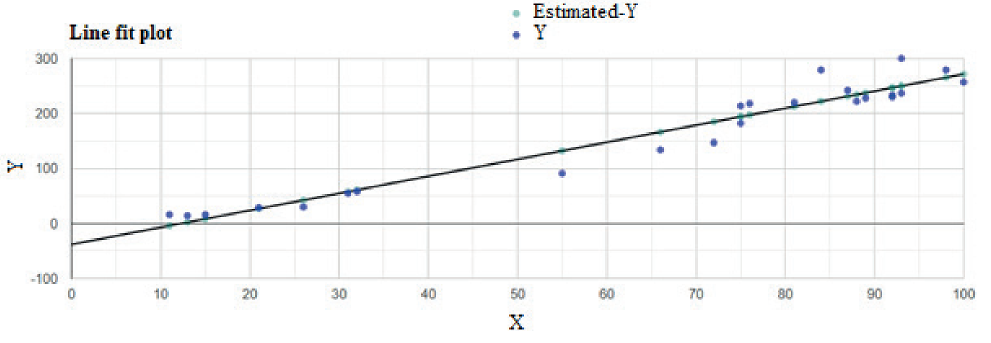


Fig. 7. *Correlation between numbers of days with reports sent during particular hour (X) and numbers of users during particular hour (Y)*

The data that is presented in Figure 7 illustrates the strong correlation between the number of users that were reported during a given hour and the number of days that were reported during that hour.

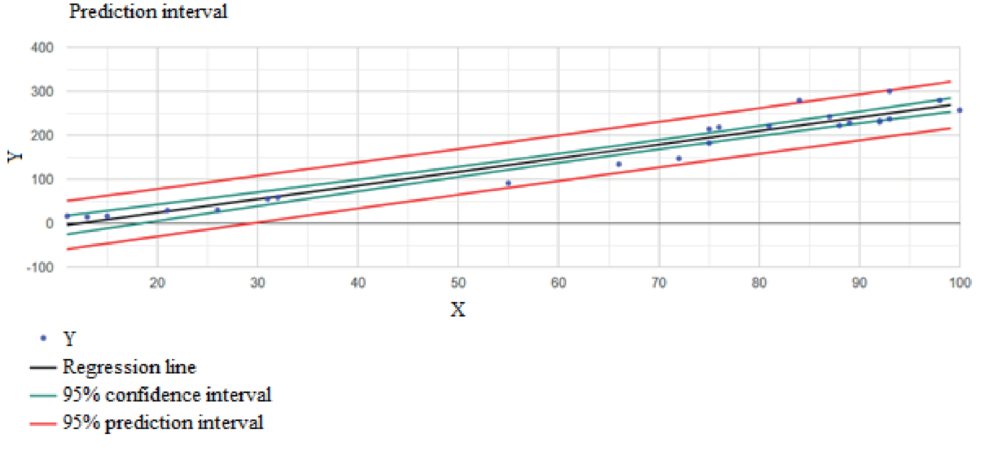


Fig. 8. *Prediction interval based on Figure 7, N = 313*

From the data that is shown in Figure 7, it follows that the data in Figure 8 was within the 95% prediction interval. Furthermore, Table 1 includes comprehensive details that concern the linear regression model that was derived from the obtained results (to enhance the understanding of the presented research).

Table 1. *Regression ANOVA*

Source	Degrees of freedom	Sum of square	Mean square	F statistic (df_1, df_2)	p-value
Regression (between $\hat{y}_i$ and $\bar{y}$)	1	209,456.076	209,456.076	351.0112 (1,22)	5.218e-15
Residual (between y_i and $\hat{y}_i$)	22	13,127.8823	596.7219	–	–
Total (between y_i and $\bar{y}$)	23	222,583.9583	9677.5634	–	–

The regression analysis (Table 1) showed the following:

1) Y and X relationship:
   - R-squared (R^2) equaled 0.941; this meant that 94.1% of variability of Y was explained by X;
   - correlation (R) equaled 0.9701; this meant that there was very strong direct relationship between X and Y;
   - standard deviation of residuals (S_{res}) equaled 24.4279;
   - slope $b_1 = 3.097$ CI [2.7542, 3.4398] meant that, when one increased X by 1, value of Y increased by 3.097;
   - y-intercept $b_0 = -38.1582$ CI [−62.7887, −13.5277] meant that, when X equaled 0, prediction of Y's value was −38.1582;
   - x-intercept equaled 12.321.
2) Goodness of fit:
   - overall regression – right-tailed, $F(1.22) = 351.0112$, p-value $= 5.218\text{e-}15$: as p-value $< \alpha$ (0.05), this meant that H_0 was rejected;
   - linear regression model $Y = b_0 + b_1X + \varepsilon$ provided better fit than model without independent variable that resulted in $Y = b_0 + \varepsilon$;
   - slope (b_1) – two-tailed, $T(22) = 18.7353$, p-value $= 5.218\text{e-}15$: for one predictor, this was same as p-value for overall model;
   - y-intercept (b_0 – two-tailed, $T(22) = -3.2129$, p-value $= 0.004008$: hence, b_0 was significantly different from zero.
3) Residual normality – the linear regression model assumed normality for the residual errors. The Shapiro–Wilk p-value equaled 0.2374, and it was assumed that the data was normally distributed.
4) Outliers – the data did not contain any outliers.

Further, complementing Figures 5 and 6 is Figure 9, where information on what the users preferred to use during their activity times can be obtained.

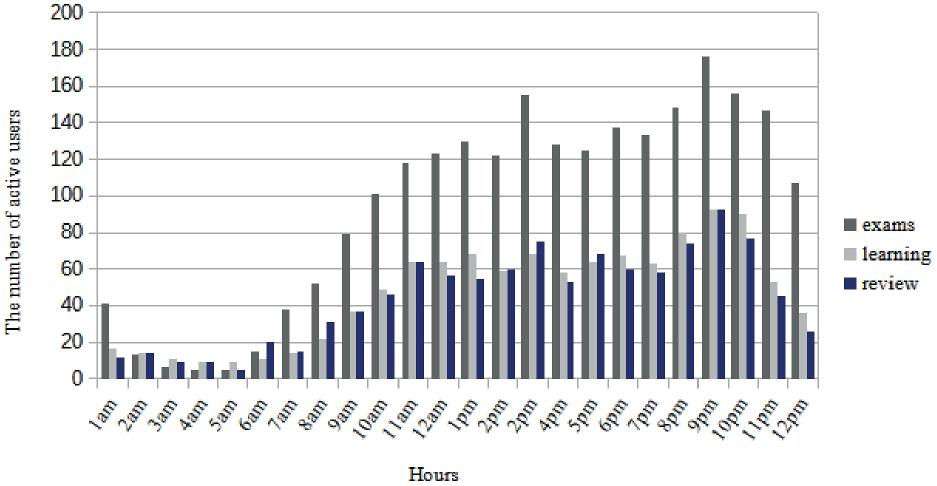


Fig. 9. *Numbers of people who used selected module during particular hours*

Subsequent analyses revealed that the vast majority of the users chose the exam module, while nearly half as many chose the learning module. It is worth noting at this point that the "review" option was part of the "learning" module (Chart 10); this appeared when a user provided an incorrect answer to at least one of the questions during a learning session. Then, they received the number of questions that they answered incorrectly in the feedback, along with the "repeat" option (the review applied only to those questions with the previously given incorrect answers).

More data about the users' activities is provided by an analysis of Figure 9, which shows that clear peaks in the quantities of the responses that were given happened during the hours of 11 a.m., 2 p.m., 8 p.m., 1 p.m., and 1 a.m. (despite the fact that the greatest user activity occurred during the hours of 12 p.m.–3 p.m. and 7–10 p.m. [Fig. 5]). By complementing this data with the information from Figure 6, it can be assumed that, despite the large number of responses that were given, the sessions were not being terminated. This situation occurred in two scenarios:

1) In the "exam" module: when a user started an exam and began to provide answers but interrupted it before completing the whole exam (they could then exit the exam without finishing it and move to the learning or question review).
2) In the "learning" module: when a user started a specific set of questions but interrupted it before completing it (they could then proceed to another set of questions, the questions with incorrect answers, or the exam module).

In both of the mentioned cases, the numbers of answers that were provided were counted, but completions of the full sessions were not recorded.

Furthermore, it was checked whether the numbers of provided answers decreased during the night-time hours; this would suggest a decrease in attention due to one's natural circadian rhythm.

Interestingly, the research showed that the users provided similar average numbers of answers regardless of the time of day/night (Table 1). A slight increase could be observed during the hours of 1:00, 3:00, and 4:00 a.m.; however, this fell within the margin of error. Considering the daily activity of the users (Fig. 10), clear extremes occurred during the following hours: 1 a.m., 11:00 a.m., 2:00 p.m., 8:00 p.m., and 11:00 p.m. As with the information that was mentioned above, this also requires further research and deeper analysis.

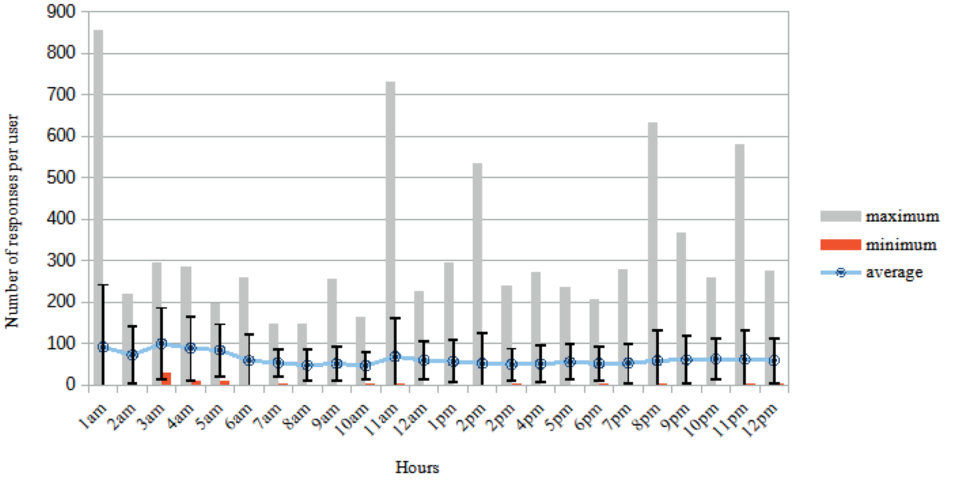


Fig. 10. *Average numbers of user responses during active hours*

Another significant aspect under study were the durations of the sessions themselves (Fig. 11). The median of the average duration of time that was spent on a session (exam) was 326 seconds (approximately 5.4 minutes), which was calculated based on the following formula:

$$M_e = \frac{x_n + 1}{2}.$$

This meant that, on average, the users spent about 10 seconds on each question (326 seconds/32 questions).

The average time that was spent on an active examination session (i.e., from the start to the end) was 336 seconds (5.6 minutes); this was based on the following equation:

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{2}\sum_{i=1}^{n} x_i,$$

where:

$x_1, x_2, \ldots$ – individual values for which average was calculated;

n – ample size equal to 313.

The average was calculated according to the formula for the arithmetic mean (see Fig. 11). This time was measured from the beginning of an exam session to its end (including the time that was taken to play any video material if it was part of a question).

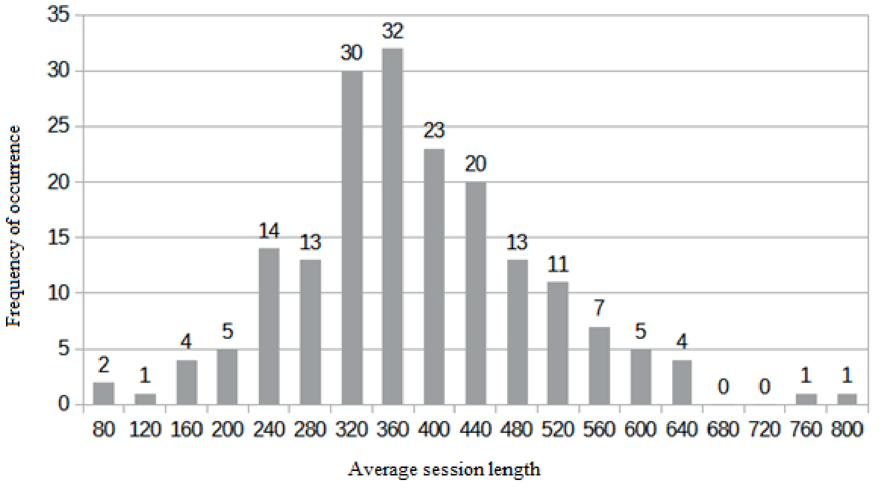


Fig. 11. *Histogram of average lengths of learning sessions, $N = 313$*

In the cases of the learning and exam modules, the users also spent an average of 10 seconds per answer.

The maximum length of a reported session among the average values for the users was 1678 seconds – approximately 28 minutes (based on the data from both the exam and learning modules), and the minimum average was 1 second; this suggested that the user finished right after starting a test/exam or that another event occurred to cause the closure of the application. Both the maximum and minimum reported session durations were not included in Figure 4, as these were individual cases.

5. CONCLUSIONS

An analysis of the user activity in the knowledge-test application for a Category B driving license revealed that the highest user activity occurred between 12:00 p.m. and 10 p.m. This result aligned with our predictions, as it corresponds to the human circadian rhythm wakefulness cycle.

The average time that was spent on the exam sessions (5.6 minutes) as well as the median of the average exam duration (5.4 minutes) aligned with Statista's research (2018) that was conducted during the years of 2017 and 2018, which showed that the average time that was spent by users in the application was 5.6 minutes for iOS devices and 6.6 minutes for Android devices. Similar results (7 minutes) were obtained in the study by Deng et al. (2018).

This data can provide a significant indication for designing the lengths and frequencies of displaying potential advertisements as well as for further research on session durations.

It is interesting that not only was the length of a session constant but also the number of responses that were given regardless of the time of day/night. This fact was surprising in that one could hypothesize that late hours and the resulting fatigue and loss of concentration would affect the duration of a session or the number of responses that are given, yet these values were relatively constant.

Furthermore, the data that is presented in Figure 7 indicated a strong correlation between the number of users that were reported at a given hour, and the numbers of days that were reported at that hour raised questions:

– Is there a correlation between the days of the week and specific hours in user activity?
– What is the distribution of user activity by hour on a selected day of the week?
– Why are there lower levels of activity during certain hours on certain days during a selected time period while it tends to increase during other hours?

The presented results are subject to further analysis and are being expanded with studies that involve a larger research sample and a longer duration of the study.

The presented data can be used in planning tests by educational institutions and companies as well as when designing training programs. They can also be applied in educational centers (such as schools) or private companies to expand and maintain knowledge in selected sectors. It is essential to ensure that the design of educational applications includes clearly defined learning objectives and is overseen by experts in the relevant field. Exam sessions should be tailored to the content but should not last longer than six minutes. Additionally, it is beneficial to divide large amounts of learning material into smaller sections and provide real-time feedback; these help increase user motivation.

The peak activity times of users (generally, from the morning until the evening) allows for the real-time implementation of various quizzes in order to reinforce one's knowledge.

6. DIRECTIONS FOR FURTHER RESEARCH

The collected data came from reports that were sent by the system every 24 hours. Real-time reporting was not used in the study, which could have had a significant impact on the presented results – especially on the ratio of the numbers of users to their activities (in the case of uninstalling an application before the end of a day, this usage data was not reported). Further research is also subject to the activities of users who started sessions but whose closures were not reported. This means that a partial analysis of hourly activities and the activities within the application itself is still possible.

In the presented studies, the type of the presented material was not taken into account in relation to the session length (answering questions that require the viewing of a short video may take users more time than simply answering text-based questions).

Another aspect that was considered was whether users achieved better results over time while using the application. Ongoing research should also examine whether any progress in one's achieved results can be observed in the longer term.

REFERENCES

Bakhtiyari K., Salehi H., Embi M.A., Shakiba M., Zavvari A., Shahbazi-Moghadam M. & Mohammadjafari M. (2014). Ethical and unethical methods of plagiarism prevention in academic writing. *International Education Studies*, **7**(7), pp. 52–62. **DOI**: https://doi.org/10.5539/ies.v7n7p52.

Bangert-Drowns R.L., Kulik J.A. & Kulik C.C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85(2), pp. 89–99. **DOI**: https://doi.org/10.1080/00220671.1991.10702818.

Böhmer M ., Hecht B., Schöning J., Krüger A. & Bauer G. (2011). Falling asleep with Angry Birds, Facebook and Kindle: A large scale study on mobile application usage. *MobileHCI'11: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, Stockholm, Sweden, pp. 47–56. **DOI**: https://doi.org/10.1145/2037373.2037383.

Butler A.C. & Roediger H.L. III (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, **19**(4/5), pp. 514–527. **DOI**: https://doi.org/10.1080/09541440701326097.

Chung M., Sharma L. & Malhotra M.K. (2022). Impact of modularity design on mobile app launch success. *Manufacturing & Service Operations Management, INFORMS*, **25**(2), pp. 756–774. **DOI**: https://doi.org/10.1287/msom.2022.1181.

Craik A., He Y. & Contreras-Vidal J.L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. *Journal of Neural Engineering*, **16**(3), pp. 1–10. **DOI**: https://doi.org/10.1088/1741-2552/ab0ab5.

Deng T., Kanthawala S., Meng J., Peng W., Kononova A., Hao Q., Zhang Q. & Prabu D. (2019). Measuring smartphone usage and task switching with log tracking and self-reports. *Mobile Media & Communication*, **7**(1), pp. 3–23. **DOI**: https://doi.org/10.1177/2050157918761491.

Ickin S., Petersen K. & Gonzalez-Huerta J. (2017). Why do users install and delete apps? A survey study. In: A. Ojala, H. Holmström Olsson, K. Werder (Eds.) *Software Business. ICSOB 2017. Lecture Notes in Business Information Processing*, 304, Springer, Cham, pp. 186–191. **DOI**: https://doi.org/10.1007/978-3-319-69191-6_13.

Larsen D.P., Butler A.C., Roediger H.L. III (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Medical Education*, **47**, pp. 674–682. **DOI**: https://doi.org/10.1111/medu.12141.

Liftoff (2020). Number of apps released worldwide between 3rd quarter 2019 and 2nd quarter 2020, by category. **URL**: https://statista.com [17.03.2023].

McDaniel M.A., Anderson J.L., Derbish M.H. & Morrisette N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, **19**(405), pp. 494–513. **DOI**: https://doi.org/10.1080/09541440701326154.

Naveed Q.N., Choudhary H., Ahmad N., Alqahtani J. & Qahmash A.I. (2023). Mobile learning in higher education: A systematic literature review. *Sustainability*, **15**(18), art. no. 13566. **DOI**: https://doi.org/10.3390/su151813566.

Purnama Y., Lewa I., Syilvia W.J., Fauza A. & Wahyuni S. (2024). Mobile learning revolution: Harnessing the potential of smartphones in the learning process. *Journal International of Lingua and Technology*. 3(2), pp. 313–328. **DOI**: https://doi.org/10.55849/jiltech.v3i2.683.

Roediger H.L. III & Butler A.C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, **15**(1), pp. 20–27. **DOI**: https://doi.org/10.1016/j.tics.2010.09.003.

Roediger H.L. III & Karpicke J.D. (2006a). The power of testing memory basic research and implications for educational practice. *Perspectives on Psychological Science*, **1**(3), pp. 181–210. **DOI**: https://doi.org/10.1111/j.1745-6916.2006.00012.x.

Roediger H.L. III & Karpicke J.D. (2006b). Test-enhanced learning: taking memory tests improves long-term retention. *Psychologica science*, 17(3), pp. 249–255. **DOI**: https://doi.org/10.1111/j.1467-9280.2006.01693.x.

Spitzer H.J. (1939). Studies in retention. *Journal of Educational Psychology*, **30**(9), pp. 641–656. **DOI**: https://doi.org/10.1037/H0063404.

Statista (2018). Average duration of in-app engagement sessions worldwide from September 2017 to August 2018, by operating system. **URL**: https://www.statista.com/statistics/202485/average-in-app-session-durations-operating-systems-worldwide/ [1.11.2023].

Wiklund-Hörnqvist C., Jonsson B. & Nyberg L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, **55**(1), pp. 10–16. **DOI**: https://doi.org/10.1111/sjop.12093.

Yue N. (2017). Computer multimedia assisted English vocabulary teaching courseware. *International Journal of Emerging Technologies in Learning*, **12**(12), pp. 67–78. **DOI**: https://doi.org/10.3991/ijet.v12i12.7955.

Yang C., Luo L., Vadillo M.A., Yu R. & Shanks D.R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, **147**(4), pp. 399–435. **DOI**: https://doi.org/10.1037/bul0000309.

Zhang X. (2022). The influence of mobile learning on the optimization of teaching mode in higher education. *Wireless Communications and Mobile Computing*, **1**, pp. 1–9. **DOI**: https://doi.org/10.1155/2022/5921242.

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 83–98

Application of Basic Machine-Learning Classifiers for Automatic Anomaly Detection in Shewhart Control Charts

Aleksander Woźniak*, Klaudia Krawiec**, Roger Książek***

Abstract. In today's dynamic technological environment, innovation plays a crucial role – especially for manufacturing enterprises that constantly strive to improve the quality of their products. This article examines the quality-management issue in a company producing car rims. It was identified that real-time quality control can sometimes be unreliable due to controller fatigue, leading to erroneous data interpretation or delayed responses to deviations in the production process. The study aimed to investigate the possibility of eliminating or significantly reducing these errors by employing a tool that is based on artificial intelligence. The article covers the preparation of training data, the training of classifiers, and the evaluation of their effectiveness in analyzing control charts in real time. The adopted hypothesis assumes that machine-learning classifiers can be effective methods of support for quality controllers. The research began with collecting measurement data from the machine and dividing it into training and test sets. The obtained results were evaluated using standard quality measures for machine-learning models. The results showed that the use of artificial intelligence can bring significant benefits in improving quality supervision in the production process of car rims.

Keywords: Machine Learning, Artificial Intelligence, AI, Statistical Process Control, SPC, Quality Control, Classifiers, Quality Metrics, Python Programming, Car Wheel, Quality Issues

Mathematics Subject Classification: 68T10

JEL Classification: C44

Submitted: June 5, 2024

Revised: October 28, 2024

* AGH University of Krakow, Faculty of Energy and Fuels, Krakow, Poland, e-mail: alekwozniak@student.agh.edu.pl
** AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: klakrawiec@student.agh.edu.pl
*** AGH University of Krakow, Faculty of Management, Krakow, Poland, e-mail: roger@agh.edu.pl

DOI: https://doi.org/10.7494/dmms.2024.18.6345

1. INTRODUCTION

The origins of Shewhart control charts can be traced back to the 1920s, when they were first developed for quality control. These represent a statistical tool that monitors process stability by tracking sequence samples and identifying patterns that deviate from expected norms. By distinguishing between common-cause and special-cause variations, this method enables the early detection of shifts in production, thus ensuring consistent quality control over time. Shewhart control charts have been employed for qualitative analysis, becoming a key tool for collecting data during production and later analyzing it (Shewhart, 1926).

It needs to be stressed that Shewhart control charts remain in use even today; however, their roles have evolved with modern digital tools. This shift has introduced several important considerations regarding the role of AI in quality control. It remains uncertain whether AI-based systems can fully replace human operators or if they will function more effectively as supplementary tools. The potential of AI to improve the accuracy and precision of measurements in rim production is significant; however, challenges and limitations come with its implementation. The effectiveness of different machine-learning algorithms in identifying and eliminating measurement errors must also be evaluated, along with the necessary data for adequately training AI models. Addressing these uncertainties is crucial for determining how AI can be optimally integrated into modern manufacturing processes (Malindzakova et al., 2023; Tran et al., 2022).

In the rim-manufacturing company to which this paper refers, their rims are sold on a global scale (both branded and for retail purposes). Quality control operators in the company were responsible for manual production measurements and the use of a special machine during the MD860 testing phase, which measured the necessary dimensions of each produced rim. Despite the use of Shewhart control charts, a key issue persisted in continuous quality control: the system relied heavily on human cognitive capabilities. Employee fatigue became a significant concern, as the constant manual supervision that was required for monitoring rims through the measurement device led to decreased efficiency and a heightened risk of errors. This human-centered limitation underscored the need for more-automated and -reliable solutions in the quality control process. This became the motivation for using the data sets from this process for research into the potential advantages of using simple AI classifiers over decision-making algorithms.

This paper showcases the process of selecting the most suitable machine-learning classifiers for supporting the automated detection and recognition of the patterns that were recorded in the Shewhart control charts that were used in the quality control in the above-mentioned rim-manufacturing process. Data sets that detailed the features of the rim-manufacturing process were utilized to train and evaluate selected machine-learning classifiers to detect and recognize classic seven-sample sequences that signal process degradation. Based on the performed literature review (which is presented in more detail in Section 2, showing the scarcity of works on the application of basic machine-learning classifiers for automatic analyses of seven-sample sequences in Shewhart control charts), data from the rim-manufacturing process was used to conduct the research. The study focused on basic ML classifiers due to their simplicity, thus allowing the presented experiment to be repeated on data sets from other processes without requiring significant hardware resources nor the performance of tedious advanced-programming work. The preparation of simple classifiers for pattern analysis is less time-consuming and less engaging than, for example, the training of an artificial neural network; so, the classifiers can be used to identify potentially promising research areas such as an n-sample sequence that precedes a seven-sample sequence that is indicative of process degradation. As this was a pilot study, the main research question (RQ) that was posed was as follows: which basic ML classifiers achieve acceptable performance (F1-score) in recognizing seven-sample patterns that are indicative of process degradation?

This paper is structured as follows. In Section 2, the research gap of insufficient research on the application of machine learning for enhancing statistical-process control via the automated analysis of Shewhart control charts is identified based on a systematic literature review. In Section 3, the control sequences that were selected for the experiments are presented, while in Section 4, the approach for training and evaluating the performance of the selected ML classifiers is explained. Section 5 reports and discusses the obtained results, and the main conclusions are presented in Section 6.

2. PAPER POSITIONING

Since the early 1990s, there has been growing interest in applying machine learning (ML) to recognize patterns on Shewhart control charts. These studies have typically approached the subject from the perspective of enhancing the detection accuracy and efficiency of the deviations from the process norms and anomalies. Techniques like support vector machines (SVMs) have emerged as particularly effective in automating the analysis of data from control charts – often outperforming traditional statistical methods in many use cases. However, much of the existing research has tended to overlook the crucial aspect of interpreting the results generated by ML-based systems – especially in the context of more dynamic and complex production environments.. This gap in the literature is significant because interpreting ability is key to ensuring that these systems can be effectively implemented in real-world scenarios; therefore, this article seeks to fill this gap by conducting a systematic review of the literature, focusing on the integration of ML with Shewhart control charts. The aim is to identify any recent developments that address the interpret ability of ML outcomes and explore how these approaches can be applied in actual production processes (Hwarng, 1992; Shewhart, 1992)

For this study, scientific articles from the SCOPUS database that covered the period from 2010 through 2024 were selected. The search criteria included the use of key phrases that were related to ‘Machine learning,’ ‘Machine learning + SPC,’ and ‘Machine learning + automotive.’ The study was conducted in September 2024 using the results of database queries from the day of September 4, 2024; the data filtering followed these steps:

1. **Machine learning** – included articles where the phrase ‘Machine learning’ was present. The number of results increased year by year (from 46,162 publications in 2010 to 780,574 in 2024).

2. **Machine learning + SPC** – included articles that contained both 'Machine learning' and 'SPC'. The number of results grew from 232 in 2010 to 2279 in 2024.
3. **Machine learning + automotive** – included articles that contained both 'Machine learning' and 'automotive'. The number of results rose from 125 publications in 2010 to 3060 in 2024.
4. **Machine learning + automotive + SPC** – included articles where the phrase 'Machine learning' and 'automotive' and 'SPC' was present. In total, 37 articles were identified that met all of the above criteria. Most of these were rejected due to not fitting within the thematic scope of the case at hand. A few were selected for analysis.

In those articles that focused on the utilization of machine-learning classifiers for automatic anomaly detection on Shewhart control charts, the data was analyzed and utilized to recognize irregularities (in a similar fashion as with other studies). However, Staněk et al. (2023) adopted the real-time recognition and classification of wheels and rims as the primary goal, while Rameshkumar et al. (2021) focused on using acoustic-emission features to predict the conditions of grinding wheels; the article concentrated on using Shewhart control charts to detect any anomalies in the processes. Compared to Lee et al. (2023), which examined the recognition of wheel-component conditions using machine learning, the research that this article focused on was anomaly detection in general (rather than specific wheel components). Additionally, there are articles such as Fang et al. (2023) and Krummenacher et al. (2017), which also employed machine-learning techniques to improve efficiency and safety in the railway industry. Despite the differences in the application and research area, all of these articles harnessed the potential of machine learning to enhance the effectiveness and quality of production processes and identify any abnormalities.

Moser et al. (2021) presented a machine-learning-based emission model for diesel engines that supported on-board diagnostics. This application of ML in automotive systems contributes to improving the control of emissions and is relevant to the automotive industry's performance and environmental sustainability. In Oh et al. (2019), the authors introduced a real-time quality-assessment system for an automotive-production process that used support vector machines (SVMs). The research was highly relevant to automotive safety and control (SPC), as it leveraged machine learning to ensure manufacturing precision in the assembly of sunroofs. Sharmin et al. (2022) explored a machine-learning approach to intrusion detection in automotive CAN networks. This research focused on cybersecurity for vehicle networks, making it highly relevant to automotive SPC by ensuring that secure data flows in in-vehicle systems. In Staf and McKelvey (2018), the authors proposed predictive models for brake performance in vehicles. While this paper did not heavily focus on ML, the predictive nature of the models ties into vehicle control systems and contributes to performance optimization, making it somewhat relevant to SPC. Lampe and Meng (2024) discussed a curated data set for improving automotive cybersecurity via machine learning. The relevance to SPC was clear, as it provided a foundational data set for developing ML models that enhance automotive network security. In Dettu et al. (2024), the focus was on integrating physical and virtual models for optimizing vehicle dynamics. While machine learning was not a core aspect, the approach was relevant to SPC through its emphasis on improving vehicle performance and control.

True et al. (2021) explored THz technology for semiconductor testing; this was not directly related to automotive SPC or machine learning, so this paper was less relevant to the automotive field. Lokman et al. (2019) provided a comprehensive review of intrusion-detection systems for automotive CAN networks, focusing on the use of ML for network security. It was highly relevant to automotive SPC – particularly in the context of enhancing vehicle cybersecurity that uses machine-learning techniques.

In Xu et al. (2024), the focus was on digital twin technologies for materials science. While interesting, it did not directly apply to automotive SPC or ML, making it less relevant to this domain. Zhou et al. (2023) reviewed monitoring techniques for selective laser melting in manufacturing. Though relevant to advanced manufacturing processes, it did not have a direct connection to ML or automotive SPC, thus making it less applicable to the field. The article by Mjimer et al. (2023) discussed the role of ML in industrial continuous improvement, but it was not focused on automotive applications; therefore, it was not strongly relevant to automotive SPC. Gong et al. (2023) reviewed machine learning in the context of laser-based manufacturing. Although it touched on ML, the focus was more on manufacturing processes outside the automotive industry, making it less relevant to automotive SPC. The paper by Benzaza et al. (2023) discussed improving quality-management systems in the automotive industry, but it did not focus heavily on machine learning. While relevant to automotive SPC in terms of quality control, the lack of a strong ML focus made it less applicable. Tsenev and Ivanova (2022) provided insights into using ML to evaluate control systems in automotive production lines; this was directly relevant to both automotive SPC and machine-learning applications. The paper by Lestyán et al. (2019) applied machine learning to identify drivers based on CAN network data. It was highly relevant to automotive SPC and ML – particularly in the context of vehicle data security and driver identification. Minawi et al. (2020) discussed a machine-learning-based system for detecting intrusions in automotive CAN networks. This work was highly relevant to automotive SPC, focusing on cybersecurity. In Habibullah et al. (2024), the authors explored software engineering for automotive perception systems. Although relevant to automotive systems, the paper did not focus on ML, making it only partially aligned with automotive SPC. Park and Baek (2023) presented a cyber-attack detection system using decision trees for automotive cyber-physical systems. This paper was highly relevant to both automotive SPC and ML due to its focus on vehicle network security. Lampe and Meng (2023) introduced a data set for automotive intrusion detection, leveraging ML for enhanced vehicle security; this was directly applicable to both automotive SPC and ML. Shi et al. (2016) focused on validating ML systems for autonomous driving, thus ensuring reliability in automotive applications. This paper was highly relevant to automotive SPC and ML – particularly in the field of autonomous vehicles. Escobar et al. (2021) discussed how AI and ML could enhance automotive manufacturing quality control, thus making it relevant to both automotive SPC and ML. In Kidmose and Meng (2024), the authors examined and evaluated data sets for intrusion-detection in automotive systems using machine learning. This was directly applicable to both ML and automotive SPC.

Cui et al. (2021) applied machine learning for predictive modeling in materials science, but it was not specifically focused on automotive SPC. In Robinson et al. (2020), the authors discussed testing techniques for automotive component reliability; however, there was little connection to machine learning, thus making it less relevant to automotive SCP. Hofmann et al. (2024) focused on ML applications in additive manufacturing, but it was not directly relevant to automotive SCP. Kalyanasundaram et al. (2018) applied machine learning for detecting denial of service (DoS) attacks on automotive CAN networks, thus making it highly relevant to both automotive SCP and ML. The review by Paturi et al. (2023) focused on machine-learning applications in manufacturing, but it was not directly related to automotive SCP (making it less relevant). Finally, Alfardus and Rawat (2023) proposed a hybrid machine-learning method for detecting cyberattacks in automotive networks, thus making it highly relevant to both ML and automotive SCP.

The reason for choosing simple classifiers is that they can analyze production data and detect subtle patterns that might indicate any potential quality issues that a human controller might overlook. These minor deviations can be easily identified by an algorithm, but they may be missed by manual inspections. The manual inspection of each rim in mass production is time-consuming and prone to errors, whereas ML can quickly analyze vast amounts of production data – especially when combined with large data sets; this increases process efficiency. Additionally, machine-learning models can be trained on various data sets, allowing them to adapt to changing production conditions or new rim specifications, providing greater flexibility as compared to static quality control methods. Following this approach, concise explanations of confusion matrixes and quality metrics such as accuracy and precision are included to ensure that those readers without programming knowledge can better understand the text.

3. SEQUENCE DETECTION IN STATISTICAL PROCESS CONTROL

To conduct the experiment, specific software that is capable of calculating and simulating certain results for analysis is needed (such as Python or Excel). Without the appropriate methodology, however, these tools will be useless. To adequately conduct the experiment, Statistical Process Control (SPC) was utilized; this is a tool that is used for monitoring and controlling production processes in order to ensure that products meet specified quality standards. SPC relies on the collection and analysis of process data, enabling quick responses to deviations from the expected quality of a product.

The main quality issues that SPC helps address include process deviations, process instability, and variability control. Deviations can arise from factors such as changes in raw materials, machinery, or environmental conditions. SPC allows for the early detection of these deviations, enabling corrective actions to be taken before defective products leave the production line.

Process instability refers to fluctuations that can affect product quality. SPC enables the monitoring of these fluctuations and the identifications of their causes, allowing for appropriate adjustments to be made. Control of process variability is crucial, as excessive variations can lead to unpredictable quality outcomes. SPC helps control and reduce this variability, resulting in more stable product quality. SPC can also be used to monitor processes that are influenced by multiple variables; this allows for identifying significant variables and controlling them in order to ensure stable quality. In Figure 1, the use of one of the SPC methods is illustrated; namely, having seven or more points on one side of the nominal line without crossing it.

The triangles in Figure 1 highlight sequences of seven or more consecutive points that exceed the nominal line, thus signaling deviations from the expected process behavior. The first such deviation begins at Sample 5, which has a value of 6. Following this, the next six samples (from Sample 6 through Sample 11) also show values that are above the nominal line (with Sample 11 reaching a value of 7). As per the rule in place, when seven or more points consecutively exceed the nominal line, they are classified as being erroneous; this means that all of the measurements from Samples 5 through 11 are considered to be incorrect and are assigned a classification of 1. The sequence concludes at Sample 14 (with a value of 6); this is the last measurement above the nominal line, making the previous six measurements also part of the error pattern.

In a similar manner, the circles in Figure 1 represent sequences of seven or more consecutive points that are below the nominal line. These points signal a drop below the expected range (again, indicating process deviations). The first such deviation occurs at Sample 17 (which has a value of 1). The following six measurements (up to Sample 20, with a value of 2) remain below the nominal line, thus fulfilling the rule for classifying the entire sequence as being erroneous. The final erroneous measurement in this sequence occurs at Sample 24 (with a value of 0), as it completes a series of seven or more points that are consecutively below the nominal line. As a result, this group of measurements is also classified as 1 (for incorrect results) (Helmold, 2021).

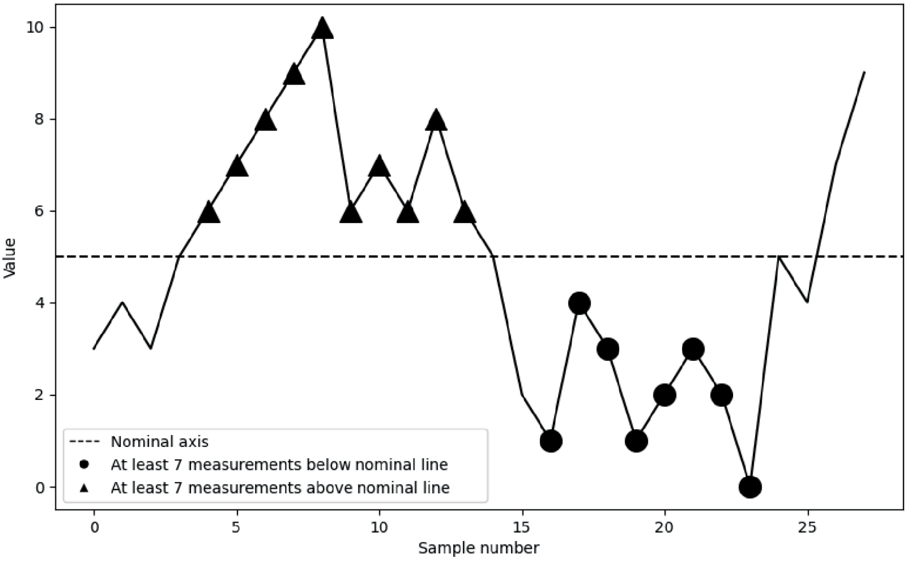


Fig. 1. *Graphical representation of method on chart*

4. MATERIAL AND METHODS

In the context of artificial intelligence and machine learning, classifiers are algorithms or models that assign objects or data to one of several classes or categories. Their main goal is to learn decision rules based on available training data to classify new unknown data. This is useful in many fields, such as image recognition, disease diagnosis, email spam filtering, sentiment analysis in social media, and more. Classifiers represent a crucial component in many artificial-intelligence-based systems. In the research on the presented problem, the following classifiers were utilized:

- AdaBoost (Adaptive Boosting) is a machine-learning algorithm that is primarily used for binary classification (but it can also be adapted for multi-class problems). The idea of AdaBoost involves sequentially training weak classifiers (known as "base classifiers") and assigning greater weights to misclassified examples in order to focus on any difficult-to-classify areas of the data.
- Decision Tree is a machine-learning algorithm that is used for both classification and regression problems. It divides a data set based on features to predict a target value.
- Gradient Boosting is a machine-learning algorithm that combines multiple weak models to create a strong predictive model. It operates sequentially, correcting the errors of previous models.
- Random Forest is an algorithm that is based on the concept of ensemble learning. It constructs multiple decision trees and combines them into one, making decisions by voting.
- Stochastic Gradient Descent (SGD) is an optimization algorithm that is used in machine learning – especially with large data sets.
- Support Vector Classifier (SVC) is a classification algorithm that uses support vector machines (SVM) to find the hyperplane that best separates data from different classes.
- Voting Classifier combines multiple different classification models and uses voting to make the final decision about the predicted class of a new sample (Cichosz, 2000).

4.1. Confusion matrix

In the conducted research, the confusion matrix was used to assess the quality (see Table 1). The confusion matrix is a tabular representation of the classification results, showing the number of correct and incorrect classifications for each class.

Table 1. *Confusion matrix*

	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

The above terms (especially in confusion matrices or performance metrics such as precision, recall, and F1 score) are commonly used to evaluate the performance of classification algorithms:

- TN (True Negatives): the number of observations that are correctly classified as negative. These are cases that are actually negative and have been classified as negative by the classifier.
- TP (True Positives): the number of observations that are correctly classified as positive. These are cases that are actually positive and have been classified as positive by the classifier.
- FP (False Positives): the number of observations that are incorrectly classified as positive. These are cases that are actually negative but have been incorrectly classified as positive by the classifier.
- FN (False Negatives): the number of observations that are incorrectly classified as negative. These are cases that are actually positive but have been incorrectly classified as negative by the classifier (Kurp, 2023).

4.2. Measures of classifier quality

By using the confusion matrix, the quality of the results was strengthened, thus expanding the scope of a further analysis. The prepared data served as modern methods for evaluating binary classifiers based on various indicators, thus allowing for a comprehensive analysis of their performance. Four key measures (precision [P], recall [R], average F1-score, and accuracy [A]) constitute a significant set of tools in this analysis.

Also known as a positive predictive value, precision focuses on how many predicted positive cases actually belong to the positive class. This is particularly useful in situations where the consequences of false positive errors are significant. In other words, precision measures how many of the instances that are classified as positive by the algorithm actually belong to the given class. Precision can take values from 0 to 1, where 1 indicates perfect precision and 0 indicates no precision. Precision is defined as the ratio of the number of correctly assigned instances of a given class to the sum of all of them:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Also known as sensitivity or the true positive rate, recall evaluates the classifier's ability to detect all actual positive cases. Recall values range from 0 to 1, where 0 indicates the model's inability to detect positive cases and 1 indicates the perfect ability to identify all positive cases. Recall is particularly important in situations where there are high risks that are associated with failing to identify positive cases:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The average F1-score (also known as the harmonic mean) is a measure that combines precision and recall, offering a single comprehensive value for an overall assessment of the balance between these two indicators. This approach is especially important in situations where classes are imbalanced, meaning that one class occurs much more frequently than the other. In practice, a high value of an average F1-score indicates an effective classifier that achieves a balance between minimizing false positives and false negatives:

$$F_1 score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

Accuracy is a general measure of classifier effectiveness that measures the ratio of the correct predictions (both positive and negative) to the total number of cases. High accuracy generally indicates the classifier's overall correct performance; however, it can be misleading in situations where classes are unevenly distributed (the classifier may be "accurate" by assigning most cases to the dominant class). In certain cases (especially when the costs of the errors differ between classes), other measures such as precision and recall may be more informative (Géron, 2020):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

5. RESULTS AND DISCUSSION

At the beginning, the data was collected; it consisted of 1040 measurements that were gathered over a specified period of time. Then, this data was prepared in Excel, where each measurement was assigned an x value (with x representing the values of seven consecutive measurements). The y value was calculated using the SPC method, where y was assigned a value of 1 (if the sequence of seven measurements stayed consistently above or below the nominal line) or 0 (if the results fluctuated above and below the nominal line). The data was then divided into 80% training data and 20% test data.

The next step involved preparing the code in Python. The data was also scaled to check for any impact on the final result for each classifier. NS represented "no scaling," while WS indicated "with scaling"; thus, 16 different classifier models were created. Libraries such as csv (for handling CSV files), NumPy (for numerical operations), AdaBoostClassifier from the scikit-learn library (for AdaBoost classification), and the confusion matrix and classification report (for evaluating the classifiers) were used. The code was developed in such a way that it could be used for each of the classifiers in an identical manner. The necessary libraries were imported, the CSV file that contained the data was opened and read, and each row was displayed. Next, the data was divided into two NumPy arrays: X, and Y. The content of Y, the length of the data, and the sum of the values were displayed to verify the correctness of the data-loading. The data was split into training and test sets, where the first 800 samples were used for the training and the rest for the testing.

Table 2. *Comparison of values for evaluating models using F1-score metric*

No.	Tested models	NS/WS	Results Dataset 1	Results Dataset 2	Results Dataset 3
1	AdaBoostClassifier	NS	0.550	0.715	0.635
2	DecisionTreeClassifier		0.520	0.860	0.935
3	GradientBoostingClassifier		0.550	0.895	0.895
4	**RandomForestClassifier**		**0.590**	**0.900**	**0.980**
5	SGDClassifier		0.445	0.450	0.435
6	SGDClassifierStratifiedShuffleSplit		0.170	0.450	0.435
7	SVC		0.175	0.835	0.930
8	VotingClassifier		0.500	0.840	0.710
9	AdaBoostClassifier	WS	0.550	0.745	0.635
10	DecisionTreeClassifier		0.515	0.875	0.960
11	GradientBoostingClassifier		0.530	0.895	0.895
12	RandomForestClassifier		0.590	0.890	0.980
13	SGDClassifier		0.170	0.450	0.295
14	SGDClassifierStratifiedShuffleSplit		0.210	0.450	0.490
15	SVC		0.175	0.835	0.930
16	VotingClassifier		0.500	0.840	0.725

Prediction was performed on the test set, and then the prediction results and actual labels were displayed in Table 2. The effectiveness of the classifiers clearly depended on the quality of the input data. The data set that was labeled Number 1 exhibited relatively poor results as compared to the other two data sets; this was mainly due to the insufficient number of errors that were identified in this data set. The classifiers did not have enough information to properly learn and assign incorrect results. In contrast, the classifiers for Datasets 2 and 3 achieved very good results, thus suggesting that the numbers of errors that were made and detected were sufficient for effective result assignment. It is worth noting that a slight improvement in results for those classifiers that were applied with data scaling could be observed; this may have been due to the smaller data range, which allowed the classifiers to better adapt to relevant classification information. The best results in all of the conducted simulations were achieved by the RandomForestClassifier. This classifier achieved values that were close to ideal in Datasets 2 and 3; and despite its lower effectiveness in this case, Dataset 1 still performed relatively well. However, it should be remembered that perfection cannot be expected in every experiment. The data that is used for experiments must contain sufficient numbers of errors for the classifiers to effectively learn the operating patterns. Under real production conditions, ensuring adequate quantities and qualities of the data can be crucial for the effective applications of classifiers. The next step after conducting these analyses could be to develop a computer application that utilizes trained classifiers to provide guidance to the controllers. Such a tool would not replace the work of the controllers entirely but could be a valuable aid in their daily work.

The main objective of this work was to prepare a training data set as well as train and evaluate the selected classifiers regarding their applicability for real-time control chart analysis. Both of the set objectives were achieved in this work, and the thesis that the known machine-learning classifiers could serve as useful methods for supporting the work of a quality controller was confirmed. During the research, it was proven that the classifiers could effectively identify the defective products on the production line in the analyzed enterprise. The use of simple classifiers along with Shewhart control charts for analyzing real production data opens up new perspectives in the field of quality control. It is also worth emphasizing that the conducted research contributes to the development of the field, where artificial intelligence becomes not only a tool that facilitates work but also an effective mechanism for eliminating human errors. From a practical perspective, defect-detection systems (though generally effective) cannot guarantee that a product is defective with absolute certainty; this is because they operate within predefined static frameworks. As such, these systems serve as valuable support tools for quality control inspectors but not as standalone solutions. However, these systems can surpass human capabilities in terms of measurement accuracy; this is due to their ability to maintain consistent stability in measurement processes.

The primary limitation when adopting these improvements lies in their cost-effectiveness – whether or not the implementation of such technology is truly necessary for a given production line. Additionally, it is crucial to consider that introducing these systems requires specialized software; this raises the question of whether the investment in such technology is justified based on the specific production needs.

The conducted research opens the door for new possibilities in the field of quality control automation, bringing benefits to statistical process control – both in terms of economic efficiency and improvements in accuracy. In detecting those patterns that signal process degradation (especially those that are defined by short seven-sample sequences), the choice of a classifier can significantly influence the predictive accuracy and model reliability. Among the basic machine-learning classifiers, random forest proved to be exceptionally effective, achieving close to 100% accuracy in identifying such degradation patterns in our experiments. This high level of performance arose from two main factors: first, the availability of a sufficient number of error samples provided a rich data set for the classifier from which to learn. When a data set includes enough instances of both normal and degraded process conditions, a random forest model can distinguish patterns that signify the onset of degradation with high precision. Each tree within the random forest ensemble captures nuances in these patterns, thus enabling robust recognition across varied instances. Second, the Random Forest algorithm's inherent structure is well-suited for this type of classification task. By creating multiple decision trees and averaging their outputs, Random Forest effectively captures complex relationships and dependencies within the data; this is especially valuable in process degradation, where subtle changes across samples might otherwise go undetected. The ensemble approach reduces the risk of overfitting on specific anomalies, leading to more generalizable insights into the degradation process. Moreover, Random Forest's ability to handle noise and outliers ensures that sporadic measurement errors do not overshadow true degradation signals. When paired with interpretability in feature importance, this resilience allows researchers and practitioners to better understand which variables most strongly indicate process decline.

6. CONCLUSIONS

All of the proposals that were suggested by this research were realized: the classifiers were well-trained, and it was found that their performance was highly satisfactory when recognizing any data anomalies that were related to process degradation; this is evidence that the implemented machine-learning models can deliver reliable support for finding defects during the production of wheel rims, thus serving as yet another tool for quality controllers.

With such promising results, the obvious next step will be to put this into practice in a real production setting. The practical and scalable application of the trained classification algorithms can be integrated into an application that is specifically designed for this purpose. Such an application would feature the real-time monitoring of production data, thus alerting operators and quality inspectors to potential defects at the times of such occurrences. By incorporating such a model-driven approach directly onto a production line, manufacturers can restrict their reliance on manual inspections to only higher detection accuracy, thereby yielding significant dividends in product quality and operational efficiency.

Further refinement could also be performed on the models in this pilot deployment itself; again, this would provide continuous improvement in anomaly-detection accuracy and adaptability for various production scenarios. This application of machine learning can become a very good avenue for bringing changes in the methods of quality control by applying data-driven insight for improving the reliability and streamlining of manufacturing.

REFERENCES

Alfardus A. & Rawat D.B. (2023). Detection of cyberattacks in controller area network (CAN) bus system using a hybrid ML approach. In: *2023 IEEE 12th International Conference on Cloud Networking (CloudNet)*, pp. 53–60. **DOI**: https://doi.org/10.1109/CloudNet59005.2023.10490074.

Benzaza L., Alfathi N. & Lyhyaoui A. (2023). Selecting key product characteristics to improve the QMS in automotive sector. In: J. Kacprzyk, M. Ezziyyani, V.E. Balas (Eds.), International Conference on Advanced Intelligent Systems for Sustainable Development. *Lecture Notes in Networks and Systems*, **714**, pp. 523–531. **DOI**: https://doi.org/10.1007/978-3-031-35245-4_48.

Cichosz P. (2000). *Systemy uczące się*. Warszawa: WNT.

Cui J., Liu W., Zhang Y., Gao C., Lu Z., Li M. & Wang F. (2021). A novel method for predicting delamination of carbon fiber reinforced plastic (CFRP) based on multi-sensor data. *Mechanical Systems and Signal Processing*, **157**, art. no. 107708. **DOI**: https://doi.org/10.1016/j.ymssp.2021.107708.

Dettú F., Formentin & Savaresi S.M. (2024). The twin-in-the-loop approach for vehicle dynamics control. *IEEE/ASME Transactions on Mechatronics*, **29**(2), pp. 1217–1228. **DOI**: https://doi.org/10.1109/TMECH.2023.3292503.

Escobar C.A., Chakraborty D., Arinez J. & Morales-Menendez R. (2021). Augmentation of body-in-white dimensional quality systems through artificial intelligence, In: *2021 IEEE International Conference on Big Data (Big Data)*, pp. 1611–1618. **DOI:** https://doi.org/10.1109/BigData52589.2021.9671610.

Fang X.-Y., Gong J.-E., Zhang F., Zhang H.-N. & Wu J.-H. (2023). Machine learning assisted materials design of high-speed railway wheel with better fatigue performance. *Engineering Fracture Mechanics*, **292**, art. no. 109586. **DOI**: https://doi.org/10.1016/j.engfracmech.2023.109586.

Géron A. (2020). *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*. Wydanie drugie. Gliwice: Helion.

Gong W., Zhao W., Wang X., Li Z., Wang Y., Zhao X., Wang Q., Wang Y., Wang L. & Chen Q. (2023). Machine learning for laser micro/nano manufacturing: Applications and prospects. *Chinese Journal of Lasers*, **50**(20), art. no. 2000001. **DOI**: https://doi.org/10.3788/CJL230827.

Habibullah K.M., Heyn H.-M., Gay G., Horkoff J., Knauss E., Borg M., Knauss A., Sivencrona H. & Li P.J. (2024). Requirements and software engineering for automotive perception systems: an interview study. *Requirements Engineering*, **29**(1), pp. 25–48. **DOI**: https://doi.org/10.1007/s00766-023-00410-1.

Helmold M. (2021). Statistical, quality and resource management tools. *Management for Professionals*, Part F473, pp. 71–79. **DOI**: https://doi.org/10.1007/978-3-030-77661-9_7.

Hofmann J., Li Z., Taphorn K., Herzen J. & Wudy K. (2024). Porosity prediction in laser-based powder bed fusion of polyamide 12 using infrared thermography and machine learning. *Additive Manufacturing*, **85**, art. no. 104176. **DOI**: https://doi.org/10.1016/j.addma.2024.104176.

Hwarng H.B. (1992). Pattern recognition on Shewhart control charts using a neural network approach. Arizona State University [PhD Thesis].

Kalyanasundaram P., Kareti V., Sambranikar M., Ss N.K. & Ranadive P. (2018). Practical approaches for detecting DoS attacks on CAN network. *SAE Technical Papers*, paper no. 0019. **DOI**: https://doi.org/10.4271/2018-01-0019.

Kidmose B. & Meng W. (2024). can-sleuth: Investigating and evaluating automotive intrusion detection datasets. In: *2024 European Interdisciplinary Cybersecurity Conference*, pp. 19–28. **DOI**: https://doi.org/10.1145/3655693.3655696.

Krummenacher G., Ong C.S., Koller S., Kobayashi S. & Buhmann J.M. (2017). Wheel defect detection with machine learning. *IEEE Transactions on Intelligent Transportation Systems*, **19**(4), pp. 1176–1187. **DOI**: https://doi.org/10.1109/TITS.2017.2720721.

Kurp F. (2023). *Sztuczna inteligencja od podstaw*. Gliwice: Helion.

Lampe B. & Meng W. (2023). can-train-and-test: A new CAN intrusion detection dataset. In: *2023 IEEE 98th Vehicular Technology Conference*. **DOI**: https://doi.org/10.1109/VTC2023-Fall60731.2023.10333756.

Lampe B. & Meng W. (2024). can-train-and-test: A curated CAN dataset for automotive intrusion detection. *Computers and Security*, **140**, art. no. 103777. **DOI**: https://doi.org/10.1016/j.cose.2024.103777.

Lee J.-H., Lee J.-H., Yun K.-S., Bae H.B., Kim S.Y., Jeong J.-H. & Kim J.-P. (2023). A study on wheel member condition recognition using machine learning (support vector machine). *Sensors*, **23**(20), art. no. 8455. **DOI**: https://doi.org/10.3390/s23208455.

Lestyán S., Acs G., Biczók G. & Szalay, Z. (2019). Extracting vehicle sensor signals from CAN logs for driver re-identification. In: *Proceedings of the 5th International Conference on Information Systems Security and Privacy*, pp. 136–145. **DOI**: https://doi.org/10.5220/0007389501360145.

Lokman S.-F., Othman A.T. & Abu-Bakar M.-H. (2019). Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review. *EURASIP Journal on Wireless Communications and Networking*, **2019**(1). **DOI**: https://doi.org/10.1186/s13638-019-1484-3.

Malindzakova M., Čulková K. & Trpčevská J. (2023). Shewhart control charts implementation for quality and production management. *Processes*, **11**(4), art. no. 1246. **DOI**: https://doi.org/10.3390/pr11041246.

Minawi O., Whelan J., Almehmadi A. & El-Khatib K. (2020). Machine learning-based intrusion detection system for controller area networks. In: *DIVANet '20: Proceedings of the 10th ACM Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, pp. 41–47. **DOI**: https://doi.org/10.1145/3416014.3424581.

Mjimer I., Aoula E.-S. & Achouyab E.H. (2023). Contribution of machine learning in continuous improvement processes. *Journal of Quality in Maintenance Engineering*, **29**(2), pp. 553–567. **DOI**: https://doi.org/10.1108/JQME-03-2022-0019.

Moser M., Kipping S., Higuchi K. & Hirayama H. (2021). Machine-learned emission model for diesel exhaust on-board diagnostics and data flow processor as enabler. *SAE Technical Papers*, paper no. 5108. **DOI**: https://doi.org/10.4271/2021-01-5108.

Oh Y., Ransikarbum K., Busogi M., Kwon D. & Kim N. (2019). Adaptive SVM-based real-time quality assessment for primer-sealer dispensing process of sunroof assembly line. *Reliability Engineering and System Safety*, 184, pp. 202–212. **DOI**: https://doi.org/10.1016/j.ress.2018.03.020.

Park S.-H. & Baek Y. (2023). CADS: Cyber-attack detection system using decision tree for automotive cyber-physical systems. In: *2023 International Conference on Electrical, Communication and Computer Engineering*. **DOI**: https://doi.org/10.1109/ICECCE61019.2023.10442984.

Paturi U.M.R., Palakurthy S.T., Cheruku S., Vidhya Darshini B. & Reddy N.S. (2023). Role of machine learning in additive manufacturing of titanium alloys: a review. *Archives of Computational Methods in Engineering*, **30**(8), pp. 5053–5069. **DOI**: https://doi.org/10.1007/s11831-023-09969-y.

Rameshkumar K., Mouli D. & Shivith K. (2021). Machine learning models for predicting grinding wheel conditions using acoustic emission features. *SAE International Journal of Materials and Manufacturing*, **14**(4), pp. 387–406. **DOI**: https://doi.org/10.4271/05-14-04-0026.

Robinson J.C., Sherman K., Price D.W. & Rathert J. (2020). Inline Part Average Testing (I-PAT) for automotive die reliability. In: O. Adan, J.C. Robinson (Eds.), *Metrology, Inspection, and Process Control for Microlitography XXXIV. Proceedings of SPIE*, **11325**, art. no. 113250D. **DOI**: https://doi.org/10.1117/12.2551539.

Sharmin S., Mansor H., Abdul Kadir A.F. & Aziz N.A. (2022). Using streaming data algorithm for intrusion detection on the vehicular controller area network. *Communications in Computer and Information Science*, 1557, pp. 131–144. **DOI**: https://doi.org/10.1007/978-981-19-0468-4_10.

Shewhart M. (1992). Interpreting statistical process control (SPC) charts using machine learning and expert system techniques. In: *Proceedings of the IEEE 1992 National Aerospace and Electronics Conference@m_ NAECON* 1992, Dayton, OH, USA, **3**, pp. 1001–1006. **DOI**: https://doi.org/10.1109/NAECON.1992.220472.

Shewhart W. (1926). Quality control charts. *Bell System Technical Journal*, **5**(4), pp. 593–603. **DOI**: https://doi.org/10.1002/j.1538-7305.1926.tb00125.x.

Shi W., Alawieh M.B., Li X., Yu H., Arechiga N. Tomatsu N. (2016). Efficient statistical validation of machine learning systems for autonomous driving. In: *2016 IEEE/ACM International Conference on Computer-Aided Design,* 7–10 November 2016, Austin, TX, USA. **DOI**: https://doi.org/10.1145/2966986.2980077.

Staf E. & McKelvey T. (2018). Introducing compressed mixture models for predicting long-lasting brake events. *IFAC-PapersOnLine*, **51**(31), pp. 840–845. **DOI**: https://doi.org/10.1016/j.ifacol.2018.10.115.

Staněk R., Kerepecký T., Novozámský A., Šroubek F., Zitová B. & Flusser J. (2023). Real-time wheel detection and rim classification in automotive production. In: *2023 IEEE International Conference on Image Processing (ICIP)*, pp. 1410–1414.

Tran P.H., Ahmadi Nadi A., Nguyen T.H., Tran K.D. & Tran K.P. (2022). Application of machine learning in statistical process control charts: A survey and perspective. In: K.P. Tran (Ed.) *Control Charts and Machine Learning for Anomaly Detection in Manufacturing. Springer Series in Reliability Engineering*, pp. 7–42. **DOI**: https://doi.org/10.1007/978-3-030-83819-5_2.

True J., Xi C., Jessurun N., Ahi K. & Asadizanjani N. (2021). Review of THz-based semiconductor assurance. *Optical Engineering*, **60**(6), art. no. 060901. **DOI**: https://doi.org/10.1117/1.OE.60.6.060901.

Tsenev V. & Ivanova M. (2022). Statistical and machine learning approach for evaluation of control systems for automatic production lines. *Bulletin of Electrical Engineering and Informatics*, **11**(5), pp. 2527–2536. **DOI**: https://doi.org/10.11591/eei.v11i5.3664.

Xu X., Wang G., Xuan S., Shan Y., Yang H. & Yao X. (2024). Digital twin and cross-scale mechanical interaction for fabric rubber composites considering model uncertainties. *Composites Science and Technology*, **248**, art. no. 110431. **DOI**: https://doi.org/10.1016/j.compscitech.2024.110431.

Zhou A., Liu X., Liu Q., Zhu G. & Zhang S. (2023). Progress of in-process monitoring techniques for selective laser melting. *China Surface Engineering*, **36**(4), pp. 36–50. **DOI**: https://doi.org/10.11933/j.issn.1007-9289.20221212001.

Decision Making in Manufacturing and Services
Vol. 18 • 2024 • pp. 99–117

Modeling Market Value of Product Based on New Technologies – Preliminary Tool Concept

Paweł Filipowicz*

Abstract. Changes in the way market advantage is achieved and the emphasis on value creation that has been introduced by technological change are forcing companies to seek new and effective management tools when it comes to innovative activities; hence, the discussion and presentation of the conceptualization of such a decision-making model. Its basic premise is to take the technology life cycle as a determinant of the market diffusion of a product based on the use of new technologies. Consequently, a decision matrix is developed – one dimension of which is innovation measured by the dynamics of technical debt and the other by customer-perceived value. The product as an analytical object is considered as the sum of the utility functions while allowing for each utility function to originate from the use of a different technology (which is in an adequate phase of its life cycle). After presenting the theoretical model, an example of an analysis of a fictitious product is given, which is a device for cooling and printing food substances. The presented analysis example shows the practical possibilities of using the developed decision-making tool in the areas of the technological and marketing activities of the enterprise and, in particular, the design of new products or new utility functions.

Keywords: product conceptualization, perceived value, innovation, technical debt, new technology

Mathematics Subject Classification: 60K25

JEL Classification: C53

Submitted: July 15, 2024

Revised: October 21, 2024

* AGH University of Krakow, Faculty of Management, Poland, e-mail: pawelfil@agh.edu.pl.

DOI: https://doi.org/10.7494/dmms.2024.18.6641

1. INTRODUCTION

Understanding the customer impact on a company's product-value-strategy-formation process means also perceiving the relationship between product innovativeness and its customer. Understanding the positive impact of this relationship on the commercialization of innovation should provide reason for developing a new insight that cements a new technology-based conceptualization of new product model tools. Innovative customers have more experience in change adoption in product development, and they lead the innovation process by predetermining a company's value creation. Innovation-based strategies pose a challenge for the creation of new markets based on innovation- and differentiation-based products or services and can help deciders develop new innovative products in an unconventional ways. At the same time, developing and introducing strongly differentiated products can result in a certain risk of commercialization failure that is directly related to the unrecognition of the value by the customer. This phenomenon highlights the significance of client involvement in new technology commercial usage, which can form the basis of new management that evaluates the potential value that is attributable by the new users to an analyzed innovation. Hence, the presentation of such a decision model is given, which allows for a visualization of the level of innovation of a new product based on its utility functions and a presentation of the resulting relevant customer-perceived values.

2. CONCEPTUAL BASIS OF MODEL AND ITS PARAMETRIZATION

Technological innovation implementation implies changes in a company's strategic thinking –particularly at the level of fundamental market-behavior models. In the case of the new technology's company sector, the model often used is the technology Foster s-curve (Foster, 1986). This curve shows the progress of a new technology as the source of the R&D future development. The first stage of this development is very slow, which reflects the very weak diffusion of the concept being developed. Furthermore, exceeding a certain level of knowledge resources results in a significant increase in the diffusions of developed solutions; this increase is exponential. After this, stagnation can be observed – this is the state of the maturity of the developed technology. An observation of the theoretical course of the curve allows us to conclude that the performance of the proposed solution decreases following the inflection point of the function; this point is characteristic of the described phenomenon of technological development (Fig. 1). According to Asthana (1995), it is also possible to use this curve to describe certain principles of technology management. He stated that the primary barrier to the adaptation of a new technology is unfamiliarity by the wider user base, which translates into uncertainty of acceptance by the market. The described new-technology-diffusion function is currently finding numerous applications and interpretations; according to Sood and Tellis (2005), this supports the need for some standardization of this model. As part of a holistic view of this model, one can propose a view of it based on defining the curve as a life-cycle model that is based on three phases. The first phase is the introduction, at which time a new technological standard is put into use; its propagation is slow because it remains unknown, and, although new, it does not attract users. Another possible interpretation of this state of affairs is that there are too few practical applications of new technological standards, resulting in a small number of new products that are based on technological innovation.

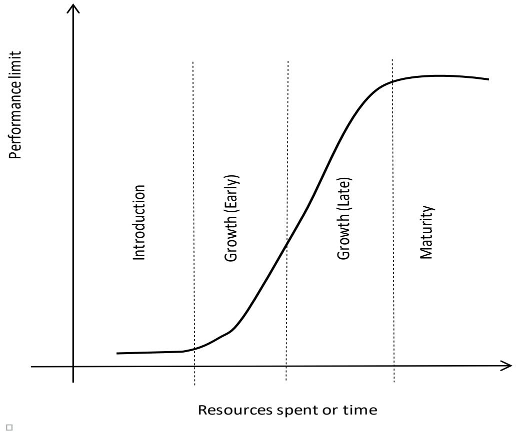


Fig. 1. *Technology s-curve*
Source: based on Bowden (2004)

The second stage is the (early and late) growth stage; a rapid propagation of the innovative technology (Filipowicz, 2015). This is the phase of the dominant standard emergence, which determines the specifics of market-dominating-product standards and consumer preferences. This progress leads to increased product sales based on the use of the new technology, which increases revenues and profits and offers further support for research and performance amelioration. The third stage is the maturity phase of slow technology diffusion and then market saturation. This maturation stage is due to a reduction of innovation activities because of low-cost imitations, competitive offers, and a growing lack of customer interest. Pearce and Robinson (1994) highlighted that, in a dynamic market, even a small or relatively weak business is often able to find an interesting market niche.

Thus, innovation-strategy formulation should be associated with technical performance, which can be measured both by the capacity to deliver customer value and the creation of an adaptable product for tomorrow. According to Highsmith (2009), the idea of exposing technical debt – improving the ability to adapt – is an important part of the innovation process. It appears that the use of technical debt could be proposed as the new technology-based product-innovativeness measure. The aforementioned exponential growth in the diffusion of new technology implementations is, in such a situation, a reflection of the increase in the value to the customer of the new implementations of the developed solution. An important issue for the effectiveness of the innovation process is the understanding that the increase in the value of the innovation as perceived by the user is the result of a compromise between the cost of its production and the value for the customer. This implies the need for the continuous optimization of value for the customer, which may translate in specific cases into a decrease in the value that is acquired by the company (Chen & Quester, 2009). This process results in the growing importance of the concept of the value-management process, which emphasizes the importance of the contribution of customer-perceived value to shareholder-value maximization (Porter, 1985). Therefore, customer-perceived value (CPV) can be defined as the ratio between the level of customer satisfaction and the perceived value of a product as well as satisfaction with the price that was paid for this product. A company thus creates added value for the customer when its products or services hold greater value to the customer than those that its competitors have offered under similar market conditions. CPV can be measured by surveying customer opinions or be calculated as the ratio of a company's performance relative to its utilities and quality (Blut et al., 2024). The deep individualization of this approach (suggested by the authors) can become a key factor in the effectiveness of the commercialization of new technologies. The often-indicated strong relationship between customer participation in the development of the innovation process and the parallel development of commercialization is expected to contribute to the significant reduction in the risk of the market failure of the venture (Ritter & Walter, 2012). Hence, the suggested connection of technology development with the conceptualization of customer value results in a contradiction with respect to the classic linear-innovation process. In effect, this leads to the possible identification of a certain congruence between the value that is perceived by the customer and the level of the innovation of the technology of a given enterprise. This congruence results in the possibility of determining the level of the innovation of the proposed technology and its perceived value, which in turn determines the level of financing that is needed for its development (Radford & Sridhar, 2005). The concept was confirmed by Cunningham (1993), and it results in the practices of enterprises with great importance being placed on the timeliness of the operations in the areas of the developed technologies. Thus, the increasing emphasis on maximizing customer value leads to an increase in the number of utility functions of an offered product. This results in the need for changes in the management and organization of an enterprise that applies such a concept (Nord et al., 2012).

3. PRESENTATION OF CONCEIVED MODEL

A competitive advantage based on new technology is not always a key factor in a company's success; this is due to the too-rapid diffusion of this technology and low-cost international imitation. Under such conditions, some companies attempt to expand the value proposition based on technological innovation and develop various commercial proposals; this is the main reason for giving innovative importance to not only the development of a new product but also the creation of complementary elements of a commercial offer (including new services) to meet market expectations. This is why the importance of the previously mentioned descriptive parameter – customer-perceived value – is so important. When analyzing the possibility of expanding an offer, it may be important to identify the set of utility functions of an innovative product that are available on the market during a given time period (Ho et al., 2014). The identified set of utility functions will, thus, form the basis of the developed product strategies based on the use of the new technology. An important factor that shapes these strategies for developing attrition functions will be the use of their evaluations by potential users, which can become part of a standard innovation development. Such an approach to product development based on a new technology can be a very effective dimension for assessing market opportunities for new applications of the technology that is being developed; therefore, this is crucial when it comes to the dynamics of diffusion of especially radically innovative products or innovations. It is important in this context to consider the organizational dimension of such an approach to the diffusion of innovations; i.e., the need to take the formation of customer value into account in the process of financially evaluating a new technology (Ho, 2011).

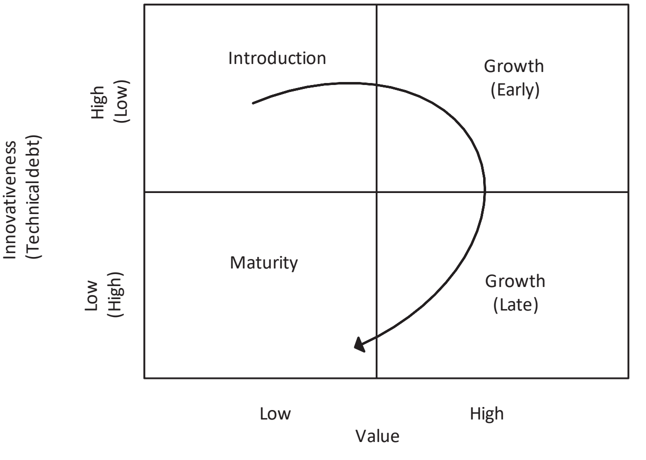


Fig. 2. *Presented tool based on technology life-cycle model*
Source: Tool-concept proposition based on s-curve holistic interpretation (Filipowicz, 2014)

The easiest way to understand the issues that were described above will be to use the developed tool (Fig. 2). The application of this model requires estimating the values of the levels of innovation and customer value that are associated with each product based on a collection of empirical data (or rather with the utility function that is offered). This tool is a very important decision factor for determining the involvement of potential users in the development of a new technology that forms the basis of the innovative products that are offered. The main purpose of its application is to anticipate the optimal value of the innovative technology that is used in the new product; this, in turn, allows the company to conduct an effective cost policy and is the basis for co-production. As a consequence, there is a further need to concretize and operationalize the presented concept.

The relevance of analyzing the current needs of customers, therefore, becomes indisputable only if the organization intends to reconfigure in order to introduce technologically innovative products; it is then important to take the fact that today's customer needs are only a base for future customer needs into account. Hence, it becomes very important when adopting a prosumption-based perspective of the innovation process to be able to anticipate that the desired new utility functions are what will then guide the development of the new technologies needed by the company (which can be sources for creating new markets) (Keinonen & Takala, 2006). At the operational and strategic levels, it therefore becomes important for the enterprise to be able to configure and evaluate new product portfolios based on innovations and the resulting need for a possible change in the structure of the designed products, thus allowing it to take the possibility of implementing the new technologies that have been developed or acquired by the organization into account. Other things to be taken into account are the possibility of analyzing the level of effectiveness of the integration of all types of innovations into the portfolio of the owned products or technologies. (Fuchs & Golenhofen, 2019); hence, the high relevance of the nature of the interactions that take place with the customers or users (who henceforth constitute the desired feedback for the company). The earlier that the nature of the reaction of one's potential customers is known (even at the level of product ideation), the greater the importance of these comments for constructing a market advantage (Sánchez-González & Herrera, 2015). When a company's offerings are subjected to customer evaluations, innovation is usually identified as an important component of the value-creation process. Accordingly, the development potential of a company is considered to be dependent on the number of new products that are introduced to the market (Schultz et al., 2013). This becomes relatively important in the case of a very dynamic environment when companies are looking for different forms of configurations of their existing portfolios of derogatory products that take the new technologies into account (Mul & Di Benedetto, 2011). The ability to analyze current customer preferences is becoming increasingly important – especially when a company wants to commercialize innovative products. In such a situation, the currently declared needs are only the basis for sealing their new forms in the future. Then, it becomes necessary to apply the prosumption approach to the course of the innovation process to search even at the strategic level of the organization – for new technological innovation based utility functions possible to incorporate in the product according the potential user wants – one that is expected by the emerging market (Keinonen & Takala, 2006). This need to analyze current customer needs becomes clear when a company recomposes its offerings to commercialize innovative products. Hence, it follows that interactions with customers are a source of feedback; it is important to develop them as early as possible at the conceptual stage of designing a new product; after all, these interactions will be an important factor in the process of shaping the company's market advantage (Sánchez-González & Herrera, 2015). According to Moore (1965), the technology propagation curve (which describes the value of an innovative product due to the effective communication with the customer) is the primary factor in the effectiveness of technological innovation diffusion (Millier, 2011). The adoption of the exponential function as a model function for the diffusion of technological innovation highlights the importance of the speed of a product's launch and, thus, replaces the classical understanding of product development (which focuses primarily on its quality) (McNally et al., 2014).

In addition, the assumption that the first phase of innovation diffusion is the most significant source of acquired value is true when the innovativeness of a proposed commercial offer is the same as the lack of familiarity with the new product, thus signifying its originality and lack of any references. This can also mean that the decline in the innovativeness of a newly introduced product occurs as the user becomes familiar with it and knowledge of it increases. The longer a product is used, the more the user's knowledge of its flaws and shortcomings grows; this, in turn, promotes interactions with the company and the reporting of needed modifications to better meet customer needs (which is the main reason for creating technical debt). The changes that are made by the enterprise reduce the value of this debt; this leads to the conclusion that the level of innovation of a market product is inversely proportional to its technical debt. This approach becomes very important when, for strategic or financial reasons, a company wants to determine the market value of the technologies that it introduces or uses (Artmann, 2009).

The role of customers in the innovation process is already widely recognized and often used to maximize the values of products. The fact of being a leader in a new market provides a unique opportunity for creating an autonomous pricing policy that fits into the strategy of value creation – as long as one's customers perceive the uniqueness of the proposed innovation in satisfying their needs (Kumar & Phrommathed, 2005). Commercialization based on innovation represents a unique opportunity for a company to create a new market and its further development if the product is modified. Hence, the great interest in applying customer relationship management as a kind of connection between the customer and the enterprise mainly serves the development of any valuable commercial applications of the introduced technologies (Huang & Wang, 2013). These new applications become a potential source of new utility functions or even new product meanings; this approach gives even more weight to the roles of customers and their experiences of using new utility functions, and it allows for redefining product utility as the sum of the utility functions (Eversheim, 2009). Thus, linking customer value to the process of developing new technologies becomes a key factor in the market success of new products and is essential if a company is thinking about growth; it also influences the decision-making rationalization of the led NPD, reducing risk and conditioning the consumer acceptance of market applications of the new technologies.

In some situations, the configuration of utility functions – the new technologies that are used – can change the application of the product and lead to the emergence of new users. Then, the principle according to which more complex and extensive modularity translates into a higher level of individualization and differentiation, which can translate into an increase in interest (Jensen et al., 2014). This approach to value creation through utility function analysis provides new opportunities for optimizing product policy – even at the level of the structure of a single product; this significantly reduces the risk of failure during the first stages of NPD. It becomes evident that the possibility of different product applications in the value-creation process is based on different ways of perceiving them by reflecting customer preferences for new products. The integration of the customer's optics in the process of designing a product's utility functions should be implemented with an understanding of not only the hedonistic approach but also as a compromise of both (Verhagen et al., 2010). The implementation of these assumptions should cover all phases of the life cycle and implies a balance between the development of the utility functions and the suggested way in which they are perceived by the user (depending on the situation being once hedonistic and once utilitarian). In addition, the process of utility-function development represents an opportunity for the enterprise to optimize the value offered in a manner that is consistent with the degree of the mastery of the technology that is used (Ha & Park, 2013). The use of such an application makes the enterprise ready to perceive the commercial offer not only as a wide selection of products but rather as an offered collection of utility functions. The described hedonic/utilitarian dualism can then serve as a tool for the value-creation-mapping process and become a starting point for the conceptualization of a new product (particularly, when it comes to designing the core functions of the product) (Pieper, 2019). Such an approach is, therefore, in line with the essence of the innovation process and the reason why it will be beneficial to use customer-designed functional characteristics and increase the market-utility value or create links between product performance and user comfort that are new in nature (Townsend et al., 2013).

Hence appears the proposal of the concept of functional value in relation to the possible uses of the product, which can be defined as the accumulation of the utility functions that represent benefits for the customer (Fig. 3). The assumption that the customer is at the center of the value-creation process also allows the customer to be involved in the product-design process by creating the configurations of the utility functions of the offered device that are most interesting to him/her (Liu et al., 2020). An opportunity that is formulated in this manner is also attractive from the perspective of innovation product development (IPD) and can provide an attractive basis for organizations to conceptualize the structures of their products based on possible user perceptions (Moon et al., 2015). Therefore, an in-depth understanding of customer preferences for the value of product applications can significantly reduce the risk of market failure and should be treated as a starting point for a company's future product and innovation portfolio. Another dimension of such a perception of innovation-product development is the ability to understand and gain the customer's perception of product innovation through different compositions of the possible applications of the designed device (Tang et al., 2018).

Therefore, it will be very helpful when designing a product to know the utility functions that are desired by users and, thus, determine the market values that are used and know how much the company is able to respond to these needs given the level of investment being made or the potential to develop technological capabilities (Kumar & Puneet, 2019). The presented perspective on the development of innovative products provides an opportunity to present the company's offer as a set of offered use functions, which allows it to present a map of product functions that are offered or developed based on new technologies. In this case, a set of use functions is represented by fu(t,v), where t is the technology that is applied to obtain the use function, and v is the customer-perceived value of the use function. When fu(t,v) = 0, this means that a specific use function does not exist or has yet to be invented (Fig. 3). The possibility of writing down the product structure in this way results in the ability to analyze use functions in terms of the company's current production capabilities as well as any projected changes in user preferences or possible technological changes. In extreme situations, conducting an analysis can be used to develop completely new production processes or a new formulation of the value offered to the customer. The ability to design the structures of new products based on the map of utility functions greatly supports the decision-making process of the enterprise, which can also provide the basis for cost optimization for the entire product family or the basis for outsourcing (Fan et al., 2015).

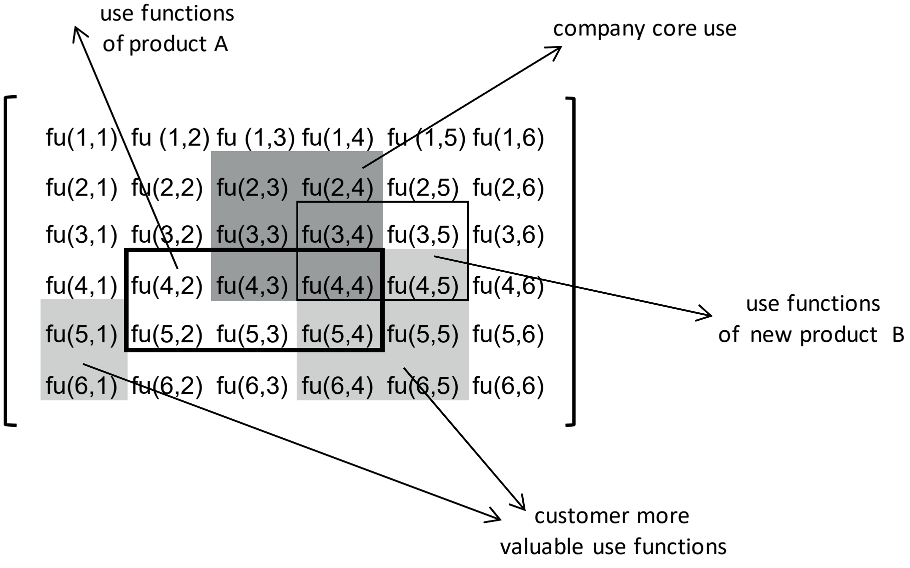


Fig. 3. *Utility functions offered by enterprise and distribution of their customer preference – analytical matrix*
Source: (Filipowicz, 2019)

4. APPLICATION EXAMPLES AND DISCUSSION

At the beginning of commercialization, an innovative product based on the use of a new technology is characterized by a low value of technical debt; when the customer does not know the product's capabilities and limitations well, any new utility functions often remain unknown to him/her. Due to the limited availability of this product and its uniqueness, its value is perceived as being high. Its utility functions are unique, hardly available – simply rare. Due to its new mix of utility functions, an innovative product often has a completely new meaning, which translates into a high innovative market value. The assumption that a highly innovative product's technical debt is low or equal to zero is derived directly from the assumptions of Jim Highsmith. As the product becomes more and more diffuse, its innovativeness decreases; this translates into an increase in the value of its technical debt, as more and more users or customers get to know product's properties and applications. This affects the wider area of the proposed modifications and improvements. The concept of the product then becomes more and more widely imitated, which translates into a decrease in its perceived value. A distinctly unique meaningfulness becomes standard, and its adaptations become more and more frequent; this is the result of imitation by other companies. Thus, these phenomena contribute to an increase in a product's technical debt or a decrease in its innovation (and its value on the market). In such a view, product innovation is a reflection of the low degree of the diffusion of the new meaningfulness and, thus, the high value of its perception. To illustrate the model, the example of an enterprise that produces household appliances is presented below. The company is considering the launch of a product that is a refrigerator with the ability to 3D print organic food – "Frizprint." The device's concept is based on the utility functions that are presented in Table 1. As a result, each utility function will be characterized by two variables – the value of its technical debt (describing the level of technological innovation), and its value (reflecting the market attractiveness); this allows for an estimation of the market price of a product that offers this type of utility. The presented perspective provides an opportunity to draw analogous conclusions that cover the entirety of the designed product. Based on the presented assumptions, the listed device functions can, therefore, be divided into core and innovative. Core functions are those functions that are sub-activities of the core business in the sense of organization theory and the resource-based view of the firm and are derived from the past utility functions of an innovative nature that have spread and are now standard activities of the company. These are characterized by a low level of innovation and, therefore, high technical debt and low perceived value for the customer. However, they represent basic utility functions in the proposed new product's concept, thus building a positive brand image of the manufacturer at all times (Mulder-Nijkamp, 2020). The innovative functions of a device are functions whose inclusion in the product concept is based on the application of new technologies; these functions have zero technical debt and high value for the user. The innovative functions of a designed product represent its technological innovation; for the potential user, they represent high value and are often the reason for buying this device. They are the ones that create a "not-seen-before" effect. The core functions of a given device are currently implemented based on in-house technologies or externally sourced technologies on an out-sourcing basis, remaining perfectly mastered in the production process of the enterprise. On the other hand, the innovative functions are based on completely new technologies, the application of which is the result of the carried out innovation process, and the implementation is reflected in zero or very low levels of technological debt. Such a value of the technological debt of a given innovative utility function can be interpreted as a current barrier against imitation by other companies. Another situation is possible when a given utility function is innovative and characterized by a zero level of technical debt of new use of externally acquired technology; this is due to the lack of organizational competence to enable its own innovation process for a given technological innovation.

Table 1. *Classification of presented product utility functions according to innovativeness and customer-perceived value: CUF – core utility function; IUF – innovative utility function*

No.	Utility function	Description	Innovation category	CPV
1	F(u1)	Cooling of food products	CUF	LOW
2	F(u2)	Freezing of food products	CUF	LOW
3	F(u3)	Decontamination of contents	CUF	LOW
4	F(u4)	3-D printer suitable for printing organic substances	IUF	HIGH
5	F(u5)	Storage of food substances for printing	IUF	LOW
6	F(u6)	Control to synthesize food products according to preset recipes	IUF	HIGH
7	F(u7)	Weight of stored products and printed products	CUF	LOW
8	F(u8)	Grinding of food ingredients feeding printer	IUF	HIGH
9	F(u9)	Information about quantity and status of food resources, composition, caloric value	CUF	HIGH
10	F(u10)	Internal and external temperature measurement	CUF	LOW
11	F(u11)	Reading of barcodes enabling acquisition of information on expiration dates	CUf	HIGH
12	F(u12)	Wi-fi communication of measuring and control device with user's phone	CUF	HIGH
13	F(u13)	Optical interface – touchscreen with information about stored, printed products	CUF	HIGH
14	F(u14)	Digital notice board (via e_communicator)	CUF	HIGH

The customer's perceived value of a given utility function is determined by the degree to which he/she perceives its attractiveness as compared to other products that are known to him/her or his/her own needs or expectations that are caused by his/her lifestyle or level of perceived utility satisfaction, with the product offering the utility function in question. Value is determined by the customer or user outside the company; this is often the result of an intensive process of communication with the customer or the result of the presumption process being carried out. The above-described utility functions of the designed device are placed in a matrix according to the value of their technical debts and perceived values by the customer. Under real conditions, both coordinates have their specific values expressed in monetary units. The function outline that is presented here provides a tool to help interpret the characteristics of the designed product. The matrix shows the distribution of utility functions according to the values of their technical debts and their customer-perceived values. Thus, the chart reflects the state of the product as is perceived by each utility function as well as according to the phase of the technology life cycle (as assumed in the model described above). Three functions are in the introduction phase; that is, brand new technologies have been used to make them available to users. For the purposes of the example, it is assumed that they are being developed based on internal funding sources. Three more are in the early-growth phase, and another three are in the late-development phase. These last six utility functions are highly perceived by customers. There are five functions in the maturity phase, except that all of them are core functions or functions that are related to the endogenous development of the company (whose main business is the production of home appliances and, therefore, its own refrigerator production). In the late-growth and maturity phases there are eight utility functions with low innovation potentials that, therefore, remain easy to imitate. The number of utility functions in the introduction and early-development phases may indicate a very cautious initiative of the enterprise in question to gain a market leadership position in the area of refrigeration and food-substance 3D-printing equipment.

There is some concern, however, from the perspective of the life-cycle phases of the used technologies regarding the very concept of such a designed appliance based on the assumption of a combo; that is, a multifunctional product (Fig. 4). The core functions are rather perceived by customers as being not very innovative; this is not a problem in the case of the refrigeration equipment, but their value in relation to the designed equipment definitely has a rather low perceived value on average, which may indicate a high level of competitiveness in the sector. It should also be taken into account that it is the sum of the perceived values of the utility functions (that is, the quantity) that can affect the price of the product. On the other hand, the sizes of the technical debts of the individual utility functions will similarly determine the subsequent needs for financing research and development activities in order to adapt the product to the possible increase in the customer's adaptation needs. The increase in these needs will grow with the increase in the technical debt, indicating a decrease in the innovativeness of the function or product and resulting from the increasing knowledge of the customers as to the possible utility of the product or its utility functions.

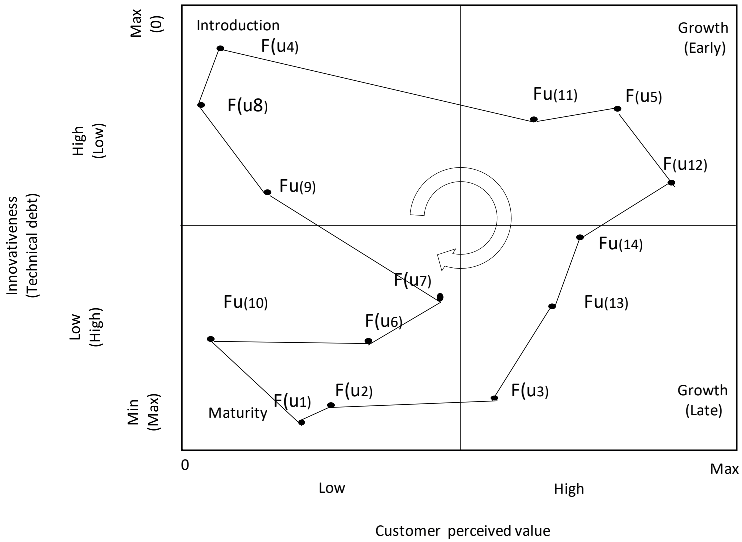


Fig. 4. *Mapping of innovative product-utility functions – case of fictitious product "Frizprint"*

Thus, the presented tool provides an opportunity to analyze the distribution of the individual utility functions in relation to the levels of the innovation of the used technologies, which effectively allows for determining the degree of the differentiation of a given product; in the case of mapping a greater number of derived products, there is the possibility of determining the levels of the differentiations of the entire range of products or product offerings. The example that is presented here refers to the conceptualization of an innovative product as the first step in the commercialization process. It is easy to imagine simulating products for the future by taking past feasible technologies into account.

On the other hand, a given enterprise's existing skills in the use of available technologies will enable the design of future utility functions and the study of the customer reactions to the new product utility mix. As a consequence of such simulations, an enterprise can profile not only current but also future customer needs so that the values of their innovative products are optimal for the enterprise at a given moment (as well as in future periods). Therefore, this tool can be used to evaluate the current range of products and to design future ones.

Undoubtedly, any further development of the presented model should focus on the possibility of using it to design new product meaningfulness through the use of innovative technologies as bases for new utility functions, which will be a very interesting tool for the conceptualization and design of new products – especially those that are based on the use of new technologies. Another direction of urgency is the full verification of the model based on real data that is derived from enterprises. Due to the high sensitivity of data that is related to the use of new technologies (and especially in relation to the value of technical debt), it is not easy to disclose this data in a publication. Another problem that occurs is the lack of data on the technical debt of a product; this applies mainly to companies outside the IT industry. Determining its value in relation to the utility function of the product and the technology that is used is very difficult and labor-intensive; it also requires a surveyed company to recognize the structure of its direct costs. Additionally, this data is treated by companies as being very sensitive – especially in the cases of innovative technologies. When it comes to measuring the value of a product as perceived by the customer, it is possible to visualize the innovative product or use computer graphics thanks to the use of artificial intelligence. Acquiring this data does not seem to be difficult; however, it may be more complicated to study the perceived values of new meaningful innovative products.

The designed tool can be used for visualizations in the case of a comparative analysis of a designed product with a competitive product on the basis of benchmarking of offered utility functions. Thanks to the results of such an analysis, it is possible to identify gaps in the range of utility functions that are offered and, therefore, potential directions for the development of the underlying technology. Based on such an analysis, it is also possible to determine potential directions of out-sourcing when it comes to the desired technologies. It is also possible to identify potential directions of out-sourcing in terms of the desired technologies, the possible increase in the value offered, and the possible directions of the development of marketing communications with potential customers. It is possible to use the model to present the company's entire portfolio, including all of the products that are offered as well as their functionalities while taking their levels of innovation or perceived value into account. In the case of perceived value, it is possible to visualize each product, the entire portfolio, and all of the utility functions that are offered and, thus, the technological level of the product offering.

Thanks to this, we also have the possibility of forecasting the financial results to be achieved in time as well as the financial outlays that are needed for this and the time that is required to implement any technological changes.

Another important direction of the possible use of the designed matrix is a detailed analysis of manufacturing capabilities and, therefore, the technological potential of the company in the field of their designed products in relation to the needs that are declared by the customer. In fact, it is possible to create a map of a product that is desired by the customer (or simply invented by him/her). This would make it possible to carry out an analysis of the potential deviations between the company's current technical capabilities and those that are expected by the customer at the level of the desirable utility functions and the product's transferable value (which is, of course, linked to the shaping of the prices of the designed or manufactured products and their changes in the future). The change in the level of the technical debt that accompanies an analysis of customer needs therefore makes it possible to predict specific technological needs or simply determine changes over time in the need for new technologies as well as in the view of changing customer needs.

Taking the life-cycle phase of the various functions of the products into account, this gives an indication of the dynamics of the development time of the various technologies that are already used by the company and allows for the identifications of specific new technologies in terms of their uses for the development of existing functions (or even the creations of concepts for new functions and their testing by potential users) by using the presented tool as a basis for the simulation and drilling of new innovative products or the development of existing ones. On the basis of technical predictions, the model allows one to determine the technical changes in any offered products in the future, which allows one to better adapt production offers to future needs or market requirements.

The suggested possibility of using matrix calculus allows for a complete mathematical notation of the model of the designed product, which can provide a basis for time-based optimization, while changes in the technical debt of the technology determine the prospects for changes in its utility functions. Not without significance is the fact that the mathematical notation also makes it possible to determine the new meaningfulness of products based on new technologies. A simultaneous analysis of a company's entire product portfolio on the basis of the proposed method creates the possibility of simulating changes in its technological and production potential depending on changes in the market characteristics or the emergence of completely new technologies. The identified problem of quantifying the presented model is the development of a complete mathematical notation using matrix calculus, which allows for the presentation of all possible changes in the evolution of the individual utility functions of the analyzed product.

The presented concept of the tool refers to the logic of portfolio-analysis methods (in particular, to the BCG matrix), which is a determinant all of the time – not only of those activities at the operational level in the marketing area. An extension of this method is the adoption of the technology life-cycle model as a determinant of product development, which is a significant modification in the sense of current business conditions. Also, the introduction of the parameterization of the matrix based on the technical debt and the perceived value of a product is a very important modification that corresponds to the current dynamics of the market processes. The use of technical debt in relation to the technology that is used as the basis of the utility function is an interesting approach to such an ambiguous issue of the level of innovation of the designed product. In particular, it refers to the assumptions of agile management and reflects the problems of IT product development. The second parameter, which is an important novelty, is the use of the value that is perceived by the customer; in particular, for the evaluation of the product utility function in this case and, therefore, the interchangeability of the estimeted value of the new product. An important issue to be developed in the model is to take the impact of the industrial design on the value of the product into account as well as the interactions that occur between it and the various utility functions of the product. It becomes an important research problem in this situation to analyze and periodically characterize the relationship between product design and its innovativeness – especially in terms of the use of new technologies. The conceptualization of the decision-making tool is an attractive proposition in the area of decisions at the operational level, enabling not only the effective designs of innovative products but also giving the possibility of assessing their potential values. In addition, it provides a wide range of possibilities to assess its evolution over time and provides an interesting tool for assessing the technological potential of the company's production processes.

Consistently based on the preparation of a matrix and taking the entire portfolio of the products or all of the functionalities that are offered by the company into account, it provides the possibility of a strategic assessment of the current technological potential of the company – especially in terms of its market value and level of innovativeness.

5. CONCLUSION

Companies that launch innovative products should have a strictly defined target clientele, which they can achieve by developing interactive communications with the customers or users; such linkages result in minimizing the risks of commercialization of innovations, and such a role can be played by the presented and described tool, which streamlines the process of developing innovative products thanks to the very possible visualization of the designed product and the possibility of assessing the levels of innovation and customer-perceived values that are attributed to its utility functions. The presented possibilities are to be used in empowering the decision-making process at the operational level with regard to the commercialized product and at the strategic level through the possible analysis of the entire portfolio of all of the offered utility functions that come from the technologies that are used. An important feature of the presented model is its possible full quantification; thus, it is possible to carry out a quantitative analysis of the designed product with regard to its level of innovation, the customer value proposition, and the value that is perceived by the customer. This is so important because a quantitative description that is based on these parameters is interesting from the perspective of marketing activities. Regarding the use of the concept of technical debt as a measurement of the level of an innovation, it is worth noting the importance of this approach, which sanctions the innovation of the technology that is used in terms of its perception by the organization. It is interesting to look at this issue in terms of commercialization decisions. The described approach facilitates the implementation of the technology innovation process; thus, firms should provide the best-possible solutions for the individual customer. To assess this, a standard evaluation tool is needed that highlights any changes in the value perception for innovative technology implementation, thus helping a company to observe and analyze changes in user behaviors. In addition, the model provides the opportunity to take the projected future segmentation of the market into account and, thus, offers an overview of different concepts for new products and even the definition of new usability. This translates into the possible conceptualization of new utility functions based on any new technologies under development – even at the beginning of an innovation process that is already underway. The role of the customer then becomes anticipating the new product's meaning.

REFERENCES

Artmann C. (2009). *The Value of Information Updating in New Product Development*. Berlin – Heidelberg: Springer-Verlag. **DOI**: https://doi.org/10.1007/978-3-540-93833-0.

Asthana P. (1995). Jumping the s-curve. *IEEE Spectrum*, June 1995, pp. 49–54. **URL**: https://api.semanticscholar.org/CorpusID:109777335.

Blut M., Chaney D., Lunardo R., Mencarelli R. & Grewal D. (2023). Customer perceived value: A comprehensive meta-analysis. *Journal of service Research*, **27**(4), pp. 1–24. **DOI**: https://doi.org/10.1177/10946705231222295.

Bowden M.J. (2004). Moore's Law and Technology S-Curve. *Stevens Alliance for Technology Management, SATM*, **8**(1), pp. 1–4. **URL**: https://gwern.net/doc/technology/2004-bowden.pdf [10.07.2024].

Chen S.-Ch. & Quester P.G. (2009). A value-based perspective of market orientation and customer service. *Journal of Retailing and Consumer Services*, **16**(3), pp. 197–206. **DOI**: https://doi.org/10.1016/j.jretconser.2008.11.015.

Cunningham W. (1993). The WyCash portfolio management system. In: J.L. Archibald, M.C. Wilkes (Eds.), *Addendum to the Proceedings: Object-Oriented Programming Systems, Languages, and Applications*, OOPSLA'92, New York: ACM Press, pp. 29–30. **DOI**: https://doi.org/10.1145/157710.157715.

Eversheim W. (Ed.) (2009). *Innovation Management for Technical Products. Systematic and Integrated Product Development and Production Planning*. Berlin – Heidelberg: Springer-Verlag. **DOI**: https://doi.org/10.1007/978-3-540-85727-3.

Fan B., Qi G., Hu X. & Yu T. (2015). A network methodology for structure-oriented modular product platform planning. *Journal of Intelligent Manufacturing*, **26**, pp. 553–570. **DOI**: https://doi.org/10.1007/s10845-013-0815-1.

Filipowicz P. (2014). Customer commitment to value creation process: case of innovation based differentiation strategies. *International Journal of Business and Management Study*, **1**(1), pp. 10–13. **URL**: https://www.zarz.agh.edu.pl/pfilipow/publikacje/pozycja5.pdf [10.07.2024].

Filipowicz P. (2015). Technical debt and customer value added as the parameters of technology innovation based strategies. *Central and Eastern European Journal of Management and Economics*, **3**(4), pp. 255–269.

Filipowicz P. (2019). Conceptualization of product development model based on use function evolution. *Central and Eastern European Journal of Management and Economics*, **7**(1), pp. 9–20.

Foster R. (1986). *Innovation: The Attacker's Advantage*. New York: Summit Books.

Fuchs C. & Golenhofen F.J. (2019). *Mastering Disruption And Innovation in Product Management: Connecting the dots (Management for Professional).* Heidelberg: Springer International Publishing AG. **DOI**: https://doi.org/10.1007/978-3-319-93512-6.

Ha Y.W. & Park M.C. (2013). Antecedents of customer satisfaction and customer loyalty for emerging devices in the initial market of Korea: An equity framework. *Psychology and Marketing*, **30**(8), pp. 676–689. **DOI**: https://doi.org/10.1002/mar.20637.

Highsmith J. (2009). *Agile Project Management: Creating Innovative Products*. Second Edition. Redwood City: Addison-Wesley.

Ho J.C. (2011). Technology evaluation in Taiwan's technology industries: Strategies, trajectories, and innovations. *Technological Forecasting & Social Change*, **78**(8), pp. 1379–1388. **DOI**: https://doi.org/10.1016/j.techfore.2011.03.008.

Ho J., Shahnewaz S. & Ruhe G. (2014). A Prototype Tool Supporting When-to--release Decisions in Iterative Development. In: 2nd International Workshop on Release Engineering (RELENG), Mountain View, CA, USA, pp. 1–3. **URL**: https://releng.polymtl.ca/RELENG2014/html/proceedings/releng2014_submission__11.pdf [10.07.2024].

Huang M.-H. & Wang E.T.G. (2013). Marketing is from Mars, IT is from Venus: Aligning the worldviews for firm performance. *Decision Sciences Journal*, **44**(1), pp. 87–125. **DOI**: https://doi.org/10.1111/j.1540-5915.2012.00396.x.

Jensen M.B., Hienerth Ch. & Lettl Ch. (2014). Forecasting the commercial attractiveness of user-generated designs using online data: An empirical study within the LEGO User Community. *Journal of Product Innovation Management*, **31**(S1), pp. 75–93. **DOI**: https://doi.org/10.1111/jpim.12193.

Keinonen T. & Takala R. (Eds.) (2006). *Product Concept Design. A Review of the Conceptual Design of Products in Industry*. Springer Science + Business Media. **DOI**: https://doi.org/10.1007/978-1-84628-126-6.

Kumar P. & Puneet T. (2019). A paradigm for customer-driven product design approach using extended axiomatic design. *Journal of Intelligent Manufacturing*, **30**, pp. 589–603. **DOI**: https://doi.org/10.1007/s10845-016-1266-2.

Kumar S. & Phrommathed P. (2005). *New Product Development: An Empirical Approach to Study of the Effects of Innovation Strategy, Organization Learning, and Market Conditions*. Springer Science+Business Media. **DOI**: https://doi.org/10.1007/b101081.

Liu Q., Du Q., Hong Y., Fan W. & Shuang W. (2020). User idea implementation in open innovation communities: Evidence from a new product development crowdsourcing community. *Information System Journal*, **30**(5), pp. 899–927. **DOI**: https://psycnet.apa.org/doi/10.1111/isj.12286.

McNally R.C., Akdeniz M.B. & Calantone R.J. (2014). New product development processes and new product profitability: Exploring the mediating role of speed to market and product quality. *Journal of Product Innovation Management*, **28**(S1), pp. 63–77. **DOI**: https://doi.org/10.1111/j.1540-5885.2011.00861.x.

Millier P. (2011). *Stratégie et marketing de l'innovation technologique*. 3ème edition. Paris: Dunod.

Moon H., Park J. & Kim S. (2015). The Importance of an innovative product design on customer behavior: Development and validation of a scale. *Journal of Product Innovation Management*, **32**(2), pp. 224–232. **DOI**: https://doi.org/10.1111/jpim.12172.

Moore G.E. (1965). Cramming more components onto integrated circuits. [reprinted from *Electronics*, **38**(8), pp.114–117]. *IEEE Solid-State Circuits Society Newsletter*, **11**(3), pp. 33–35. **DOI**: https://doi.org/10.1109/N-SSC.2006.4785860.

Mul J.& Di Benedetto C.A. (2011). Strategic orientations and new product commercialization: mediator, moderator, and interplay. *R&D Management*, **41**(4), pp. 337–359. **DOI**: https://doi.org/10.1111/j.1467-9310.2011.00650.x

Mulder-Nijkamp M. (2020). Bridging the gap between design and behavioral: (Re) searching the optimum design strategy for brands and new product innovations. *Creativity and Innovation Management*, **29**(4), pp. 11–26. **DOI**: https://doi.org/10.1111/caim.12393.

Nord R., Ozkaya I., Kruchten P. & Gonzales-Rojas M. (2012). In search of a metric for managing architectural technical debt. In: *2012 Joint Working IEEE/IFIP Conference on Software Architecture & European Conference on Software Architecture*, pp. 91–100. **DOI**: https://doi.org/10.1109/WICSA-ECSA.212.17.

Pearce J.A. & Robinson R.B. (1994). *Strategic Management Formulation, Implementation, and Control*. Fifth Edition. Burr Ridge – Boston – Sydney: Richard D Irwin.

Pieper T. (2019). *User Innovation Barriers' Impact on User-Developed Products. An Empirical Investigation on User Innovation Processes*. Springer Gabler. **DOI**: https://doi.org/10.1007/978-3-658-25506-0.

Porter M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press.

Radford S. & Sridhar S. (2005). All co-production is not created equal: A value congruence approach for examining the degree of co-production. In: B.A. Walker, M.B. Houston (Eds.), *AMA Summer Educators' Conference 2005. Enhancing Knowledge development in Marketing*, American Marketing Association, pp. 244–251.

Ritter T. & Walter A. (2012). More is not always better: The impact of relationship functions on customer-perceived relationship value. *Industrial Marketing Management*, **41**(1), pp. 136–144. **DOI**: https://doi.org/10.1016/j.indmarman.2011.11.020.

Sánchez-González G. & Herrera L. (2015). User cooperation effects on firm's innovation outputs. *Canadian Journal of Administrative Sciences Revue Canadienne des sciences de l'administration*, **32**(2), pp. 86–101. **DOI**: https://doi.org/10.1002/cjas.1319.

Schultz C., Salomo S., Talke & K. (2013). Measuring new product portfolio innovativeness: How differences in scale width and evaluator perspectives affect its relationship with performance. *Journal of Product Innovation Management*, **30**(S1), pp. 93–109. **DOI**: https://doi.org/10.1111/jpim.12073.

Sood A. & Tellis G.J. (2005). Technological evolution and radical innovation. *Journal of Marketing*, **69**(3), pp. 152–168. **DOI**: https://doi.org/10.1509/jmkg.69.3.152.66361.

Tang D., Yin L. & Ullah I. (2018). *Matrix-Based Product Design and Change Management*. Singapore: Springer Nature. **DOI**: https://doi.org/10.1007/978-981-10-5077-0.

Townsend J.D., Wooseong Kang W., Montoya M.M. & Calantone R.J. (2013). Brand-specific design effects: Form and function. *Journal of Product Innovation Management*, **30**(5), pp. 994–1008. **DOI**: https://doi.org/10.1111/jpim.12042.

Verhagen T., Boter J. & Adelaar T. (2010). The effect of product type on consumer preferences for website content elements: An empirical study. *Journal of Computer-Mediated Communication*, 16(1), pp. 139–170. **DOI**: https://doi.org/10.1111/j.1083-6101.2010.01536.x.