sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 15 December 2024
accepted: 15 May 2025

pages: 1-22



# Integrating policy agendas for sustainable innovation: a quadruple helix approach to the European Green Deal and sustainability transitions

Monika Popper
Rafael Popper
Carina Alejandra Rapetti
Josep Miquel Pique

**Monika Popper**
Futures Diamond Ltd,
9 Bold Street, Warrington, WA1 1DN,
Manchester, United Kingdom
ORCID 0009-0000-7138-0002

Smart Society Research Group,
La Salle – Ramon Llull University,
Sant Joan de La Salle 42,
08022, Barcelona, Spain

Corresponding author:
e-mail: monika.popper@futuresdiamond.com

**Rafael Popper**
Faculty of Mechanical and Industrial Engineering, Warsaw University ofTechnology, 02-524 Warsaw, Poland
e-mail: Rafael.Popper@pw.edu.pl

Smart Society Research Group,
La Salle – Ramon Llull University,
Sant Joan de La Salle 42,
08022, Barcelona, Spain
e-mail: Rafael.Popper@salle.url.edu

Alliance Manchester Business School,
The University of Manchester,
Oxford Rd, Manchester M13 9PL,
United Kingdom
e-mail: Rafael.Popper@manchester.ac.uk
ORCID 0000-0002-4653-4994

**Carina Alejandra Rapetti**
Smart Society Research Group,
La Salle – Ramon Llull University,
Sant Joan de La Salle 42,
08022, Barcelona, Spain
ORCID 0000-0001-6423-7478
e-mail: Carina.Rapetti@salle.url.edu

**Josep Miquel Pique**
Smart Society Research Group,
La Salle – Ramon Llull University,
Sant Joan de La Salle 42,
08022, Barcelona, Spain
ORCID 0000-0002-0215-0139
e-mail: jm.pique@salle.url.edu

ABSTRACT

This paper explores the integration of DIY (Do-It-Yourself) and crowdsourcing-driven innovation management with top-down policy frameworks to advance the European Green Deal and sustainability transitions. Drawing on the CASI (Common Framework for the Assessment and Management of Sustainable Innovation) project, it examines how CASI-F and its CASIPEDIA platform — a repository of over 700 sustainable innovation initiatives — can bridge local and global sustainability goals. Using a quadruple helix approach, the study highlights the role of innovation ecosystems in driving systemic change through the alignment of technological, economic, social, and environmental objectives. The paper compares the CASI SI agendas with the European Green Deal and sustainability transition agendas, identifying synergies, gaps, and opportunities for further alignment. Foresight is emphasised as a crucial tool to ensure innovation strategies remain adaptable, future-oriented, and resilient to emerging challenges. The findings underscore the importance of integrating grassroots innovations into broader policy frameworks to enhance sustainability transitions. Lastly, the paper recommends strengthening the Green Deal by fostering innovations from diverse actors, including those from the quadruple helix, and promoting the application of SMART foresight processes.

KEY WORDS

**sustainability transitions, innovation ecosystems, European Green Deal, crowdsourcing-driven innovation, quadruple helix, sustainable innovation (SI) agendas, foresight methodology, climate-neutral future**

**10.2478/emj-2025-0009**

## Introduction

The CASI project, which concluded in 2017, laid the foundational groundwork for significant contributions to sustainable innovation across Europe through its sustainable innovation (SI) agendas. Although the project itself has now ended, CASI-F and its associated tools, such as the CASIPEDIA mapping platform, continue to attract users, with over 100 new SI cases added annually by a growing community of SI enthusiasts (Martini et al., 2020a). This sustained engagement with CASI-F, coupled with its increasing adoption across Europe, provides the impetus for this

Popper, M., Popper, R., Rapetti, C. A., & Pique, J. M. (2025). Integrating policy agendas for sustainable innovation: a quadruple helix approach to the European Green Deal and sustainability transitions. *Engineering Management in Production and Services, 17*(2), 1-22. doi: 10.2478/emj-2025-0009

paper. In particular, the aim of this paper is to reanalyse the 10 CASI SI agendas within the context of the sustainability transition agendas and the European Green Deal agendas, with a particular focus on identifying commonalities, unique features, and synergies between these agendas. The CASI SI agendas, developed through a bottom-up, empirical approach, offer valuable insights into crowdsourcing-based innovation initiatives that have emerged across Europe. These initiatives are deeply rooted in local communities, businesses, and civil society, emphasising community-driven efforts, local knowledge, and practical solutions that are tailored to specific regional contexts. The ongoing application of CASI-F and its platform further underscores the relevance of these agendas in driving innovation processes, effectively bridging the gap between local, bottom-up innovation and broader sustainability goals.

Foresight plays a crucial role in connecting local innovation with broader sustainability goals. As positioned by Popper (2008), foresight is a systematic and methodologically rigorous process requiring careful design and implementation. This complexity necessitates its thoughtful integration into policymaking, ensuring that sustainable innovation agendas are both analytically grounded and adaptable to emerging challenges. Foresight is a participatory, prospective, and policy-oriented process that actively engages key stakeholders through environmental and horizon scanning techniques, aiming to anticipate, recommend, and transform (ART) futures across technological, economic, environmental, political, social, and ethical (TEEPSE) dimensions (Popper, 2011). This participatory and forward-looking nature makes foresight an essential tool for guiding long-term sustainability transitions, especially within the framework of the European Green Deal, where it informs policy decisions that address the interconnected challenges of achieving a climate-neutral future.

In contrast, the European Green Deal, initiated by the European Commission, represents a top-down, policy-driven framework that sets out ambitious targets for climate neutrality by 2050. It identifies strategic actions across key sectors such as energy, transport, agriculture, and biodiversity, with the aim of catalysing systemic transformation across Europe. Similarly, the sustainability transition agendas focus on understanding the political, social, and economic drivers and barriers of sustainability transitions, offering a comprehensive theoretical framework for facilitating transformative societal change. While these top-down and theoretical agendas are critical in guiding Europe's overarching sustainability strategy, their effectiveness relies on deeper integration with local-level innovations to achieve tangible results on the ground.

This paper compares these three sets of agendas, focusing on how the crowdsourcing-driven CASI SI agendas can complement and enhance the policy-driven frameworks of the European Green Deal and the research-focused sustainability transition agendas. By examining areas of alignment, identifying gaps, and exploring opportunities for further development, the analysis offers a novel perspective on how local innovations and top-down policies can be aligned to accelerate sustainability transitions. Moreover, mapping initiatives such as CASI-F and its CASIPEDIA platform are instrumental in understanding the intersection of bottom-up and top-down efforts. As Popper (2009) asserts, "Foresight research is not only interdisciplinary in theory but also in practice", emphasising the importance of mapping platforms that facilitate interdisciplinary connections and promote collaboration across sectors. These initiatives play a pivotal role in ensuring that local innovations are effectively mapped and integrated into broader policy frameworks, fostering synergies that can drive sustainable development.

An additional purpose of this paper is to demonstrate the importance of tools like CASIPEDIA as an instrument for capturing and disseminating crowdsourcing-based innovations. By leveraging CASI-F and its ever-expanding database of sustainable innovation initiatives, CASIPEDIA offers valuable insights into successful local projects, enabling their integration into broader policy frameworks. The potential of such tools for fostering interdisciplinary collaboration and advancing sustainability transitions in Europe cannot be overstated. Ultimately, the paper aims to illustrate how the continued use of CASI-F, underpinned by SMART foresight principles, can support a more integrated and inclusive approach to achieving Europe's climate-neutral goals. This approach ensures that systemic change is both actionable and widely supported, paving the way for a more sustainable and inclusive future across Europe.

## 1. Literature review

### 1.1. Innovation ecosystems: catalysts for systemic change

Innovation ecosystems are increasingly recognised as pivotal drivers of systemic change, particu-

larly within the context of sustainability transitions. These ecosystems are dynamic networks of interconnected actors — governments, businesses, research institutions, and civil society — that collaborate to address complex, cross-cutting challenges. By facilitating the flow of knowledge, resources, and technology, innovation ecosystems have the potential to instigate transformative shifts across various sectors. The concept of systemic change, or transformation, is a fundamental principle of sustainability science. It involves far-reaching shifts across economic, technological, and social systems, enabling societies to effectively respond to global challenges such as climate change and resource depletion. Unlike incremental changes that refine existing systems, systemic change necessitates the complete rethinking and redesigning of these systems. Achieving such transformation requires a realignment of policies, practices, and institutional frameworks to align with sustainability objectives. As Kivimaa and Kern (2016) suggest, transitions involve not only the development of disruptive innovations but also the formulation of policies designed to drive broader changes within socio-technical systems. This underscores the critical role of integrating technological innovation with policymaking to facilitate large-scale transformations. It is through the strategic alignment of these elements that societies can navigate the complexities of sustainability transitions and achieve meaningful change.

Within this context, innovation ecosystems serve as the driving force behind systemic shifts, enabling the scaling of sustainable innovations by integrating technological advancements with new governance models, business practices, and societal norms. An innovation ecosystem is the evolving set of actors, activities, and artefacts, and the institutions and relations, including complementary and substitute relations, that are important for the innovative performance of an actor or a population of actors (Granstrand & Holgersson, 2020). This dynamic and evolving nature is crucial for facilitating the innovation required for sustainability transitions, and these ecosystems can compete with or complement each other (ibid.). This is particularly relevant for the European Green Deal, which spans sectors like energy, transport, agriculture, and industry, requiring collaboration and competition to drive technological solutions for sustainability. As Oh et al. (2016) highlight, innovation ecosystems are purposefully designed and engineered systems aiming to achieve specific goals, such as sustainability transitions, rather than evolving naturally. They thrive on the balance of competition and collaboration, ensuring that diverse actors both cooperate and compete to drive innovation. Furthermore, the era of the laissez-faire approach to market development is over, and today's complex challenges necessitate active governance and intervention to align innovation with sustainability objectives. This shift requires clear policy frameworks that guide and regulate the ecosystem's development, ensuring that market dynamics support the overarching goals of the European Green Deal.

Transitions encompass changes both across value chains and within each part of them, highlighting the need for multi-dimensional transformations in technological and societal structures (Markard & Truffer, 2008). Furthermore, the innovation systems approach provides greater analytical power by examining the structural and functional dynamics within these ecosystems, offering valuable insights into how different actors collaborate and navigate complex challenges. A key feature of these transitions is the fostering of competition among diverse suppliers, which in turn can stimulate market-driven solutions that enhance efficiency and effectiveness in addressing sustainability goals. The innovation ecosystem concept emphasises both collaboration and competition, offering a more holistic and nuanced understanding of how innovation ecosystems function. The European Green Deal could greatly benefit from supporting such a variety of technological solutions rather than focusing on a single pathway, as this diversity fosters more robust and adaptive innovation ecosystems.

A key characteristic of successful innovation ecosystems is their ability to effectively address the interdependence between various actors. As highlighted by recent studies on ecosystem dynamics, the challenges faced by different stakeholders — from technology developers to end-users — are distinct yet deeply interconnected. The challenges in upstream components, such as those faced by suppliers and technology developers, are qualitatively different from those faced by downstream complementors who enable customers to fully utilise new innovations. For example, innovations in clean energy technologies may require not only advances in the core technology itself but also changes in infrastructure, such as energy grids or charging stations, which are essential for their widespread adoption (Adner & Kapoor, 2010).

In the context of the European Green Deal, these distinctions are particularly relevant. Achieving sustainability requires more than just technological breakthroughs; it also requires systemic changes in

how innovations are integrated into existing societal and economic systems. This means addressing not only the internal innovation challenges within firms but also considering how the surrounding ecosystem — including suppliers, regulators, and customers — must evolve to support these innovations. Understanding the structure of technological interdependence in such ecosystems allows for better identification of the barriers to scaling sustainable solutions and formulating policies that target these bottlenecks (Adner & Kapoor, 2010).

As the focus shifts from individual innovations to broader systems of collaboration, it becomes clear that successful sustainability transitions depend on the alignment of diverse actors across sectors. The European Green Deal provides a framework to address the interconnected challenges of climate change, biodiversity loss, and inequality, offering an opportunity to foster the kind of multi-stakeholder collaboration needed for systemic change. This approach moves beyond firm-level innovation and highlights the importance of a broader, more inclusive ecosystem perspective, which enables technology evolution and creates pathways for achieving long-term sustainability goals.

### 1.2. European Green Deal (EGD): a framework for collaboration

The European Green Deal (EGD), launched in 2019, presents an ambitious roadmap to make Europe the world's first climate-neutral continent by 2050. Central to this vision is a commitment to reducing greenhouse gas emissions across key sectors such as energy, agriculture, transport, and industry while addressing the profound challenges posed by climate change. However, the success of this transition depends on its inclusivity and the ability to mitigate social and economic impacts, particularly for vulnerable regions and workers. The European Commission (2019a, 2019b, 2020) highlights the importance of addressing these inequities through mechanisms like the Just Transition Fund and Mechanism, which aim to mobilise at least EUR 100 billion to support the regions and sectors most affected by the green transition. Despite these efforts, concerns remain regarding the adequacy of the EGD's focus on social sustainability and equity, which are often overshadowed by decarbonisation goals (Hereu-Morales et al., 2024; Kwilinski et al., 2024)).

A key element of the EGD is its emphasis on systemic innovation, calling for an integrated, cross-sectoral approach that fosters synergies between technologies, policies, and societal practices. This vision requires the development of robust innovation ecosystems that enable multi-stakeholder cooperation, knowledge exchange, and the scaling of sustainable solutions. By facilitating the integration of sustainable practices across sectors such as energy, transport, and agriculture, the EGD envisions the transformation of economic systems prioritising sustainability. However, for a truly just and inclusive transition, this transformation must also address the social and economic dimensions of sustainability (Koundouri et al., 2024). Despite its ambitious goals, the EGD has been criticised for insufficiently incorporating policies to tackle inequalities, poverty, and social justice, even though these are closely linked to environmental sustainability.

The EGD is often described as a collection of evolving targets, intentions, and objectives rather than a fixed strategy, reflecting the complexity and long-term nature of the transition. This dynamic approach acknowledges that policy implementation and societal adaptation will unfold over decades. Thus, the development of innovation ecosystems that bridge the gap between advanced technologies and their real-world applications is crucial. These ecosystems will play a vital role in advancing clean energy, sustainable mobility, and circular economy practices, with initiatives like Destination Earth (DestinE), which creates a digital simulation of Earth to better understand climate impacts, offering essential tools for decision-making and policy development.

To ensure the EGD delivers on both its environmental and social promises, it is critical to develop innovation ecosystems that extend beyond technology to include social innovation. These ecosystems must promote collaboration across all sectors of society, ensuring that no one is left behind in the green transition and that the benefits of sustainable growth are shared equitably. As Szélpál and Varga (2024) emphasise, such collaborative frameworks are essential for fostering both environmental and social sustainability, ensuring that the EGD's ambitions are met comprehensively. A key aspect of the Green Deal's success is its focus on energy independence and security, reducing Europe's reliance on energy imports from politically unstable regions. By promoting renewable energy sources like wind and solar power, the EGD not only contributes to decarbonisation but also enhances Europe's economic competitiveness and political stability. Furthermore, the transition to a circular economy, which replaces the

traditional linear model of resource use, is integral to decoupling economic growth from resource consumption, supporting the EGD's target of net-zero emissions by 2050. This focus on systemic innovation ensures a sustainable and economically viable transition. Additionally, the Green Deal's financial mechanisms, such as NextGenerationEU (NGEU) and green bonds, provide critical support for achieving climate and sustainability targets, with investments directed towards long-term resilience. In this context, initiatives like the "Fit for 55" package, which seeks to reduce net greenhouse gas emissions by at least 55% by 2030, must be viewed as part of a broader effort to create a just, inclusive, and resilient European economy. Achieving this requires integrating measures to address inequalities, promote decent work, and support communities most affected by the green transition. Only by addressing the full spectrum of sustainability — environmental, social, and economic — can the EGD be realised as a transformative framework, positioning Europe as a global leader in sustainable development.

## 2. Research methods

This study adopts a mixed-methods approach, combining both qualitative and quantitative datasets to assess the role of innovation ecosystems in advancing the objectives of the European Green Deal. The methodology draws upon the CASI project, a European Union-funded initiative that systematically mapped and analysed over 500 sustainable innovation cases across Europe. The study compares the ten CASI sustainable innovation (SI) agendas with the European Green Deal and sustainability transition agendas to identify synergies and gaps that could guide more effective sustainability transitions.

### 2.1. Literature review and documentary analysis

The first step of the methodology involved an extensive literature review and documentary analysis. This was critical in establishing a comprehensive understanding of the theoretical frameworks and policy directions relevant to the study. The European Commission's reports and communications about the European Green Deal were examined to identify the core objectives, priorities, and planned actions within the Green Deal. This involved a detailed review of key documents such as the European Green Deal Communication (European Commission, 2019a, 2019b), the Climate Law, and the Green Deal Implementation Plan (European Commission, 2020). These documents provided crucial insight into the EU's strategic plans for achieving climate neutrality by 2050, focusing on decarbonisation, resource efficiency, and technological innovation.

Additionally, the sustainability transitions (ST) agendas outlined in Köhler et al. (2019) were reviewed to understand the academic perspectives on the dynamics of sustainability transitions. This framework, which examines the role of power, governance, social movements, and institutional transformations, was crucial in contextualising the CASI SI agendas

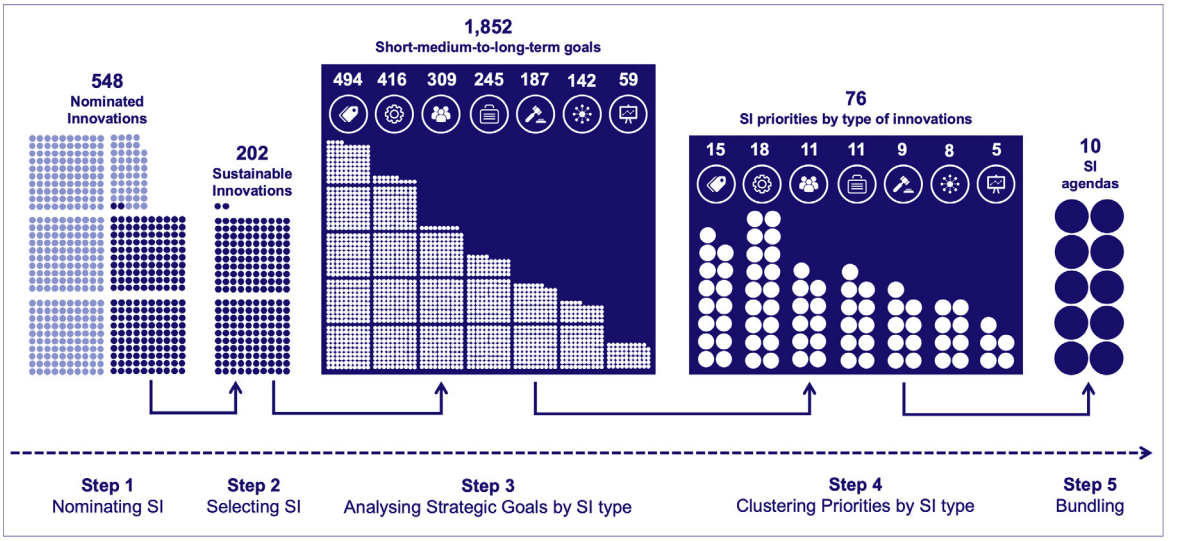



Fig. 1. From 1,852 quadruple helix innovation goals to ten sustainable innovation agendas

within broader theoretical discussions on systemic change.

The SI agendas themselves are based on a review of key documents, including CASI-F and the CASI project's comprehensive reports on sustainable innovation mapping and policy advice, which outline the core principles for fostering crowdsourcing-based innovation assessment and management (Popper et al., 2017a, 2017b, 2017c).

### 2.2. Data collection

During the life of the CASI project (Popper et al., 2020; Martini et al., 2020b; Velasco et al., 2020), the CASI team, consisting of 19 partners from 12 EU countries and 16 country correspondents, nominated a total of 548 initiatives, each country representative selecting 15–22 initiatives to ensure broad coverage of the pre-defined seven types of sustainable innovation (SI) and public participation, as well as engagement with the quadruple helix of stakeholders. Of the nominated initiatives, 49% were led by business actors, 21% by government entities (including intergovernmental organisations such as the EU and the UN), 20% by civil society organisations (including NGOs), and 10% by research and education stakeholders. These initiatives focused on social, economic, and environmental dimensions and aligned with one or more EU Framework Programme for Research and Innovation priorities related to climate action, environment, resource efficiency and raw materials while spanning key sectors such as energy, agriculture, mobility, and the circular economy, ensuring a diverse representation of local, regional, and national-level innovations (Popper et al., 2020). This diversity was vital for addressing sustainability challenges across Europe.

From this pool, 202 initiatives were selected from CASIPEDIA (Box 1) for detailed mapping and analysis based on criteria such as sustainability impact, public participation, scalability, and novelty. The aim was to capture innovations with substantial potential for societal impact, particularly at local and regional levels. These selected initiatives were analysed in terms of their main and supporting objectives, which were then clustered to identify key research and innovation priorities. A total of 1,852 short-, medium-, and long-term goals and aspirations of the quadruple helix of SI stakeholders were analysed, aimed at contributing positively to sustainability. These goals were derived from the different types of innovations, with 494 from product-related innovations, 416 from service innovations, 309 from social innovations, 245 from organisational innovations, 187 from governance innovations, 142 from system innovations, and 59 from marketing-related innovations (Popper et al., 2017a).

From these 1,852 goals, some 76 SI priorities were mapped against the seven types of SI, and these were subsequently clustered into ten sustainable innovation (SI) agendas (Fig. 2). The resulting key terms were carefully analysed and clustered into key SI priorities. Several research priorities per type of SI were formulated by clustering these terms, but only

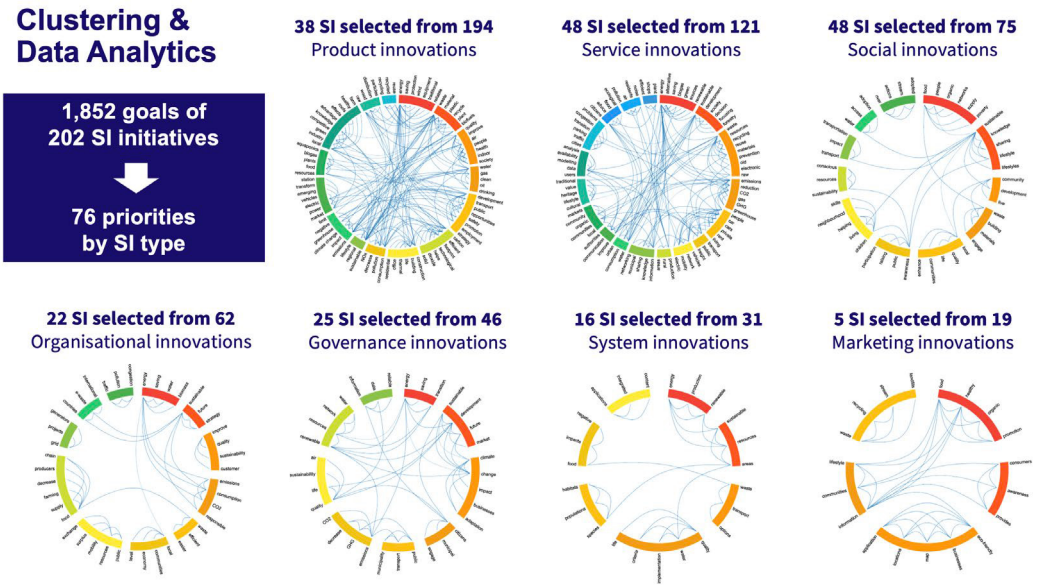



Fig. 2. From SI goals to SI priorities by type of innovation

the top five priorities from each type are discussed in this paper. These priorities were then linked to EU Framework Programme for Research and Innovation priority areas and relevant socio-economic sectors. It is worth noting that while some of the innovators' objectives have already been achieved and are driving positive environmental, economic, and social transformations, they remain as priorities for future developments and diffusion strategies of similar innovations. Thus, the combined analysis of SI objectives can be seen as the implicit, and at times explicit, ongoing and future research and innovation agendas of innovators (ibid.).

### 2.3. Data analysis and comparative framework

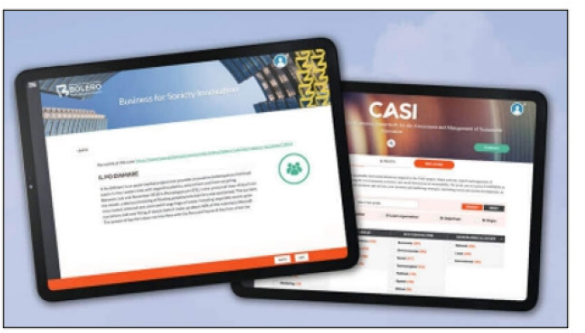
The next step of the analysis involved identifying key sustainable innovation (SI) priorities across seven innovation categories: product, service, social, organisational, governance, system, and marketing. Qualitative and quantitative methods were employed to extract meaningful patterns and trends from the dataset. Text mining tools were used to identify recurring themes, while qualitative content analysis was employed to interpret and contextualise the data, revealing the underlying connections and implications of these themes.

Once the key themes and trends were identified, a comparative analysis was conducted between the ten SI agendas, the sustainability transition agendas, and the European Green Deal agendas. This analysis focused on understanding the synergies between crowdsourcing-based innovation management and top-down policy frameworks. In particular, it aimed to identify potential gaps where further alignment or development might be necessary to ensure cohesive and inclusive sustainability transitions. The key themes from the SI agendas — including eco-community empathy, circular economy, and sustainable transport — were compared with the policy-driven agendas of the Green Deal and the research-oriented agendas of sustainability transitions.

A critical component of this analysis involved using the quadruple helix framework, which maps the roles of government, business, academia, and civil society across these sustainability agendas. Each agenda was assessed for the intensity and relevance of each stakeholder's role in driving its objectives. The comparative matrix highlighted how each stakeholder group contributed to advancing innovation and sustainability goals, focusing on areas such as govern-

By 2025, CASIPEDIA has become a unique repository of some 700 sustainable innovation initiatives, combining environmental, economic, and social dimensions of sustainability. This extensive database provides activists, experts, and sustainability advocates with access to a diverse range of innovative solutions, from new products and services to innovative business strategies, social developments, and emerging policies. CASIPEDIA's wide array of innovations enables stakeholders to explore practical solutions that address the most pressing global sustainability challenges.

- *Innovation types:* CASIPEDIA categorises innovations into seven types, with 36 % focused on products, 21 % on services, and 18% on social innovations. It also includes 10 % organisational innovations, 7 % governance-related solutions, 6 % system innovations, and 4 % marketing innovations.
- *Geographical scope:* The initiatives span diverse geographical contexts, with 41.79 % of innovations at the national level, 31.69 % local, and 26.52 % international in scope, ensuring a comprehensive global perspective on sustainable innovation.




- Key areas: The database prioritises global sustainability challenges with a focus on 36.48 % of initiatives aimed at resource efficiency, 31.03 % addressing climate action, and 25.07 % on raw materials. 7.41 % of the initiatives focus on environmental solutions.
- Success factors: CASIPEDIA captures the impact of innovations across multiple success factors, with 21.92 % emphasising economic impact, 21.86 % environmental, and 18.26 % technological advancements. Social and political success factors each contribute 18.26 % and 8.26 %, respectively, while spatial and ethical dimensions account for 6.09 % and 5.51 %.

Further access to CASIPEDIA is available online at https://www.futuresdiamond.com/casi2020/casipedia/.

Fig. 3. Box 1. CASIPEDIA — a comprehensive resource for sustainable innovation

ance, system innovation, and community-based solutions. This allowed for a nuanced understanding of how crowdsourcing-driven innovations and top-down policy initiatives can complement and reinforce each other, providing a comprehensive view of the innovation ecosystem and its role in achieving sustainability transitions.

### 2.4. Bottom-up horizon scanning and expert insights

As part of the methodology, a self-sustainability element was incorporated into CASI-F through the development of an openly accessible online tutorial designed to promote crowdsourcing-based sustainable innovation assessment and management (Box 2). This innovative, certified executive course represents an example of organisational innovation, introducing a novel approach to learning and applying CASI-F to real-world innovation cases. By providing widespread and unrestricted access to the course and enabling participants to map their own innovation cases into CASIPEDIA, a comprehensive database that aids users in organising and tracking sustainable innovations, the CASI project has established an open and scalable system for continuous innovation.

By 2017, the CASI project had successfully trained and certified 100 participants, meeting its initial target. However, the enduring relevance of the course and its accessibility via an online platform has resulted in sustained interest beyond the completion of the project. As of now, the course has expanded its reach significantly, with over 577 participants from 80 countries around the world. Countries such as Colombia, Poland, and Bulgaria lead in enrolment, underscoring the course's global accessibility and its capacity to foster bottom-up engagement in sustainable innovation practices. This growth in participation highlights the long-term impact and self-sustainability of CASI-F, ensuring its continued relevance in the absence of further EU funding since 2017.

The online tutorial and executive course form a crucial component of the bottom-up horizon scanning strategy, which generates new data for subsequent analysis through focus groups and expert insights. This approach encourages participants from a variety of sectors to contribute their innovation cases and challenges, providing a rich source of real-world data that feeds into the broader analysis. A diverse range of perspectives on local sustainable innovations is captured by engaging directly with

The Sustainable Innovation Assessment and Management course has achieved notable global participation, reflecting widespread interest in sustainability and innovation management across various regions. The distribution of participants highlights the course's reach beyond Europe, with significant engagement from Latin America, Asia, and Africa. Approximately 45 % of participants come from Europe, with countries such as Poland, Italy, and Portugal leading the way. The Americas represent around 30 %, with strong participation from Colombia, Peru, and Brazil. The remaining 25 % is spread across Africa, Asia, and Oceania, demonstrating the global relevance of sustainable innovation practices.

In terms of top participating countries, the course has seen the highest engagement from Colombia, which accounts for 19 % of total participants, followed by Poland (11 %) and Bulgaria (9 %). These three countries alone make up nearly 40 % of the course's total enrolment, showcasing the course's significant impact in these regions. Other countries with notable participation include Italy, Portugal, and Spain, contributing to the European dominance in course enrolment.

Top ten countries by participation:

- Colombia — 108 participants (19 %)
- Poland — 65 participants (11 %)
- Bulgaria — 54 participants (9 %)
- Italy — 43 participants (8 %)
- Portugal — 34 participants (6 %)
- Spain — 22 participants (4 %)
- Peru — 21 participants (4 %)
- Czech Republic — 19 participants (3 %)
- United Kingdom — 17 participants (3 %)
- France — 15 participants (3 %)


This broader geographical spread and the top countries reflect the enduring and expanding relevance of the course, contributing to its legacy of self-sustainability and its alignment with current global sustainability needs. With such widespread participation, the course continues to foster the development of sustainable innovation across borders, nurturing a community of practitioners equipped to address the pressing environmental and societal challenges of our time.

Further access to the CASI-F tutorial is available online at https://www.futuresdiamond.com/casi2020/tutorial/.

Fig. 4. Box 2. Global participation in the CASI-F course

crowdsourced participants, thereby enhancing the understanding of how these innovations evolve and align with broader sustainability objectives.

Furthermore, the methodology is supported by other types of innovation, such as social innovation and system innovation, particularly through the development of the BOLERO project (2022–2025). The BOLERO project is designed to adapt CASI-F to the needs of the University of Milano-Bicocca in Italy, offering a practical example of how system innovation can transform education and innovation management by linking local institutional needs with broader sustainability agendas. The BOLERO project serves as a sister mapping platform to CASIPEDIA, further advancing the integration of bottom-up approaches with the overarching goals of sustainability. It demonstrates how social innovation can be embedded into educational and research practices by encouraging collaborative, knowledge-sharing environments that enhance societal relationships and create new forms of collaboration within the university and beyond.

The focus groups formed as part of this methodology comprised stakeholders from local innovation ecosystems, including innovators, policymakers, and researchers, who discussed key findings from the SI agendas and their relevance to the European Green Deal. These discussions provided valuable insights into how innovations from local contexts can align with or challenge larger policy frameworks. Focus groups not only identified barriers but also highlighted enablers to the integration of local innovations with global sustainability objectives. Additionally, expert insights were gathered from leading researchers in sustainability transitions, policy advisors, and practitioners engaged in the European Green Deal. Their input provided a more nuanced understanding of the challenges and opportunities for sustainability transitions in Europe. These insights helped refine the analysis, offering perspectives on how top-down policy initiatives could better support the integration of bottom-up innovations, thus enhancing the scalability and impact of sustainable innovations across different levels. Through these combined efforts, a global community of innovators and researchers committed to sustainable change continues to grow.

### 2.5. Comparative analysis and synthesis

The final phase of the analysis synthesised the findings from the SI agendas, the European Green Deal, and the sustainability transition agendas. This synthesis aimed to develop a more integrated and inclusive approach to sustainability transitions, highlighting how the SI agendas, with their emphasis on community-driven innovation, could complement and strengthen the Green Deal's ambitious goals. The comparative analysis also identified specific areas of alignment, potential gaps, and opportunities for policy development, ultimately contributing to a more cohesive sustainability agenda that integrates local innovations with broader European objectives.

## 3. Research results and discussion

The alignment between the CASI SI agendas, the European Green Deal, and the transitions research agendas – complemented by the guidance set out in Action Roadmaps for More Resilient Research and Innovation Futures – paves a shared pathway toward a climate-resilient, sustainable future (Popper & Popper, 2024, 2025). These three frameworks share common overarching objectives, including the reduction of greenhouse gas emissions, the enhancement of resource efficiency, and the promotion of sustainability across various sectors. By comparing and contrasting these agendas, this section highlights the synergies that can foster more integrated approaches for the implementation of the Green Deal (Szpilko & Ejdys, 2022). The identification of these commonalities not only reveals areas of alignment but also emphasises the significance of combining crowdsourced innovation management with top-down policy frameworks to drive systemic transformations. This analysis provides valuable insights into how local innovations, as outlined in the CASI SI agendas, can meaningfully contribute to the ambitious goals of the European Green Deal, ensuring that both local and global sustainability objectives are effectively pursued.

### 3.1. Synergies for sustainability: a comparative analysis of key agendas

This subsection presents a comparative analysis of ten sustainable innovation agendas, nine sustainability transition agendas, and ten European Green Deal agendas, focusing on their key areas of alignment. The five core themes explored — environmental sustainability and resource management; technological innovation and industrial transformation; social change, equity, and justice; governance, policy, and institutional transformation; and integra-

tion of sustainability across sectors — serve as the foundational pillars through which these agendas intersect. Examining the synergies between these three frameworks uncovers the potential for a more integrated approach that combines bottom-up innovation with top-down policy frameworks. This comparative analysis not only clarifies the shared priorities of these agendas but also provides insight into how localised innovation efforts, as outlined in the CASI sustainable innovation agendas, can complement and reinforce the European Union's broader sustainability objectives. In doing so, it highlights the opportunities for a more coordinated and cohesive path toward a sustainable, climate-neutral future. The following analysis delves deeper into the converging themes across these agendas, illustrating how their collective strengths can be leveraged to foster long-term sustainability.

### 3.1.1. Sustainable innovation agendas

The ten sustainable innovation (SI) agendas presented here are based on previous work from the authors in the CASI project (Popper et al., 2017a; 2017b), though they have been revised and updated to reflect current challenges and emerging trends. These agendas focus on addressing urgent global issues through systemic, multidisciplinary approaches. They encompass the transformation of energy, transportation, and resource management while fostering resilience and innovation in governance and education. Promoting sustainable practices across sectors, they advocate for circular economies, renewable energy, and localised food systems, aiming to equip communities with the necessary tools for long-term sustainability.

1. Tackling climate change and reducing greenhouse gas emissions agenda calls for innovative approaches to tackling climate change through systemic solutions at multiple levels. It advocates for a shift from viewing climate change as a problem to recognising it as an opportunity for broad societal transformation. Key strategies include the implementation of climate mitigation technologies, the promotion of low-carbon business models, and the integration of climate-resilient policies across sectors. R&I efforts should focus on creating adaptive, multilevel solutions that engage all sectors of society, empowering communities, and ensuring climate justice, while fostering collaboration across industries, governments, and citizens to accelerate the transition to a low-carbon future.

2. Enhancing governance through foresight and data-driven intelligence. In a world increasingly shaped by uncertainty and unpredictability, the need for anticipatory governance has never been more urgent. This agenda focuses on the integration of foresight methodologies, data analytics, and predictive intelligence into governance structures to better address sustainability challenges. By leveraging big data, artificial intelligence (AI), and scenario planning, it aims to create adaptive, responsive governance models that can navigate global shifts in climate policy, geopolitics, and socio-economic dynamics. The role of ICT and data-driven decision-making is crucial to enable governance that is proactive, flexible, and capable of responding to wild card events, such as policy reversals and climate commitments, ensuring long-term sustainability and resilience.

3. Building integrated sustainable infrastructures for a circular bioeconomy agenda focuses on developing sustainable urban and rural infrastructures that support a circular bioeconomy. Emphasising material efficiency, sustainable building practices, and the integration of smart technologies, it explores how infrastructure can reshape supply chains and consumption patterns. The role of governance and social innovation is crucial for scaling circular economy models, particularly in resource management. Future R&I should explore systemic approaches to creating connected economies, where sustainable technologies work in tandem with social, economic, and governance innovations to enhance resilience and reduce environmental impacts across diverse geographical and socio-economic contexts.

4. Advancing responsible environmental management and resource efficiency agenda focuses on balancing economic activity with environmental sustainability, particularly through resource-efficiency strategies. By adopting a systems approach, it addresses product innovations, governance models, and social innovations aimed at reducing emissions and managing resource consumption. It advocates for integrating circularity into industrial and residential systems, with a focus on water, energy, and waste. Moving beyond technical solutions, it calls for systemic, behavioural, and governance innovations that promote sustainable consumption and equitable distribution while also focusing on restorative approaches to natural ecosystems and promoting collaborative action for inter-sectoral sustainability.

5. Accelerating the transition to renewable energy and biofuels. The focus of this agenda is the transformation of energy systems towards sustaina-

bility, primarily through renewable energy and biofuel solutions. It addresses both supply-side innovations, such as biogas and anaerobic digestion, and demand-side strategies, such as energy efficiency and community energy initiatives. Future R&I efforts should focus on creating integrated energy systems, incorporating social and economic governance models that facilitate the transition. This agenda highlights the need for policies that support low-carbon technologies while also considering societal impacts, including equity in access to renewable energy, community involvement, and sustainable energy consumption practices.

6. Innovating sustainable transport and mobility systems agenda focuses on sustainable transport solutions, advocating for a shift towards low-carbon, accessible mobility systems. It examines innovations in transport technologies, such as electric and hydrogen vehicles, alongside social innovations that encourage shifts in behaviour and public engagement. The integration of smart technologies, such as big data and IoT, is central to optimising transport systems and promoting more sustainable urban planning. Future R&I should explore the tensions between technological advances, such as smart cities, and the social implications of mobility, including access, equity, and the challenge of reducing dependency on unsustainable transport models.

7. Advancing circular economy and waste resource management agenda promotes the transition to a circular economy where waste is viewed as a resource. It highlights the importance of systemic innovation across industries, focusing on circular business models, waste reduction, and recycling. Drawing from examples of both small-scale community-based initiatives and large-scale industrial symbiosis, it encourages scaling up successful models through improved infrastructure, financing, and design innovation. R&I efforts should address barriers to circularity, exploring solutions for product life-cycle management, waste-to-resource technologies, and the social dynamics of consumption to facilitate the transition to a sustainable, zero-waste economy.

8. Enhancing eco-community resilience and collaborative development agenda aims to foster sustainable, inclusive communities through empathy, solidarity, and local economic resilience. Emphasising multi-stakeholder engagement, it advocates for collaborative governance models that address gender, ethnic equality, and citizen empowerment. By leveraging digital platforms, this approach enhances local development through crowd-funded initiatives, cooperative business models, and sustainable regional development. Additionally, it promotes environmental stewardship in both urban and rural settings. The focus is on enabling behavioural change, with empathy as a central driver, creating institutions that are socially, culturally, and ecologically responsive to long-term sustainability challenges.

9. Embedding sustainability in education and skills development agenda stresses the critical role of education in shaping a sustainable future. It advocates for incorporating sustainability into all levels of education, from primary schools to professional training, promoting both technical and behavioural change. The focus is on fostering skills that support sustainable practices in diverse sectors, from green technologies to community-based sustainability initiatives. Future R&I should explore the potential of online platforms, gamification, and eco-feedback systems to engage citizens, businesses, and policymakers. The aim is to create an informed, active populace capable of driving and participating in the transformation towards a more sustainable world.

10. Promoting sustainable agriculture and localised food systems agenda focuses on fostering sustainable agricultural practices that support both ecological health and food security. Emphasising local food systems, it encourages innovations in farming techniques, such as aquaculture and alternative cultivation practices, that enhance resource efficiency and reduce environmental impact. By promoting circularity in food production and distribution, it seeks to reconnect producers and consumers, encouraging sustainable consumption. Future R&I should explore scaling up successful local initiatives and identifying pathways to influence global food systems towards sustainability, addressing challenges such as public health, education, and the need for a post-oil food transition.

### 3.1.2. Sustainability transition agendas

The nine sustainability transition agendas presented here are based on collaborative work from the Sustainability Transitions Research Network (STRN), evolving from earlier research and discussions within the network (Köhler et al., 2019). These agendas have been revised to reflect the rapid growth and diversification of the field, addressing emerging global challenges. They focus on the complex, multi-dimensional nature of transitions, exploring the roles of governance, civil society, and industry in reshaping energy,

food, transport, and urban systems. The agendas advocate for transformative, equitable change, aiming to build resilient, sustainable societies through interdisciplinary and participatory approaches.

1. Understanding transitions — theoretical frameworks and insights agenda focuses on the foundational frameworks in sustainability transitions, primarily the multi-level perspective (MLP), technological innovation systems (TIS), strategic niche management (SNM), and transition management (TM). These frameworks aim to capture the complexity of transitions, considering the dynamics between niches, regimes, and landscapes. They help explain the processes of system innovation, path dependency, and change. While early studies concentrated on niche innovations, newer research has increasingly explored the role of incumbent regimes and the interactions between emerging and established technologies in shaping sustainable transitions.

2. Examining power and politics in transitions. Transitions are inherently political, as various actors and groups contest the direction and speed of change. This agenda explores the power dynamics that influence sustainability transitions, particularly the resistance of incumbent industries and the lobbying efforts of emerging actors. The research draws from policy science theories such as advocacy coalitions and discourse coalitions to study how power shapes transition outcomes. Attention is given to the politics of governance, where power struggles between regime actors and innovators determine the pace and scope of transitions toward sustainability, with implications for equity and justice.

3. Governing transitions — approaches and frameworks agenda examines the governance of sustainability transitions, highlighting the role of multiple actors, from governments to civil society, in shaping the path forward. Transition management (TM) and strategic niche management (SNM) offer frameworks for guiding transitions, promoting collaboration across sectors. It emphasises the importance of governance mechanisms that can deal with the uncertainties and complexities of transitions. Recent developments in governance focus on the role of transition arenas, policy mixes, and the experimental approaches needed to address sustainability challenges at local, national, and international levels.

4. Integrating civil society, culture, and movements in transitions agenda recognises the critical role of civil society and social movements in driving sustainability transitions. These groups influence industrial change by advocating for policies, providing protective spaces for innovation, and challenging existing cultural values. Research in this field examines crowdsourced innovation assessment and management, the role of social movements in shaping public policies, and the cultural shifts necessary to support sustainability. Civil society's efforts often lead to societal transformations, influencing consumer preferences, policy frameworks, and broader cultural norms, thereby promoting a transition towards more sustainable practices.

5. Engaging organisations and industries in transitions agenda investigates the role of firms and industries in sustainability transitions, focusing on how businesses drive innovation, market formation, and the reorientation of industries towards more sustainable practices. Industries not only develop new technologies and services but also influence policy and public opinion through lobbying, discourse, and framing. This research highlights the tensions between incumbents, who may resist change, and new entrants, who drive radical innovations. Understanding how businesses interact with institutional changes, create legitimacy for new technologies, and contribute to the emergence of new industries is vital for effective transition strategies.

6. Exploring transitions in practice and everyday life agenda explores how sustainability transitions unfold in everyday practices, including consumption and lifestyle changes. It examines the role of users in innovation and transition processes, focusing on how individuals and communities adopt, adapt, and advocate for sustainable technologies. Practice theory is applied to understand the persistence of resource-intensive behaviours and the potential for transformative shifts in daily life. Research in this area investigates the social dynamics of consumption, user involvement in innovation, and the impact of everyday practices on broader sustainability transitions from the household level to society at large.

7. Understanding the geography of transitions — spaces and scales. The geography of transitions examines how sustainability transitions vary across different regions and scales. It highlights the role of local institutions, natural resources, and regional networks in shaping transitions, as well as the transfer of innovations between places. Research in this area explores how different geographical contexts-rural, urban, or global-facilitate or hinder the development and diffusion of sustainable practices. It also considers how global networks and local experiments interact, with particular attention to urban transitions

and the role of cities in driving sustainability through experimentation and innovation.

8. Ethical considerations in transitions — justice and equity agenda focuses on the ethical dimensions of sustainability transitions, addressing concerns of justice, distribution, and poverty. It explores how transitions can perpetuate or alleviate inequalities, considering the impacts of technological and social innovations on vulnerable communities. Research highlights the importance of ensuring that transitions are just, inclusive, and equitable, ensuring that marginalised groups benefit from sustainable development. It also examines the ethical implications of policy decisions, including the distribution of costs and benefits, and calls for a more explicit integration of justice considerations into sustainability transition frameworks.

9. Reflecting on methodologies for transitions research agenda reflects on the methodologies used in sustainability transitions research, emphasising the need for methodological pluralism. It addresses key dilemmas such as balancing in-depth case studies with the search for broader, generic insights, and reconciling micro-level investigations with macro-level frameworks. The agenda also considers the complexity of transitions, arguing for approaches that can capture the non-linear dynamics of system change. Methodological advancements include the integration of qualitative and quantitative methods, the use of comparative case studies, and the exploration of transdisciplinary approaches that engage with policy and real-world experiments to catalyse transitions.

### 3.1.3. European Green Deal agendas

The ten European Green Deal (EGD) agendas outline a comprehensive framework for achieving climate neutrality and sustainability across the EU. In addition to the original eight priorities, two more agendas have been added to reflect the European Commission's increasing emphasis on sustainable finance and the internationalisation of climate diplomacy, positioning the EU as a global leader in climate action. These agendas highlight the EU's role in fostering clean energy transitions, reducing pollution, preserving ecosystems, and promoting sustainable mobility and food systems.

1. Becoming climate-neutral by 2050 for Europe. The European Green Deal (EGD) sets the ambitious target of achieving a net-zero carbon European Union by 2050, with interim goals including a 55% reduction in greenhouse gas emissions by 2030. This transformation demands an overhaul of policies, including new legislation to strengthen climate targets across various sectors. Achieving climate neutrality requires broad participation from all societal actors, including businesses, governments, and individuals, ensuring a comprehensive transformation that spans the economy and everyday life for all EU citizens.

2. Driving a clean and efficient energy transition. The Green Deal prioritises the transition to renewable energy sources, such as wind and solar, to reduce the EU's dependence on fossil fuels. Aiming for affordable, clean, and secure energy for all EU citizens, this agenda addresses energy poverty while promoting energy efficiency across member states. Since energy production accounts for over 75% of the EU's greenhouse gas emissions, decarbonising the energy system is crucial to meet the EU's 2030 climate objectives and the long-term goal of carbon neutrality by 2050.

3. Mobilising industry for a clean circular economy. The Green Deal promotes a circular economy model that decouples economic growth from resource consumption. By mobilising industries and integrating green technologies, it aims to foster sustainable, job-creating economic models that align with environmental objectives. Transitioning to a circular economy will reduce pressure on natural resources, halt biodiversity loss, and stimulate sustainable growth. Local governments and small businesses have significant roles in scaling these models, benefiting both citizens and the local economies, thus contributing to the EU's 2050 climate neutrality goal.

4. Building energy- and resource-efficient structures. Improving energy performance in buildings is a key focus of the European Green Deal, aiming to reduce energy consumption and emissions. Renovating existing buildings and constructing new, energy-efficient structures is vital for meeting climate targets. With buildings accounting for approximately 40% of Europe's energy use, the need for energy-efficient solutions is critical. The Renovation Wave Strategy, launched in 2020, addresses this challenge, particularly in retrofitting 85% of European buildings built before 2001, engaging citizens in the transformation to improve sustainability across Europe's housing stock.

5. Reducing pollution for a toxic-free environment. The European Green Deal aims to cut pollution across air, water, and soil, ensuring a toxic-free environment for all. It focuses on strengthening industrial

pollution controls and enhancing regulations to reduce harmful emissions. Pollution remains a significant environmental and health issue, with detrimental effects on public health, especially for vulnerable populations. The EU's Zero Pollution Action Plan seeks to address air quality, plastic waste, chemical pesticide use, and more in collaboration with citizens to create a cleaner, healthier, and more sustainable European environment.

6. Preserving ecosystems and restoring biodiversity. The Green Deal underscores the need to protect biodiversity and reverse the loss of ecosystems. The EU's Biodiversity Strategy for 2030 aims to safeguard natural habitats, conserve species, and restore damaged ecosystems, including through the creation of protected areas. Biodiversity plays a fundamental role in food production, health, and climate regulation. With the goal of protecting 30% of European land, restoring ecosystems on land and at sea is essential for ensuring sustainable growth and enhancing public awareness of the importance of biodiversity preservation.

7. Transforming food systems for sustainability. The Farm to Fork Strategy focuses on creating a sustainable, healthy, and fair food system across Europe. This agenda promotes sustainable farming practices, reduces food waste, and aims to enhance food security. The food sector, which accounts for about one-third of global greenhouse gas emissions, must evolve to meet both environmental and public health goals. The EU's strategy addresses these challenges, using new technologies, scientific discoveries, and increased public demand for sustainable food to create a food system that benefits all stakeholders, from producers to consumers.

8. Accelerating the shift to sustainable mobility. A central objective of the European Green Deal is to reduce emissions from transport, a significant contributor to EU greenhouse gas emissions. This agenda focuses on promoting electric vehicles, alternative fuels, and sustainable public transport systems to create a more connected, low-carbon mobility system. Urban mobility is specifically targeted through the European Urban Mobility Framework, which proposes measures for cities to develop more sustainable transport solutions. Achieving climate neutrality by 2050 requires reducing transportation emissions and enhancing the sustainability of urban transport networks across Europe.

9. Leveraging sustainable finance for green goals. Sustainable finance is essential for supporting the Green Deal's goals, ensuring that investments align with climate and environmental objectives. The EU aims to mobilise substantial funds through initiatives such as InvestEU, with at least 25% of the EU budget dedicated to climate change mitigation. This agenda includes incorporating environmental, social, and governance (ESG) considerations into financial decisions, promoting long-term investments in sustainable economic activities. By addressing climate change, pollution, and social inequalities, sustainable finance aims to support the EU's transition to a greener, more inclusive economy with transparent risk management strategies.

10. Leading global climate action and diplomacy efforts. The European Green Deal positions the EU as a global leader in climate action, promoting international cooperation and climate diplomacy to address global sustainability challenges. By strengthening partnerships with neighbouring regions like Africa and the Mediterranean, the EU aims to foster green investments and lead multilateral climate negotiations. The EU's external strategy involves a blend of collaborative, coercive, and diplomatic approaches, ensuring that the global green transition is inclusive, just, and aligned with Europe's climate-neutral objectives by 2050, encouraging widespread international commitments to climate resilience.

### 3.1.4. Comparative analysis: quadruple helix stakeholder roles in sustainability agendas

The comparative matrix presented below assesses the role and intensity of the four key stakeholders - government, business, academia, and civil society — across the ten CASI SI agendas, nine sustainability transition agendas, and ten European Green Deal agendas. The purpose of this analysis is to capture the extent of each stakeholder's involvement in advancing the respective agendas, focusing on their unique contributions to sustainable innovation, transitions, and policy formulation.

The quadruple helix model, which encompasses government, business, academia, and civil society, is used to evaluate the level of engagement of each stakeholder group in the implementation and impact of the sustainability agendas. This approach acknowledges that sustainable transitions require the active participation of multiple actors who can offer diverse perspectives and expertise. Governments are typically responsible for policymaking, regulatory frameworks, and the establishment of incentives; businesses drive innovation and market solutions; academia

Tab. 1. Quadruple helix stakeholder comparative matrix

| AGENDA | GOVERNMENT | BUSINESS | ACADEMIA | CIVIL SOCIETY |
|---|---|---|---|---|
| **SIA 1. Tackling climate change and reducing greenhouse gas emissions** | **High**. Governments set policies, regulations, and frameworks for climate action. | **Medium**. Businesses implement low-carbon practices, adopting green tech. | **High**. Academia drives research and innovation in climate solutions. | **High**. Civil society advocates, mobilises, and ensures public support. |
| **SIA 2. Enhancing governance through foresight and data-driven intelligence** | **High**. Governments drive foresight planning and policymaking. | **Medium**. Businesses provide data and contribute to technological solutions. | **High**. Academia conducts research to refine foresight methodologies. | **Low**. Limited role in foresight processes, mainly a receiver of policy. |
| **SIA 3. Building integrated sustainable infrastructures for a circular bioeconomy** | **High**. Governments provide regulatory frameworks and financial incentives. | **High**. Businesses innovate, provide technology, and scale sustainable practices. | **High**. Academia leads research on circular economy models and best practices. | **Medium**. Civil society participates in local solutions, demanding sustainability. |
| **SIA 4. Advancing responsible environmental management and resource efficiency** | **High**. Governments create policies for resource management and regulation. | **High**. Businesses apply resource-efficient practices and innovate for sustainability. | **Medium**. Academia develops tools and technologies but has limited practical reach. | **High**. Civil society influences and demands sustainable resource management. |
| **SIA 5. Accelerating the transition to renewable energy and biofuels** | **High**. Governments set policies, regulations, and financial mechanisms for energy transition. | **High**. Businesses lead renewable energy production, biofuel innovation, and distribution. | **High**. Academia conducts research and provides technological breakthroughs. | **Medium**. Civil society supports transitions and advocates for energy equity. |
| **SIA 6. Innovating sustainable transport and mobility systems** | **High**. Governments create transport policies, subsidies, and infrastructure. | **High**. Businesses develop and commercialise electric vehicles and sustainable transport solutions. | **Medium**. Academia provides innovative solutions and evaluates sustainable transport impacts. | **Medium**. Civil society participates in public campaigns and adopts new mobility solutions. |
| **SIA 7. Advancing circular economy and waste resource management** | **High**. Governments regulate waste management and incentivise circular practices. | **High**. Businesses innovate circular supply chains and recycling processes. | **High**. Academia conducts research on efficient waste management systems. | **High**. Civil society demands improved waste management and engages in recycling. |
| **SIA 8. Enhancing eco-community resilience and collaborative development** | **High**. Governments create policies and support initiatives for community resilience. | **Medium**. Businesses support sustainable development projects and local economies. | **Medium**. Academia explores and evaluates resilience models for communities. | **High**. Civil society organises grassroots movements and empowers local communities. |
| **SIA 9. Embedding sustainability in education and skills development** | **High**. Governments promote policies and curricula for sustainability education. | **Low**. Businesses offer some training but have limited influence on education. | **High**. Academia is central to integrating sustainability into education systems. | **Medium**. Civil society engages in educational campaigns and promotes learning. |
| **SIA 10. Promoting sustainable agriculture and localised food systems** | **High**. Governments provide subsidies, policy support, and frameworks for sustainable agriculture. | **High**. Businesses innovate in sustainable farming technologies and local food systems. | **High**. Academia leads research on sustainable farming practices and food systems. | **High**. Civil society supports local agriculture and campaigns for food system transformation. |
| **STA 1. Understanding transitions — theoretical frameworks and insights** | **Medium**. Governments implement transition strategies based on theoretical insights. | **Medium**. Businesses adapt models from theoretical frameworks for innovation. | **High**. Academia develops and refines theoretical frameworks for sustainability transitions. | **Low**. Limited role, primarily as beneficiaries of academic research outputs. |
| **STA 2. Examining power and politics in transitions** | **High**. Governments play key roles in managing power dynamics through policy. | **Medium**. Businesses influence transitions through lobbying and political power. | **Medium**. Academia explores power dynamics and contributes research on sustainability transitions. | **High**. Civil society advocates for just transitions and political accountability. |
| **STA 3. Governing transitions — approaches and frameworks** | **High**. Governments govern and coordinate transitions at various levels. | **Medium**. Businesses adapt to governance changes, integrating sustainable practices. | **High**. Academia contributes to governance models and framework development. | **Medium**. Civil society engages in the consultation processes for governance models. |

| | | | | |
|---|---|---|---|---|
| **STA 4. Integrating civil society, culture, and movements in transitions** | **Medium**. Governments collaborate with civil society to implement cultural shifts. | **Low**. Businesses may support but have minimal involvement in cultural and social movements. | **Medium**. Academia studies and supports the integration of culture into sustainability transitions. | **High**. Civil society drives cultural shifts, policy advocacy, and community mobilisation. |
| **STA 5. Engaging organisations and industries in transitions** | **High**. Governments provide incentives and policies for business involvement. | **High**. Businesses are central to innovation, industrial shifts, and market-driven transitions. | **Medium**. Academia provides research but is more of a supporting player in this agenda. | **Low**. Civil society has a limited direct role in business and organisational transitions. |
| **STA 6. Exploring transitions in practice and everyday life** | **Medium**. Governments create enabling policies for sustainable practices at the individual level. | **Medium**. Businesses influence consumer behaviour and support everyday sustainability solutions. | **High**. Academia studies practices and provides insights on sustainable living. | **High**. Civil society is central to implementing everyday sustainability through grassroots actions. |
| **STA 7. Understanding the geography of transitions — spaces and scales** | **High**. Governments define spatial policies and regional planning for transitions. | **Medium**. Businesses adapt to different regional contexts and manage localised supply chains. | **Medium**. Academia investigates regional transitions and how spatial scales impact sustainability. | **Medium**. Civil society plays a role in implementing local initiatives but is less influential at higher scales. |
| **STA 8. Ethical considerations in transitions — justice and equity** | **High**. Governments define policies to ensure equity and fairness in transitions. | **Medium**. Businesses must align with ethical standards and address equity in operations. | **Medium**. Academia contributes research on the ethics of transitions and equity considerations. | **High**. Civil society ensures that ethical concerns and justice are central to sustainability. |
| **STA 9. Reflecting on methodologies for transitions research** | **Low**. Limited government involvement in the development of research methodologies. | **Low**. Business involvement is minimal; primarily academic-driven methodologies. | **High**. Academia leads methodological innovations for studying transitions. | **Low**. Civil society is primarily a beneficiary of the methodologies developed. |
| **EGD 1. Becoming climate-neutral by 2050 for Europe** | **High**. Governments set ambitious targets, policies, and regulatory frameworks. | **Medium**. Businesses implement solutions but are heavily influenced by government policies. | **Medium**. Academia supports policy development and technological innovations. | **Low**. Civil society's role is largely as supporters, less directly involved in policy creation. |
| **EGD 2. Driving a clean and efficient energy transition** | **High**. Governments set energy transition goals and implement regulations and incentives. | **High**. Businesses provide renewable energy solutions, technologies, and infrastructure. | **Medium**. Academia provides energy-related research and evaluates transition impacts. | **Medium**. Civil society supports the transition and advocates for energy access equity. |
| **EGD 3. Mobilising industry for clean circular economy** | **High**. Governments create frameworks for industry participation and regulation. | **High**. Businesses are key players in developing and scaling circular economy models. | **High**. Academia supports research on best practices and business models for circularity. | **Medium**. Civil society supports through advocacy for reduced waste and increased recycling. |
| **EGD 4. Building energy- and resource-efficient structures** | **High**. Governments set building codes and policies for energy-efficient structures. | **High**. Businesses innovate in construction and energy-efficient technologies. | **Medium**. Academia contributes through research on energy-efficient materials and systems. | **Low**. Civil society's role is limited to adopting energy-efficient practices in buildings. |
| **EGD 5. Reducing pollution for a toxic-free environment** | **High**. Governments regulate pollution, set limits, and enforce compliance. | **Medium**. Businesses are affected by regulations but also innovate in pollution reduction. | **Medium**. Academia studies pollution reduction technologies and their effectiveness. | **High**. Civil society advocates for pollution reduction and public health safety. |
| **EGD 6. Preserving ecosystems and restoring biodiversity** | **High**. Governments create policies, protected areas, and biodiversity strategies. | **Medium**. Businesses help through sustainability practices and biodiversity-friendly innovations. | **High**. Academia conducts research and provides strategies for ecosystem restoration. | **High**. Civil society plays a key role in advocacy, conservation efforts, and education. |
| **EGD 7. Transforming food systems for sustainability** | **High**. Governments regulate food systems and set policies for sustainability. | **High**. Businesses play a major role in transitioning the food sector towards sustainability. | **Medium**. Academia contributes research on sustainable farming and food systems. | **High**. Civil society pushes for food security, health, and sustainable agriculture. |
| **EGD 8. Accelerating the shift to sustainable mobility** | **High**. Governments create policies and infrastructure for green mobility solutions. | **High**. Businesses provide electric vehicles, sustainable transport technologies, and services. | **Medium**. Academia researches transportation systems and sustainable mobility. | **Medium**. Civil society advocates for equitable mobility and the adoption of clean technologies. |

| | | | | |
|---|---|---|---|---|
| **EGD 9. Leveraging sustainable finance for green goals** | **High**. Governments regulate green investments and provide financial mechanisms. | **High**. Businesses attract investments and drive green economic growth. | **Medium**. Academia contributes research on sustainable finance models and investment strategies. | **Low**. Civil society's role is more indirect, advocating for ethical investments. |
| **EGD 10. Leading global climate action and diplomacy efforts** | **High**. Governments take the lead in international negotiations and commitments. | **Medium**. Businesses align with international climate policies but may resist strict measures. | **Low**. Academia supports through research but plays a secondary role in diplomacy. | **Low**. Civil society plays a minimal role in international climate diplomacy efforts. |

provides research, analysis, and new technological developments; and civil society plays a crucial role in advocacy, public mobilisation, and ensuring accountability.

The comparative analysis highlights several key patterns. First, it is evident that governments have a high level of involvement in nearly all the agendas, as they set regulatory frameworks and policies and provide financial incentives. Business stakeholders, often second in influence, are central to scaling innovations, particularly in areas like renewable energy, circular economies, and sustainable agriculture. Academia's role, although significant, is mostly focused on research, technological advancements, and theoretical frameworks, which provide essential guidance for policy and business practices. Civil society, while influential in many areas, particularly in advocating for sustainability and raising awareness, often has a secondary role in the formal implementation of policy agendas.

The matrix further reveals that certain agendas require more collaborative involvement, such as those related to community resilience, circular economy, and sustainable mobility, where the active participation of all four stakeholders is essential for achieving long-term sustainability. For example, the "Advancing circular economy and waste resource management" agenda requires high involvement from both government and business for regulation and innovation, while academia and civil society contribute through research and public advocacy. Conversely, agendas like "Enhancing governance through foresight and data-driven intelligence" see a more prominent role for governments and academia, with limited direct involvement from business and civil society.

In conclusion, the comparative matrix offers valuable insights into the varying degrees of stakeholder involvement in each agenda, highlighting both synergies and potential gaps. It underlines the importance of fostering greater collaboration across these stakeholder groups, particularly in areas where grassroots-driven innovation and top-down policy frameworks intersect. The analysis sets the stage for further exploration into how these stakeholders can work more effectively together to advance the goals of the European Green Deal and other sustainability agendas.

### 3.2. Key features of each agenda

While there are notable alignments between the CASI SI agendas, the transitions research agendas, and the European Green Deal agendas, several gaps and differences have been identified across their focus areas. These gaps highlight opportunities for further integration and refinement of the Green Deal's approach. By addressing these discrepancies, the Green Deal could benefit from a more holistic and inclusive strategy, ensuring a more participatory and comprehensive transition towards sustainability. This would not only strengthen the Green Deal's objectives but also enhance its alignment with both crowd-sourced innovation management and the broader societal shifts necessary for a sustainable future.

- Sustainable innovation agendas (Popper et al., 2017a, 2017b, 2017c). The SI agendas are distinguished by their strong focus on quadruple helix and ecosystem-oriented innovation approaches, fostering collaboration across diverse sectors. They prioritise the importance of localised efforts, such as fostering eco-community empathy, crowd-funded development, and eco-local agriculture. The SI agendas focus on practical, bottom-up innovations that involve local communities and small-scale initiatives, thereby addressing sustainability from a more regional and community-specific perspective. This approach complements the more top-down, policy-driven frameworks seen in other agendas. Additionally, SI agendas place a high value on the integration of sustainability across sectors such as bioeconomy and urban infrastructures and promoting sustainability in cultural and educational contexts.

- Sustainable transition agendas (Köhler et al., 2019). Transition research agendas are distinct for their academic focus on the dynamics of sustainability transitions, especially the social, political, and cultural dimensions. These agendas investigate the role of power, governance, and social movements in facilitating or hindering transitions. They also explore ethical considerations in transitions, such as justice, equity, and poverty. This framework places a strong emphasis on understanding transitions as complex, multidimensional processes, with a focus on both theoretical and methodological advancements in transition research.
- European Green Deal agendas (EC, 2019a, 2019b, 2020). The European Green Deal agendas are unique in their overarching focus on achieving Europe's climate-neutral goals by 2050, positioning Europe as a global leader in sustainability transitions. These agendas strongly emphasise the transformation of the European economy, particularly in energy, transport, and agriculture. The Green Deal underscores the importance of achieving decarbonisation and resource efficiency at a large scale, with goals like circular economies, sustainable mobility, and clean energy. Moreover, the Green Deal highlights inclusivity and social justice in the transition, ensuring that no region or community is left behind in the shift towards sustainability.

### 3.3. Towards a harmonised agenda

CASI-F, in conjunction with CASIPEDIA, can make a substantial contribution towards harmonising the diverse sustainability agendas outlined here by serving as a central platform that integrates bottom-up innovations with top-down policy objectives. CASI-F's emphasis on community-driven projects and local innovations complements the European Green Deal's broader sustainability goals, fostering synergies between local and European initiatives. "When a technology is consumed as part of a system, the performance that matters to the consumers' assessment of value is not the performance of the focal technology on its own, but is rather a function of its interaction with the other elements of the system" (Adner & Kapoor, 2016). This highlights that the impact of local innovations, when integrated into broader systems, can be more pronounced than when considered in isolation. By systematically mapping and evaluating sustainable innovations through CASIPEDIA-like tools, these projects can be linked to wider global and European sustainability objectives, thereby facilitating better alignment with the SDGs.

Moreover, the application of foresight and anticipatory intelligence within CASI-F's governance framework can aid in assessing future challenges and opportunities, thereby offering a proactive approach to sustainability management. "The relative pace at which the new technology will substitute the old technology will depend on the joint levels of ecosystem emergence challenge for the new technology and the ecosystem extension opportunity for the old technology: fastest when both are low, slowest when both are high, and intermediate in the mixed case where one is high and one is low" (ibid.). This underscores the importance of creating optimal conditions for the co-evolution of technologies within a system, which is critical for accelerating transitions (Aarikka-Stenroos & Ritala, 2017). By collecting comprehensive data on regional innovations and practices, CASIPEDIA can provide invaluable insights into the effectiveness of these innovations, thereby contributing to ongoing policy development and guiding future research directions in sustainability transitions. Consequently, CASI-F and CASIPEDIA have the potential to serve as an effective bridge, uniting research, policy, and practice in a cohesive and integrated manner, thus fostering a more harmonised and dynamic sustainability agenda across Europe and beyond.

## Conclusion

Innovation ecosystems play a crucial role in advancing the objectives of the European Green Deal by offering pathways for systemic transformation that align technological innovation with broader social, economic, and institutional changes. These ecosystems facilitate multi-level collaboration, effectively integrating DIY and crowdsourced innovation management with top-down policies, which can accelerate Europe's transition to a climate-neutral and resource-efficient economy. The lessons derived from the CASI project and its sister BOLERO project provide valuable insights for refining the Green Deal's framework, ensuring that it remains not only ambitious but also inclusive and resilient in addressing the complex sustainability challenges we face today and in the future.

One core principle emerging from this study is the importance of adopting a global perspective while

acting locally. Although the European Green Deal presents a high-level vision for a climate-neutral Europe, it is at the local level where the most impactful changes often take place. DIY and crowdsourced initiatives – such as community-led energy cooperatives, decentralised agriculture, and locally driven circular economy projects – hold significant potential to drive meaningful change. These bottom-up solutions can be scaled up and integrated into national and regional frameworks, ultimately contributing to global sustainability objectives. By empowering local communities and fostering innovation, the Green Deal has the capacity to create a more inclusive, resilient, and adaptive transition.

The integration of local innovations into broader national and regional policies is essential for developing a flexible, responsive approach that addresses sector-specific and regional needs while still aligning with global sustainability goals. This integration promotes greater coherence between policy levels, enhancing the ability to address diverse challenges and ensuring that no region or community is left behind. Strengthening policy coherence across local, national, and EU levels is not only critical for the success of the Green Deal but also vital for driving a transition that is sustainable, equitable, and adaptable to future challenges. "The co-evolutionary logic shifts the focus from the evolution and exchange of a business network to system co-evolution... Ecosystems cross industry boundaries, connecting various types of business and innovation networks in an exchange that is embedded in an institutional and socio-technical environment" (ibid). A relevant example of this co-evolution is seen in the interaction between the CASIPEDIA and BOLERO platforms. These platforms act as bridges between European-level and regional innovation ecosystems focused on sustainability. By connecting local sustainability innovations in the BOLERO platform with the broader policy and research initiatives showcased in CASIPEDIA, they facilitate a dynamic exchange of knowledge, technologies, and best practices that fosters system-wide sustainability transitions.

Thus, this study reinforces the need for a more harmonised, multi-actor approach – blending DIY and crowdsourced innovation management, policy agendas, and cross-sectoral collaboration – ensuring that Europe can move forward as a global leader in sustainability and climate resilience.

Recommendations for policy enhancements: This section outlines key recommendations for enhancing the European Green Deal's effectiveness and inclusivity, with the goal of ensuring a transformative and sustainable transition to a climate-neutral future. The recommendations emphasise the integration of local, community-driven initiatives into broader policy frameworks, fostering an approach that is both adaptable and equitable. Key strategies include strengthening community engagement, advancing sustainability education, scaling DIY and crowdsourced innovation management, improving innovation mapping tools, and incorporating SMART foresight processes. By implementing these measures, Europe can create a more resilient, inclusive, and forward-looking sustainability strategy, empowering local communities while reinforcing global sustainability objectives.

1. Strengthening community engagement. To ensure that the Green Deal is inclusive and equitable, it is essential to develop and expand participatory mechanisms that actively engage local communities in the design and implementation of sustainability policies. Local communities possess invaluable knowledge and insights, which can help shape policies that are both effective and contextually relevant. By involving these communities in decision-making processes, policies will reflect diverse needs and foster greater ownership and accountability in the green transition. As Geels and Schot (2007) point out, "ecosystems cross industry boundaries, connecting various types of business and innovation networks in an exchange that is embedded in an institutional and socio-technical environment", which suggests that integrating local knowledge into larger systems can provide mutually beneficial outcomes, aligning community-driven innovations with broader policy frameworks.

2. Advancing sustainable innovation mapping. To optimise the European Green Deal's impact, it is crucial to further advance tools like CASIPEDIA and BOLERO, which systematically support the mapping and assessment of sustainable innovations across Europe. These platforms play a pivotal role in identifying, sharing, and scaling locally-driven initiatives, fostering collaboration and knowledge exchange. By combining qualitative and quantitative approaches, it would be possible to develop a more universal or hybrid framework for sustainable innovation (SI), which could help reduce fragmentation and divisions within the growing community of SI practitioners and scholars (Pajula & Popper, 2020). Expanding and enhancing these tools would strengthen the integration of DIY and crowdsourced innovation management into broader policy frameworks, ensuring that

Europe's sustainability efforts are both informed and inclusive. This hybrid approach, which bridges the boundaries across sectors and regions, can facilitate the alignment of local innovations with EU-wide policies. The integration of local innovations into national and regional policies is vital for developing a flexible, responsive approach that addresses sector-specific and regional needs while still contributing to global sustainability goals. The systematic assessment and management of SI provides unique and shared benefits to multiple stakeholders, making it accessible across different areas of science and industry and capable of delivering the results needed to address and meet current challenges. This underscores the importance of adapting local solutions to larger frameworks like the European Green Deal, ensuring that bottom-up, DIY-driven initiatives align with top-down policy objectives and contribute to a more cohesive, comprehensive transition.

3. Promoting sustainability education across Europe. Integrating sustainability education into all levels of the European education system — from primary schools to higher education — is essential for the long-term success of the Green Deal. By embedding eco-literacy and sustainability-focused curricula, Europe can equip future generations with the skills and knowledge required to actively contribute to a more sustainable future. This educational approach must also emphasise the ethical dimensions of sustainability, enabling citizens to understand the social, environmental, and economic implications of their actions. Education for sustainability should aim to foster a critical understanding of technology and social organisation, encouraging learners to reflect on how both can promote harmony between people and the natural world. As Huckle (1991) pointed out in one of the earliest works on this subject, such education must empower students to engage in democratic processes that can drive societal transformation. The development of new knowledge and values is crucial for building a broad-based support system for sustainability innovations, especially in the context of the co-evolution of niche innovations with societal values. Higher education institutions play a central role in preparing future leaders who will help achieve the United Nations Sustainable Development Goals (Žalėnienė & Pereira, 2021).

4. Scaling and integrating DIY and crowdsourced innovation management. The European Green Deal would greatly benefit from increased support for scaling successful DIY and crowdsourced innovation management practices, including grass-roots-level initiatives. Financial and technical support for community-driven projects — such as renewable energy cooperatives, urban farming, and circular economy ventures — will enable these local initiatives to contribute more effectively to national sustainability goals and create scalable models for broader implementation. By embedding these innovations into broader policy frameworks, Europe can ensure that sustainability strategies remain adaptable to diverse regional needs while reinforcing the global sustainability agenda. This approach can help foster resilient, adaptive systems that leverage local knowledge and resources for maximum impact. In line with Aarikka-Stenroos and Ritala's (2017) focus on value co-creation, facilitating the scaling of DIY and crowdsourced innovation management will foster collaborative relationships between multiple stakeholders, thus enhancing the overall value generated by these local projects and aligning them more closely with broader sustainability objectives. As noted by Callaghan (2016) and Popper et al. (2025), this approach is particularly valuable in crisis management, where geopolitical instability and unpredictable events, or "wild cards", are increasingly shaping every aspect of innovation ecosystems. These disruptions require more agile and responsive frameworks that can draw upon a diverse range of inputs from crowdsourced R&D and social media platforms. By integrating real-time problem-solving and probabilistic innovation, Europe and other world regions can better navigate these complexities, enabling faster, more collaborative solutions to urgent global challenges.

5. Adopting SMART foresight for future-oriented policymaking. The adoption of SMART foresight processes (Popper, 2011, 2012) offers significant potential to enhance the effectiveness of the European Green Deal by providing a structured framework for future-oriented policymaking. This approach guides decision-making through distinct phases — scoping, mobilising, anticipating, recommending, and transforming — which together create a dynamic cycle of strategic foresight. Integrating this foresight methodology (Popper, 2008) into the Green Deal will enable the European Union to ensure its policies are both forward-looking and flexible, capable of responding proactively to emerging challenges and uncertainties. This approach provides a comprehensive roadmap for the sustainability transition, helping to identify potential risks, set clear priorities, and mobilise a diverse range of stakeholders to implement effective sustainability solutions. Through the creation of a "space" for discourse, analysis, and creative vision-

ing, foresight empowers policymakers to consider long-term implications and explore innovative pathways that might otherwise be overlooked (Miles et al., 2008; Velasco et al., 2021). Furthermore, by combining foresight with other strategic policy intelligence tools such as evaluation and technology assessment, the EU can ensure that the Green Deal is grounded in actionable, evidence-based strategies that meet current needs while being resilient to future disruptions. The integration of SMART foresight ensures that the Green Deal is adaptive, resilient, and capable of responding to evolving environmental, social, and technological contexts, ultimately fostering a more sustainable and inclusive future for Europe.

By implementing these recommendations, regions worldwide can accelerate their transitions to a climate-neutral future that is transformative, inclusive, adaptive, and resilient. Empowering citizens and communities to actively contribute to a sustainable future will be crucial for achieving long-term climate objectives. The integration of bottom-up innovations and local solutions into national, regional, and global frameworks will foster a sustainable, just transition that addresses the diverse needs of communities while reinforcing global sustainability goals.

Further research and a way forward: As the European Green Deal strives to meet its ambitious sustainability targets, future research should prioritise enhancing emerging frameworks that underpin sustainable innovation. Building on lessons from established tools such as CASI-F and BOLERO, integrating the complementary Candy Innovation Model, developed at La Salle – Ramon Llull University, offers a promising avenue for advancing both methodology and impact. With its structured approach spanning four key phases — challenges, ideas, development, and scalability — the model aligns closely with the principles of CASI-F and BOLERO, particularly in fostering multi-level stakeholder engagement and supporting scalable innovation solutions. As Pique et al. (2017) underscore, the model's emphasis on addressing real-world challenges through collaborative problem-solving provides invaluable insights for navigating complex sustainability issues, ensuring that local innovations are developed, tested, and scaled effectively in alignment with broader policy frameworks.

Originally developed to address regional innovation needs in Catalonia, Spain, the approach has since expanded to other Spanish regions and is being adopted across Latin America through universities within the La Salle network. This broad implementation demonstrates the model's adaptability across diverse geographical and cultural contexts. Future research should explore how this methodology can be integrated into national and EU-level policy frameworks to enhance sustainability efforts.

Additionally, research should address the scalability potential of sustainability projects within the EU context. Integrating this model into CASI-F presents an opportunity for enhanced cross-border collaboration, pooling expertise from industry, academia, government, and civil society. This integration ensures that local solutions to sustainability challenges are effective and adaptable across regions, aligning innovations with EU-wide sustainability targets. The model's emphasis on testing, prototyping, and real-time feedback from diverse stakeholders will help develop policies that are both forward-looking and resilient to global challenges. Through continued adaptation and refinement, Europe can establish a more integrated, adaptive, and inclusive framework for achieving sustainability, empowering local communities to contribute meaningfully to the broader transformation envisioned by the European Green Deal.

## Literature

Aarikka-Stenroos, L., & Ritala, P. (2017). Network management in the era of ecosystems: Systematic review and management framework. *Industrial marketing management*, *67*, 23-36.

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic management journal*, *31*(3), 306-333.

Adner, R., & Kapoor, R. (2016). Innovation ecosystems and the pace of substitution: Re-examining technology S-curves. *Strategic management journal*, *37*(4), 625-648.

Callaghan, C. W. (2016). Disaster management, crowdsourced R&D and probabilistic innovation theory: Toward real time disaster response capability. International journal of disaster risk reduction, *17*, 238-250.

European Commission. (2019a). *The European Green Deal. Communication from the Commission. COM(2019) 640 final*. Brussels, 11.12.2019.

European Commission. (2019b). *Annex to the Communication on The European Green Deal. COM(2019) 640 final*. Brussels, 11.12.2019.

European Commission. (2020). *Sustainable Europe Investment Plan. European Green Deal Investment Plan. Communication from the Commission. COM(2020) 21 final*. Brussels, 14.01.2020.

Geels, F. W., & Schot, J. (2007). Typology of sociotechnical transition pathways. *Research policy*, *36*(3), 399-417.

Granstrand, O., & Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. *Technovation*, *90*, 102098.

Hereu-Morales, J., Segarra, A., & Valderrama, C. (2024). The European (Green?) Deal: a systematic analysis of environmental sustainability. *Sustainable Development*, *32*(1), 647-661.

Huckle, J. (1991). Education for sustainability: Assessing pathways to the future. *Australian Journal of Environmental Education*, *7*, 43-62.

Kivimaa, P., & Kern, F. (2016). Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions. *Research policy*, *45*(1), 205-217.

Koundouri, P., Alamanos, A., Plataniotis, A., Stavridis, C., Perifanos, K., & Devves, S. (2024). Assessing the sustainability of the European Green Deal and its interlinkages with the SDGs. *Npj Climate Action*, *3*(1), 23.

Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., Alkemade, F., Avelino, F., Bergek, A., Boons, F., Fünfschilling, L., Hess, D., Holtz, G., Hyysalo, S., Jenkins, K., Kivimaa, P., Martiskainen, M., McMeekin, A., Mühlemeier, M., Nykvist, B., Pel, B., Raven, R., Rohracher, H., Sandén, B., Schot, J., Sovacool, B., Turnheim, B., Welch, D., & Wells, P. (2019). An agenda for sustainability transitions research: State of the art and future directions. *Environmental innovation and societal transitions*, *31*, 1-32. doi: 10.1016/j.eist.2019.01.004

Kwilinski, A., Lyulyov, O., Pimonenko T. (2024). Sustainable development in the european union: managing artificial intelligence technology for green economic growth. *Polish Journal of Management Studies*, *30*(1), 187-207. doi: 10.17512/pjms.2024.30.1.11

Markard, J., & Truffer, B. (2008). Technological innovation systems and the multi-level perspective: Towards an integrated framework. *Research policy*, *37*(4), 596-615.

Martini, M., Holsgens, R., & Popper, R. (2020a). Governance and Management of Sustainable Innovation: Learning from Experience to Shape the Future. Berlin, Germany: Springer.

Martini, M., Marafioti, E., & Carminati, M. (2020b) Exploring the Role of Stakeholder Engagement for the Development of Sustainable Innovation Strategies. In M. Martini, R. Holsgens, R. Popper (Eds.), *Governance and Management of Sustainable Innovation: Learning from Experience to Shape the Future*. Berlin, Germany: Springer.

Miles, I., Harper, J. C., Georghiou, L., Keenan, M., & Popper, R. (2008). New Frontiers: Emerging Foresight. The Handbook of Technology Foresight: concepts and practice. *Innovation, Strategy and Sustainability*, 400.

Oh, D. S., Phillips, F., Park, S., & Lee, E. (2016). Innovation ecosystems: A critical examination. *Technovation*, *54*, 1-6.

Pajula, T., & Popper, R., (2020) Towards a Hybrid Framework for Sustainable Innovation. In M. Martini, R. Holsgens, R. Popper (Eds.), *Governance and Management of Sustainable Innovation: Learning from Experience to Shape the Future*. Berlin, Germany: Springer.

Piqué, J. M., Marcet, X., Majó, A. & Miralles, F. (2017). Candy Innovation Model: Managing the Deal Flow of Innovation. In *XXXIV IASP World Conference on Science and Technology Parks*. Istanbul, Turkey.

Popper, R. (2008). How are foresight methods selected?. *foresight*, *10*(6), 62-89.

Popper, R. (2009). Mapping Foresight: Revealing how Europe and other world regions navigate into the future. *Publications Office of the European Union*, 126.

Popper, R. (2011). *21st Century Foresight (PhD Thesis)*. The University of Manchester, Manchester.

Popper, R. (2012). Mapping futures studies. *Foresight and STI Governance*, *6*(2), 56-74.

Popper, M., Tregner-Mlinaric, A., Popper, R., Velasco, G., Schwarz-Woelzl, M., Van Eynde, S., Ramioul, M., Damianova, Z., Kozarev, V., Martini, M., Hölsgens, R., & Schultze, J. (2017a). Sustainable innovation policy advice using a quadruple helix approach to "innovations" mapping. In R. Popper, G. Velasco (Eds.), *Sustainable Innovation Policy Advice* (pp. 28-49). Brussels, Belgium: European Commission.

Popper R. & Popper M. (2024). Action Roadmaps for More Resilient Research and Innovation Futures: Strategic Pathways to Foresight-Driven and Sustainable R&I Policies in FP10 (CFI Working Paper Series No. 2024/02). Warsaw: Technology Partners.

Popper R. & Popper M. (2025). Exploring the Future of Sustainability Through Sustainable Innovations in Eco-Friendly Business Ventures, Redefining Innovation and Entrepreneurship for 21st Century Ventures. *Opolo Global Innovation*, 799.

Popper, R., Velasco, G. & Popper, M. (2017b). CASI-F: Common Framework for the Assessment and Management of Sustainable Innovation, CASI project report. Deliverable 6.2.

Popper, R., Popper, M., & Velasco, G. (2017c). Towards a more responsible sustainable innovation assessment and management culture in Europe. *Engineering Management in Production and Services*, *9*(4), 7-20. doi: 10.1515/emj -2017-0027

Popper, R., Popper, M., & Velasco, G. (2020). Sustainable Innovation Assessment and Management Framework: Principles, Methodology and Practice. In M. Martini, R. Holsgens, R. Popper (Eds.), *Governance and Management of Sustainable Innovation: Learning from Experience to Shape the Future*. Berlin, Germany: Springer.

Popper, R., Villarroel, Y., & Popper, R. W. (2025). Towards a Sustainable Disruptive Growth Model: Integrating Foresight, Wild Cards and Weak Signals Analysis. *Foresight and STI Governance*, *19*(1), 32-49.

Szpilko, D., & Ejdys, J. (2022). European Green Deal–research directions. a systematic literature review. *Ekonomia i Środowisko*, *2*, 8-38. doi: 10.34659/eis.2022.81.2.455

Velasco, G., Popper, M., & Popper, R. (2020). An Inductive and Multidimensional Approach to Sustainable Innovation: Evidence from Multiple Case Studies. *Governance and Management of Sustainable Innovation: Learning from Experience to Shape the Future*, 181-198. doi: 10.1007/978-3-030-46750-0_9

Velasco G., Popper R., & Miles I. (2021). Repositioning People in Creative Futures: A Method to Create Sound Advice with Exploratory Scenarios. *Foresight and STI Governance*, *15*(2), 25-38. doi: 10.17323/2500-2597.2021.2.25.38

Žalėnienė, I., & Pereira, P. (2021). Higher education for sustainability: A global perspective. *Geography and Sustainability*, *2*(2), 99-106. doi: 10.1016/j.geosus.2021.05.001

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 27 August 2024
accepted: 30 April 2025

pages: 23-36



# Trends and challenges of forecasting in the airline industry research

Oki Anita Candra Dewi
Nur Aini Masruroh
Budhi Sholeh Wibowo

ABSTRACT
This study aims to comprehensively review aviation forecasting research by identifying its bibliometric trends, evolving research areas, and thematic developments. It focuses on understanding the aviation industry's research gaps, highlighting emerging trends, and offering insights into future forecasting innovations. A systematic literature review in the Scopus database used Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) and bibliometric analysis. It identified key patterns, influential publications, and emerging topics. A science mapping analysis was executed to pinpoint research trends in airline forecasting using Biblioshiny to visualise the network analysis and thematic evolution keywords mapping. The study categorised research trends and identified underexplored areas for future investigation. The findings reveal significant shifts in aviation forecasting research, with three distinct phases of publication growth and a surge in output from 2016 onwards. Passenger demand forecasting remains the most researched topic, though its growth has stabilised. Emerging issues such as customer behaviour, financial forecasting, and dynamic pricing have gained prominence, driven by advancements in machine learning and big data analytics. The study also highlights transitioning from traditional statistical methods to more advanced predictive techniques, emphasising real-time decision-making and operational efficiency. Established research areas, such as air cargo forecasting and fleet scheduling, have become more standardised, reducing the need for further innovation.

KEY WORDS
**airline, forecasting, bibliometric analysis, PRISMA**

**10.2478/emj-2025-0010**

**Oki Anita Candra Dewi**
Mechanical and Industrial Engineering Department, Universitas Gadjah Mada, Grafika Street No. 2, Sinduadi, Mlati District, Sleman Regency, Special Region of Yogyakarta, Indonesia
Universitas Internasional Semen Indonesia PT Semen Indonesia Complex, Veteran Street, Gresik, East Java, Indonesia
ORCID 0000-0002-6690-0453
Corresponding author:
e-mail: oki.dewi@uisi.ac.id

**Nur Aini Masruroh**
Mechanical and Industrial Engineering Department, Universitas Gadjah Mada, Grafika Street No. 2, Sinduadi, Mlati District, Sleman Regency, Special Region of Yogyakarta, Indonesia
ORCID 0000-0003-0171-7620
e-mail: aini@ugm.ac.id

**Budhi Sholeh Wibowo**
Mechanical and Industrial Engineering Department, Universitas Gadjah Mada, Grafika Street No. 2, Sinduadi, Mlati District, Sleman Regency, Special Region of Yogyakarta, Indonesia
ORCID 0000-0002-0448-8045
e-mail: budhi.sholehwibowo@ugm.ac.id

## Introduction

The airline industry is currently experiencing substantial growth in travel demand. International Air Transport Association (IATA) reports an annual compound average growth rate of approx. 3.7%. Projections indicate that by 2035, the global number of travellers is expected to soar to around 7.2 billion, nearly doubling the number of passengers recorded in 2016 (IATA, 2016). The aviation industry can address this projected growth by leveraging accurate forecasting across all operational and strategic levels. This remarkable increase in demand presents a significant

Dewi, O. A. C., Masruroh, N. A., & Wibowo, B. S. (2025). Trends and challenges of forecasting in the airline industry research. *Engineering Management in Production and Services, 17*(2), 23-36. doi: 10.2478/emj-2025-0010

challenge for the airline industry, as it requires precise forecasting to manage and optimise revenue expectations effectively.

Forecasting is fundamental to decision-making (Winkowski, 2019) and is critical across various aviation sectors. Accurate travel demand information is important for transportation systems, especially for the airline industry (Samli et al., 2021). In airline operations, accurate forecasting is essential for determining optimal flight frequencies, route allocations, and seat capacities to avoid overcapacity or shortages, thereby maximising load factors and improving overall profitability (Wang & Gao, 2021). Short-term decisions, such as flight schedules, pricing strategies, marketing campaigns, and crew rostering, rely heavily on demand forecasts. Long-term strategic planning, including fleet acquisition, emission, workforce allocation, resource investment, and network expansion, also depends on demand projections. Airport operators use demand forecasting as the foundation of airport master plans, which guide infrastructure expansion and significant capital investments to ensure sufficient capacity and operational efficiency (Sulistyowati et al., 2018). Policymakers rely on demand estimates to allocate budgets for air transportation and related sectors, while the aerospace industry depends on accurate forecasts for developing and manufacturing aircraft, engines, and components. Moreover, demand forecasts are critical in determining the stock market performance of airlines (Caiado & Lúcio, 2023), as investors and stakeholders closely monitor demand trends to assess airline companies' financial health and future profitability. Local communities also benefit from accurate air travel forecasts, as these estimates determine air transportation networks and regional connectivity.

Numerous research efforts have been devoted to addressing forecasting-related challenges in the aviation industry. Key areas of focus include predicting and mitigating flight delays (Li et al., 2023), optimising flight routing paths, addressing the impacts of weather and climate change, and improving spare parts inventory management and maintenance strategies to reduce operational disruptions. The growing emphasis on environmental sustainability has led researchers to explore demand forecasting for CO2 emissions reduction in the aviation sector (Huang et al., 2023).

Airlines were compelled to adopt new adaptive and robust revenue management strategies that continuously monitored key metrics to respond dynamically to market fluctuations (Vinod, 2022). These adjustments underscored the importance of incorporating flexible forecasting models that can adapt to sudden disruptions and evolving market conditions. Methods relying solely on historical data without traditional demand forecasting and optimisation have also been proposed to enhance forecasting accuracy in highly uncertain environments (van Ryzin & McGill, 2000).

Forecasting became incredibly challenging during the COVID-19 pandemic, as sudden disruptions rendered historical demand patterns less meaningful and necessitated new forecasting approaches. Studies have explored passenger demand variations across segments during the pandemic (Li et al., 2023) and the recovery of airfreight operations using time series analysis (İnan, 2022). Furthermore, Caiado and Lúcio (2023) proposed clustering methods to compare financial time series and analyse the pandemic's impact on the US stock market. Wang et al. (2020) studied predictors for international stock markets during the spread of COVID-19, while Ye et al. (2023) analysed the pandemic's impact on the airline industry. Their findings highlighted the importance of workforce planning and human resource management in adapting to the industry's disruptions.

Given the critical role of forecasting and the dynamic nature of the aviation industry, systematic reviews of relevant studies have become meaningful and necessary. Over the past decades, several significant efforts have been made to review and summarise advancements in this field to provide insights into effective forecasting practices and future research opportunities. Based on Zachariah et al. (2023), forecasting future demand for the aviation industry is critical, as it relies on a thorough understanding of various demand determinants. They also conducted a notable study evaluating aviation demand research and assessing the effectiveness of various forecasting techniques, providing valuable insights into the evolving landscape of demand modelling. Wang et al. (2022) provided a detailed review of the flight operation process, identifying flight delays as the most critical irregular flight issue. Their findings highlighted that delays are primarily caused by severe weather and other factors such as airline planning, air traffic control, airspace flow restrictions, and military activities, which also contribute to the problem.

Although extensive studies have been conducted, certain emerging and evolving areas in aviation forecasting remain underexplored. Addressing these gaps is the primary motivation for this literature review, which employs a bibliometric analysis to uncover key

trends and research opportunities. This work highlights key focus areas for airline managers to gain a competitive advantage by concentrating on emerging research trends. For areas where research is becoming less relevant or more standardised, practitioners can focus on applying established best practices to maintain operational efficiency and optimise decision-making. This study aims to provide a holistic view of forecasting research in the aviation industry, with specific contributions as outlined below:

- RQ1: Identifying bibliometric research trends and global focus in forecasting or prediction in the aviation industry.
- RQ2: Evolving research areas and trends in the field of airline forecasting.
- RQ3: Thematic evolution and keyword shifts using co-citation analysis.
- RQ4: Topic mapping and emerging directions in aviation forecasting

The following chapters explore the identified challenges and trends in aviation demand forecasting, offering detailed bibliometric analysis and highlighting emerging research themes for future exploration. Chapter 1 examines the topics within the context of airline forecasting. Chapter 2 explains the methodology used for bibliometric analysis. Chapter 3 presents the findings from the bibliometric analysis, and the final chapter discusses the results and provides the conclusions.

## 1. Literature review

Forecasting presents a major challenge in the aviation industry and relies on accurate demand predictions. Moreover, revenue management (RM) depends on classifying demand based on historical demand patterns and consumer behaviour (Selçuk & Avşar, 2019). An essential focus within this industry is modelling demand over the sales horizon, spanning studies of air transport demand at international, regional, national, intercity, and airport levels (Zachariah et al., 2023). Over recent years, the collaborative efforts of academic researchers and industry professionals have significantly propelled advancements in aviation forecasting. Many studies have been conducted employing linear and nonlinear models to predict passenger and cargo demand, with some even delving into the intriguing realm of demand prediction based on customer sentiment (Iddrisu et al., 2023).

Numerous studies in the aviation sector have focused on forecasting and predictive analysis, covering aspects such as passenger and cargo demand, flight delays, weather and climate change, route adjustments, maintenance, spare parts management, emissions, and fuel consumption. These critical areas of research can generally be categorised into several key domains.

The most prominent key domain is in forecasting demand for passengers. The demand for air travel continues to rise, accompanied by dynamic changes in passenger preferences, making forecasting an exceptional challenge for the aviation industry. Consequently, commercial aviation heavily relies on reliable travel demand predictions (Zachariah et al., 2023). A comprehensive assessment has been conducted concerning the prediction of passenger demand (Zachariah et al., 2023; Banerjee et al., 2020; Wang & Gao, 2021), including ticket pricing (Abdella et al., 2021). The use of demand forecasting is leveraged by airlines and airports for capacity planning, such as predicting passenger connectivity (Guimarães et al., 2022), determining airport infrastructure needs (Nieto & Carmona-Benítez, 2021), forecasting cargo space in combination carriers (Tseremoglou et al., 2022), and personalised customer experiences by considering customer behaviour (Sznajder et al., 2023). As Zachariah et al. (2023) summarised, the methods used in demand forecasting include econometric, statistical, machine learning, artificial intelligence, and hybrid models. Additionally, many studies have compared statistical methods with machine learning approaches, such as the research by FAN et al. (2023), which compared time series forecasting methods with machine learning, assessing forecast accuracy through RMSE error values.

Following demand forecasting, flight delay prediction has emerged as the second most extensively studied area. This prominence is due to the significant impact that flight delays have on both operational efficiency and passenger satisfaction. Various factors, including severe weather conditions, air traffic control issues, airline scheduling problems, military activities, and other reasons, typically cause flight delays (Wang et al., 2022). Besides, flight delays can affect many factors, such as operational efficiency and airline service quality (Wang & Pan, 2022). The complexity of these factors necessitates the use of advanced predictive models. Wang et al. (2022) categorised flight delay prediction methods into traditional statistical analysis, simulation modelling and queuing theory, and machine learning methods. Conventional statistical analysis methods are usually used to predict data trends and characteristics. The methods used for flight

delay predictions include time series, regression models, and correlation analysis. Several studies discussing time series were conducted by Wang et al. (2019), which addressed the flight delay situation by using time delay stability.

The flight delay prediction method based on simulation uses operations research to analyse, calculate the model, evaluate the results, and compare the appropriate case to achieve better predictions (Tascón & Díaz Olariaga, 2021; Lee et al., 2020). Flight delays have become increasingly prevalent, and forecasting related to climate change to meet the needs of airlines and issues associated with global warming and carbon emissions generated by aircraft (Oguntona, 2020). The machine learning method is based on artificial intelligence derived from a large amount of flight data. Several machine learning techniques include decision trees (Kang et al., 2021), Bayesian networks (Yang et al., 2023), random forests (Tang, 2021), k-nearest neighbours, support vectors, and deep learning.

Another airline forecasting topic is that contemporary climate science and weather forecasting studies increasingly integrate advanced computational techniques to address complex environmental challenges. Several studies focus on utilising machine learning and AI techniques (Choi et al., 2016), such as random forests, to improve the accuracy of weather predictions (Williams, 2014). This weather forecasting includes applications in wind forecasting and predicting exhaust gas temperature margins in aero-engines using transfer learning techniques (YAN et al., 2022). Additionally, significant attention is being given to developing data-driven models for wind forecasting and the prediction of thunderstorms, which are critical for improving the safety and efficiency of air travel (Andrés et al., 2021). The airline industry operates in a much more complex environment, which became especially true during the COVID-19 and post-COVID era, where passenger and cargo markets have gained significance. The situation resulted in numerous disruptions and uncertainties. COVID-19 has disrupted societies and economies worldwide, with the aviation sector experiencing one of the most catastrophic impacts (Suau-Sanchez et al., 2020). The COVID-19 pandemic caused a drastic drop in passenger air transport demand due to two forces: supply restriction and demand depression. Li et al. (2023) proposed a method separating the two COVID-19 forces and evaluating the respective impact on demand. It involves dividing passengers into segments based on passenger characteristics, simulating different scenarios, and predicting demand for each passenger segment in each scenario.

## 2. Research methods

This research uses a bibliometric analysis to assess the evolving interest in this topic, focusing on the number of publications during the specified period. The study was based on data retrieved from the Scopus database, which is more comprehensive, encompasses the most extensive peer-reviewed data, and is widely used in several bibliometric analyses (Asif et al., 2020). It identified the use of forecasting in the airline industry and the methods employed. Subsequently, Biblioshiny was used to create a classification of research subareas and developed to visualise the network analysis and the thematic evolution of keyword mapping. The Scopus database was used because it is recognised as the most extensive repository of abstracts and references from peer-reviewed literature (Pérez-Acebo et al., 2018). The ensuing section provides details regarding the specific database employed in this investigation.

- Database: Scopus
- Period of data: 1964 to 2023

This study involved a comprehensive literature assessment that adopts the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) methodology. This approach allows for identifying and evaluating the most relevant and appropriate research through several criteria (Rahman et al., 2023). Fig. 1 established a knowledge base on forecasting in the airline industry. The methodology comprises the following steps:

1. Defining the search criteria. This step involves selecting keywords. While investigating the Scopus database, specific keyword strings were utilised to ensure the inclusion of all relevant publications, including "forecast" OR "forecasting" in conjunction with "airline", OR "predict", "demand", AND "airline".

2. Search and screening. The second phase entailed the quest for documents to scrutinise. Initially, 1465 documents were found. Subsequently, a language restriction was applied to focus on English. Several criteria for exclusion and inclusion were used to refine the selection and achieve a final set of articles meeting the study's parameters. These articles were further assessed and chosen based on their relevance to the study's objectives, incorporating keywords, abstracts, titles, and source type considerations.

3. Publication information processing. This step identified articles related to the author's name, journal, and title to determine the presence of any duplications.

4. Document characteristics analysis. The study involved analysing information extracted from 820

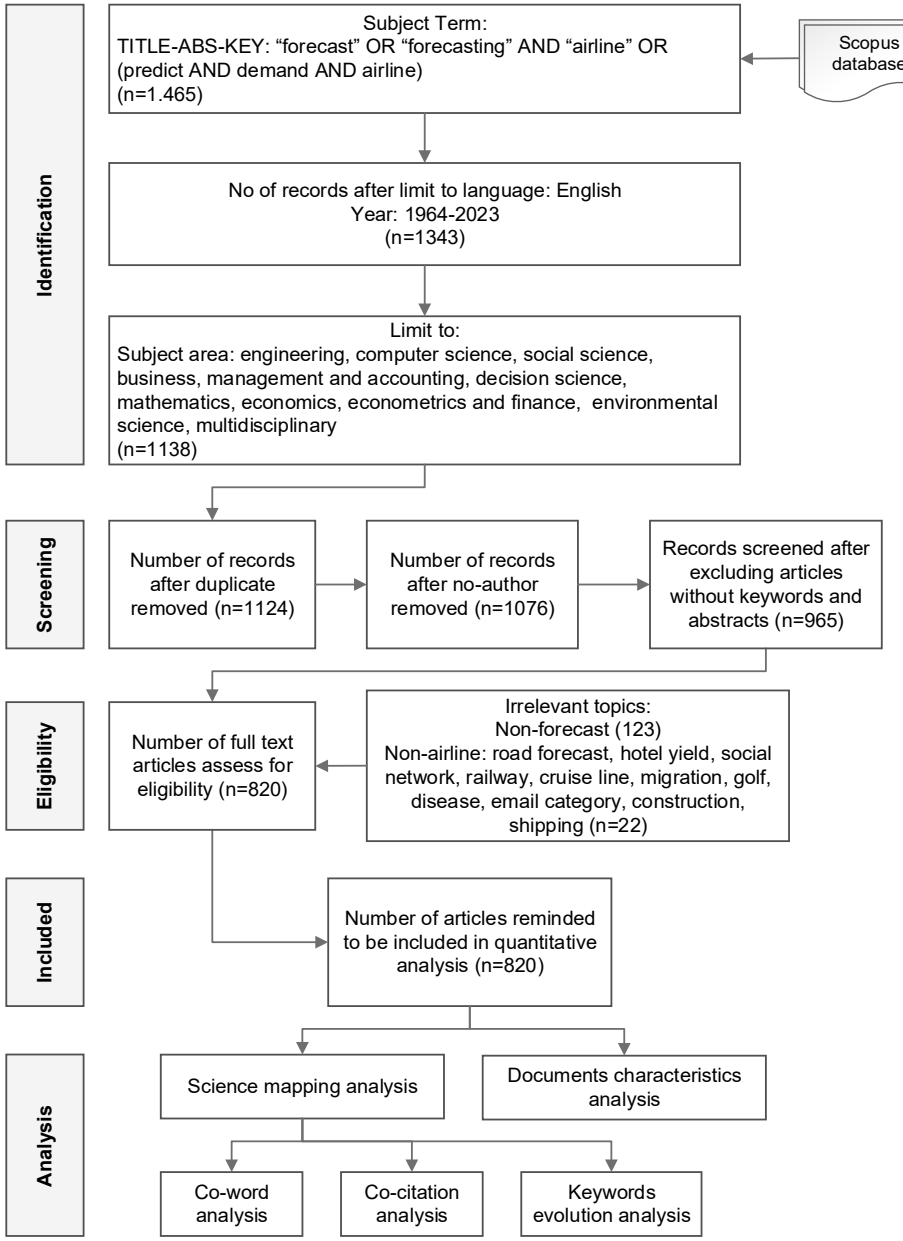



Fig. 1. Framework of the review paper

publications indexed in Scopus. Descriptive statistics were employed to conduct trend analyses of growth.

5. Scientific mapping analysis. In this step, the database was examined using various software tools. Biblioshiny was employed to analyse the network of keywords and authors and create thematic and keyword evolution maps.

6. A quantitative review of airline forecast and prediction. A quantitative analysis focused on numerous studies addressing various aspects of airline forecasting and prediction. Through this analysis, emerging subjects were identified, and the research requirements for the future of this field were charted.

# 3. Research results

## 3.1. RQ1: Bibliometric trends and global research focus

The bibliometric analysis encompasses 820 documents recorded from 1964 to 2023, providing an overview of the distribution of scientific publications identified in Scopus. The annual production of publications (Fig. 2) indicates three distinct phases. The first phase (1964–2000) had a consistent output of one to two publications annually. The second phase

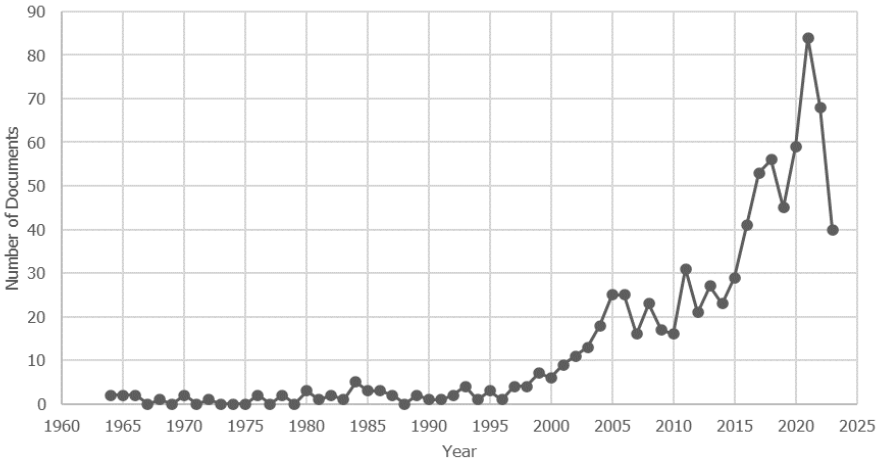



Fig. 2. Total publications per year

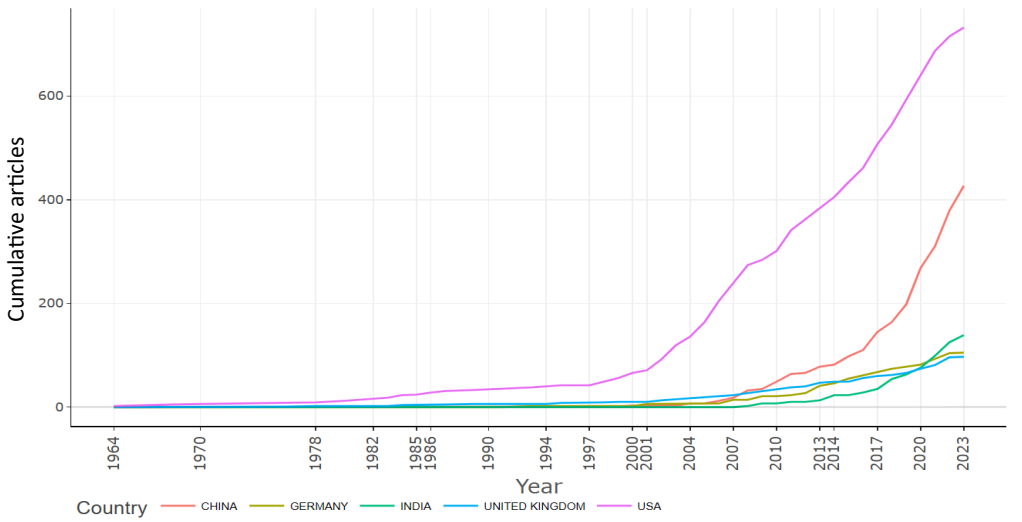



Fig. 3. Patterns of cumulative papers published in the six most productive countries

Tab. 1. Ten most globally cited documents

| MOST CITED DOCUMENT | PUBLICATION YEAR | AVERAGE CITATION PER YEAR | TOTAL CITATIONS |
|---|---|---|---|
| (Colizza et al., 2006) | 2006 | 45.78 | 824 |
| (Gui et al., 2020) | 2020 | 42.75 | 171 |
| (Gudmundsson Et Al., 2021) | 2021 | 34.67 | 104 |
| (McGill & Van Ryzin, 1999) | 1999 | 31.44 | 786 |
| (Baptista et al., 2018) | 2018 | 27.83 | 167 |
| (Lamb et al., 2020) | 2020 | 22.25 | 89 |
| (Hadavandi et al., 2010) | 2010 | 21.14 | 300 |
| (Liu et al., 2021) | 2021 | 21.33 | 64 |
| (Elsaid et al., 2018) | 2018 | 16.50 | 99 |
| (Zhang & Graham, 2020) | 2020 | 16.25 | 65 |

(2001–2015) saw a moderate increase, averaging 20 publications annually. The final phase (2016–2023) demonstrated significant growth, averaging 55 publications annually, reflecting an increased academic focus on aviation forecasting.

In the early 2000s, the USA led aviation research, primarily focusing on forecasting passenger demand (32% of studies). Flight delay prediction (15%) and weather (11%) were the second and third most prominent topics, respectively. In 2006, research on flight delays increased, while studies on economic issues declined. In China, aviation forecasting research experienced a steady upward trend starting in 2007. The topics most studied in China were flight delays (24%), passenger demand (19%), and spare parts (11%). Research on spare parts emerged in 2016, and studies on flight delays began in 2011. These trends are visually represented in Fig. 3, highlighting the topic distribution by country over time.

The top ten highly cited publications and authors on airline forecasting are listed in Table 1. The most frequently cited study is by Colizza et al. (2006), which models the global spread of diseases through large-scale airport networks using a stochastic computational framework. This framework incorporates the international air transportation network and population data to evaluate outbreak scenarios and containment policies. Another highly cited publication by Gui et al. (2020) uses big data and machine learning models to predict flight delays. Their findings indicate that while the Long Short-Term Memory (LSTM) model can handle aviation data, the random forest model provides higher accuracy and is more robust against overfitting. Gudmundsson et al. (2021) explored the relationship between economic strength and the airline industry's recovery following the COVID-19 pandemic, finding a direct correlation between financial shocks and passenger or cargo traffic.

### 3.2. RQ2: Evolving research areas and trends

This study identified and analysed 15 distinct research areas in aviation forecasting and prediction, as summarised in Table 2. The table illustrates the distribution and evolution of research focus across five periods, from 1964 to 2023. Passenger demand forecasting consistently emerged as the most researched topic, with 241 published papers reflecting its central role in airline operations and strategy. Flight delay prediction has seen significant publication growth since 2016, ranking second with 117 papers. This increase highlights the growing importance of minimising delays to enhance operational efficiency and customer satisfaction (Wang et al., 2022; Wang & Pan, 2022). Other areas, such as traffic demand, maintenance spare parts, risk analysis, and customer behaviour, show an upward trend, indicating increased interest in optimising operational aspects through personalised services and big data. In contrast, traditional topics like aircraft and engine research have

Tab. 2. Classification by field of airline forecasting

| Type of forecasting/prediction | 1964-2005 | 2006-2010 | 2011-2015 | 2016-2020 | 2021-2023 | Total papers |
|---|---|---|---|---|---|---|
| Demand passenger | 55 | 34 | 55 | 58 | 39 | 241 |
| Flight delay | 5 | 9 | 16 | 54 | 33 | 117 |
| Air traffic | 10 | 6 | 6 | 22 | 19 | 63 |
| Weather | 16 | 10 | 9 | 9 | 11 | 55 |
| Maintenance-spare part | 10 | 5 | 3 | 16 | 15 | 49 |
| Financial | 12 | 6 | 5 | 8 | 16 | 47 |
| Route | 6 | 5 | 7 | 19 | 9 | 46 |
| Customer behavior and competition | 3 | 1 | 2 | 18 | 16 | 40 |
| Price | 2 | 1 | 7 | 15 | 13 | 38 |
| Aircraft and engine | 15 | 3 | 4 | 9 | 1 | 32 |
| Fleet and scheduling | 5 | 8 | 5 | 3 | 4 | 25 |
| Airport capacity | 6 | 3 | 2 | 3 | 4 | 18 |
| Fuel and emission | 1 | 3 | 1 | 9 | 4 | 18 |
| Risk | 0 | 1 | 5 | 8 | 4 | 18 |
| Cargo | 2 | 2 | 4 | 3 | 2 | 13 |

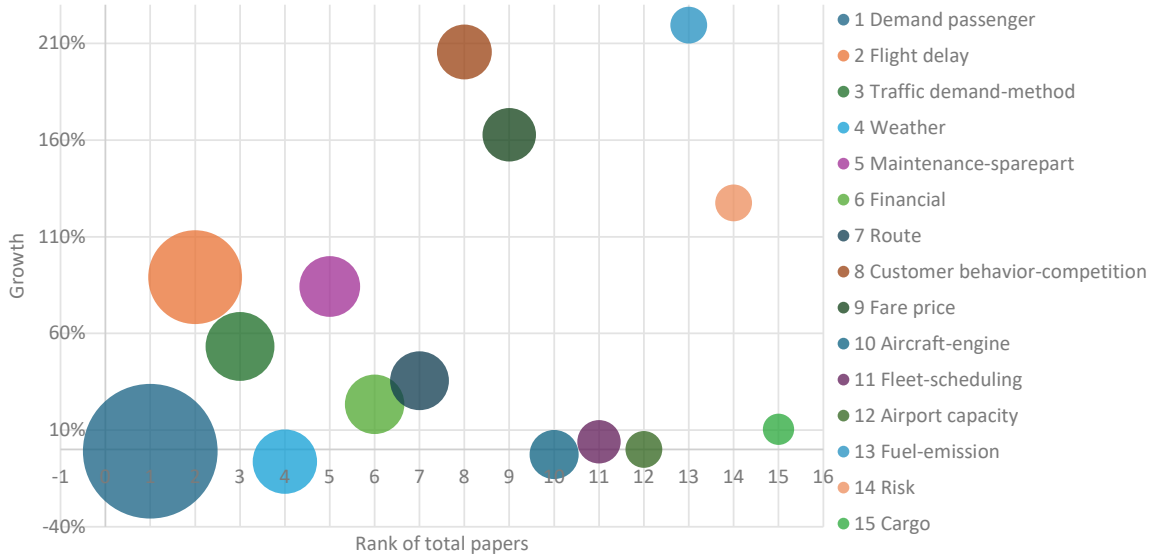



Fig. 4. Bubble chart analysis: growth vs rank of total papers

experienced a decline, suggesting advancements or shifting priorities. Fleet scheduling and airport capacity have also shown reduced publication volumes. Emerging areas, including customer behaviour, competition, and financial forecasting, have gained prominence, driven by the rise of big data and more nuanced market analysis.

Fig. 4 presents a bubble chart depicting the year-to-year growth of research publications across the 15 areas. Passenger demand forecasting ranks first in volume, as indicated by its large bubble size, but shows a low growth rate, indicating slower expansion despite its significance. Flight delay prediction ranks second and exhibits a high growth rate, signifying its rising importance. Mid-ranking topics, such as traffic demand and spare parts forecasting, display notable growth, likely due to advancements in data analytics and the need for precision in forecasting. Emerging areas, like customer behaviour, competition, and fare pricing show rapid growth, reflecting shifting academic interest. Under-researched topics, such as fuel and emissions and risk forecasting, are indicated by small bubble sizes, highlighting potential research opportunities. Traditional areas, like aircraft and engine forecasting, show minimal growth and low volume, suggesting a shift in research priorities towards emerging challenges and operational complexities.

### 3.3. RQ3: Thematic evolution and keyword shifts

The study collected and analysed keywords from research publications from 1964 to 2023. Keyword occurrence frequency was employed to identify the most used terms and to determine potential future research directions. The evolution of research themes over time is illustrated in Fig. 5, highlighting different thematic transformations from 1964 to 2023. The keyword "revenue management" demonstrated consistent use throughout this period. Keywords such as

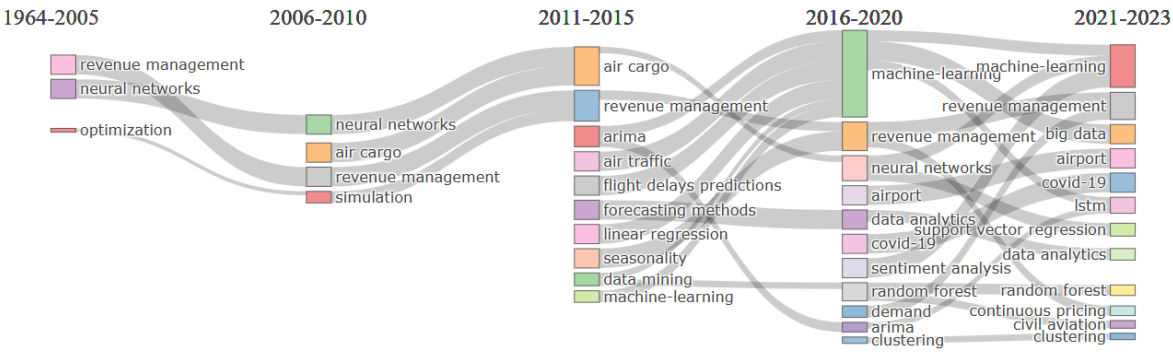



Fig. 5. Thematic evolution of author's keywords

"data mining", "machine learning", and related forecasting methodologies began to emerge around 2011 and have since evolved to include terms like "random forest", "data analytics", "sentiment analysis", and "LSTM". Air cargo research was particularly prominent between 2006 and 2015 but has declined. "Flight delay" research, prevalent from 2011 to 2015, became closely associated with machine learning studies. Between 2021 and 2023, research themes shifted towards advanced technologies and the impact of COVID-19, with machine learning remaining a central focus alongside innovations in big data, continuous pricing, and random forest methodologies.

### 3.4. RQ4: Topic mapping and emerging directions in aviation forecasting

Fig. 6 visually maps the evolution of trending research topics over time, illustrating the frequency and emergence of key terms in aviation forecasting. In the early years, foundational topics, such as technological forecasting, flight dynamics, and industrial economics, were prevalent, reflecting the industry's initial focus on operational fundamentals. Over time, there was a marked shift towards more advanced computational methods and data-driven research. By the mid-2000s, terms like machine learning, deep learning, and prediction modelling began to dominate, indicating the growing integration of emerging technologies into aviation research. The COVID-19 pandemic, which emerged as a dominant research topic by 2020, further reshaped aviation research priorities. Studies on the industry's recovery prospects emphasised several core strategies, including operational resilience (Linden, 2021), lessons learned from successful recovery stories (Czerny et al., 2021), and the critical role of digital transformation in enhancing industry adaptability (Halpern et al., 2021). Furthermore, the research explored defining attributes for resilience and sustainability in aviation, underlining the industry's need to withstand future disruptions. Colizza et al. (2006) demonstrated the applicability of stochastic models in predicting epidemic transmission through airports. Li et al. (2023) expanded on this by leveraging forecasting models to evaluate the impact of the COVID-19 pandemic on passenger demand and operational strategies. Simultaneously, research increasingly focused on enhancing prediction accuracy and improving operational efficiency. Terms such as learning algorithms, flight delay prediction, and predictive maintenance became more frequent, reflecting the industry's shift towards data-driven decision-making to optimise real-time operations. Collectively, these emerging trends illustrate how the aviation sector is evolving to integrate advanced forecasting methods and technological innovations, particularly in response to global disruptions.

Fig. 7 further illustrates the evolution of analytical methods over time. Traditional statistical methods, such as regression analysis and time series modelling, were dominant from 1964 to 2005, with a 100% usage rate. However, their prevalence steadily declined, reaching just 11% by 2021–2023. Simulation and queuing theory methods, which were first employed between 2006 and 2010, peaked at 33% during that period but subsequently declined to 6% in the latest period. In contrast, machine learning methods — including decision trees, random forests, and deep neural networks — rose rapidly, becoming the dominant analytical approach by 2021–2023 with an 83% usage rate. This trend underscores the field's increasing reliance on machine learning to manage complex aviation challenges and improve prediction accuracy.

Emerging topics and strategic opportunities. Topic mapping reveals several emerging themes that airline managers can leverage to gain a competitive edge. Machine learning, deep learning, and customer behaviour analysis are critical for developing accurate demand forecasts, predictive maintenance, and customer-centric strategies. Sentiment analysis highlights the growing emphasis on passenger feedback and experience in operational decisions, while continuous pricing reflects the trend toward dynamic and real-time pricing strategies to maximise revenue. Practitioners can enhance operational decision-making and develop adaptive, real-time responses to evolving market conditions by focusing on these emerging topics.

Sinking topics and standardisation. Conversely, topic mapping highlights areas where research interest has stabilised or declined. Air cargo forecasting, fleet scheduling, and traditional statistical methods have become more standardised, reflecting their diminished relevance in addressing new challenges. The industry has primarily developed best practices for these areas, reducing the need for further research. Practitioners in these domains should focus on applying proven solutions and optimising existing processes to maintain operational efficiency without significant innovation.

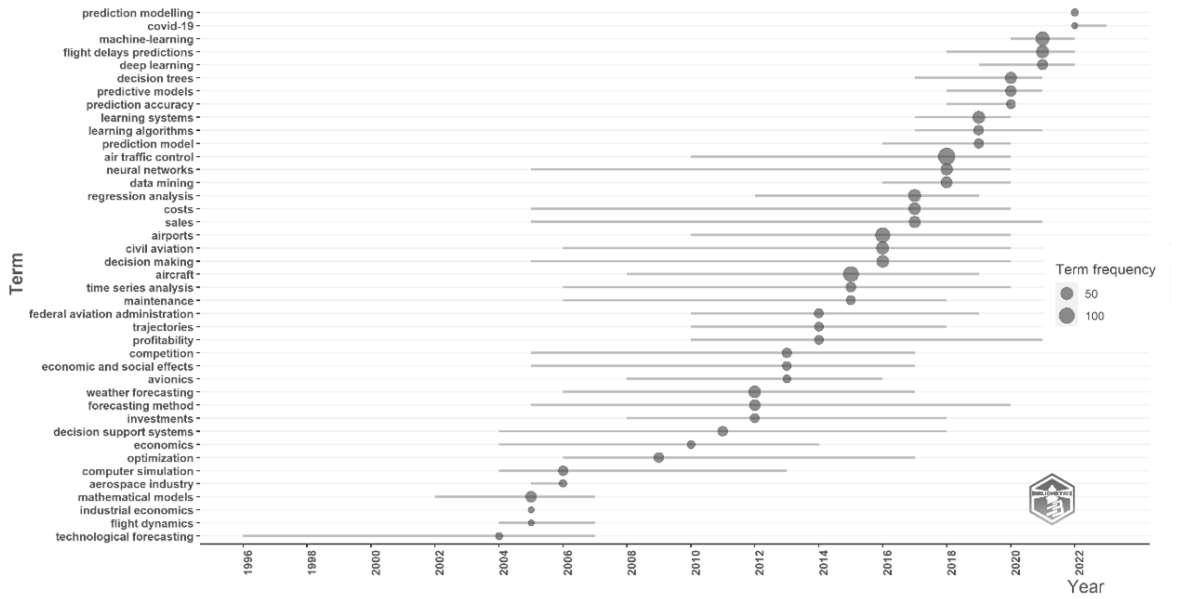



Fig. 6. Mapping of trend topics

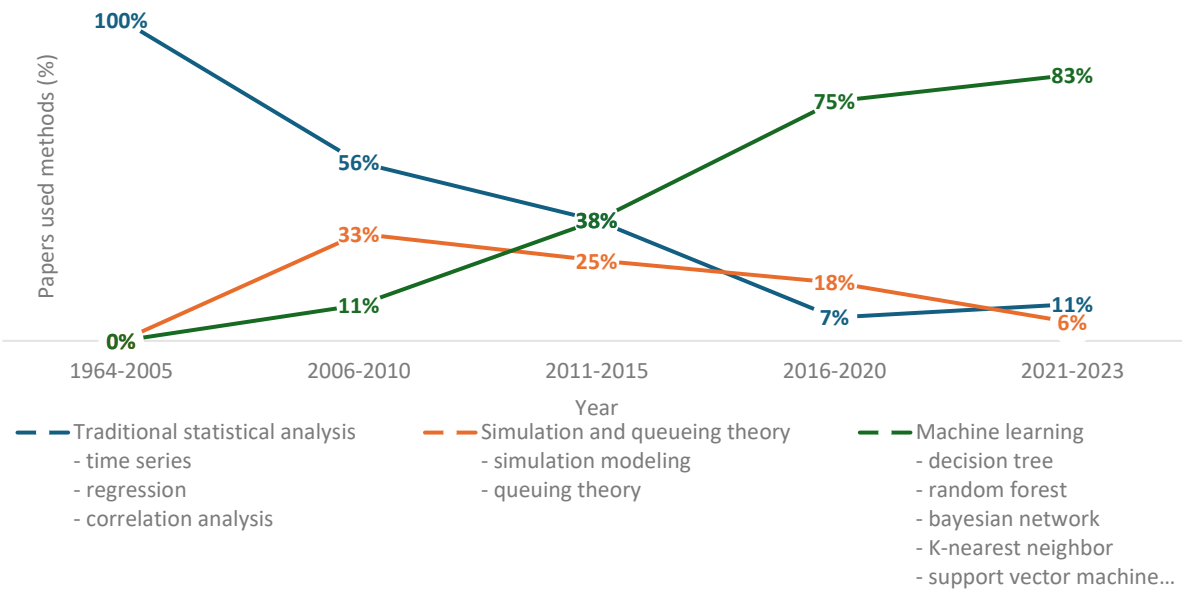



Fig. 7. Number of articles using machine-learning methods

## 4. Discussion of the results

This section synthesises key findings across bibliometric trends, thematic evolution, and technological advancements to address the four research questions (RQ1–RQ4).

RQ1. The bibliometric analysis of aviation forecasting research revealed significant shifts in publication trends and geographical research focus from 1964 to 2023. Three distinct phases of publication growth were identified, with a notable increase in research output from 2016 onwards, signalling heightened academic interest. This growth aligns with global advancements in data analytics and the increasing complexity of aviation operations. Geographically, the USA initially led aviation forecasting research, with passenger demand forecasting being

the most studied topic (32%). The focus later expanded to flight delay prediction and weather-related studies. China followed with a steady upward trend in aviation research, particularly on flight delays and spare parts forecasting, reflecting the country's expanding aviation industry. These findings indicate that regional aviation market needs and economic priorities significantly influence research directions. Highly cited publications reflect the integration of advanced methodologies such as machine learning and stochastic modelling. For instance, Colizza et al. (2006) applied network-based stochastic models to disease transmission through airports, demonstrating the relevance of aviation research in broader public health contexts. The prominence of machine learning models like random forest and LSTM in predicting flight delays underscores the shift towards more data-driven, predictive analytics approaches in aviation forecasting.

RQ2. The study identified 15 distinct research areas, with passenger demand forecasting consistently dominating the field due to its strategic importance for airline operations. Flight delay prediction has grown significantly since 2016, driven by operational efficiency improvements and rising customer expectations. Emerging topics, such as customer behaviour, financial forecasting, and competition analysis, have gained traction, influenced by the proliferation of big data and evolving market dynamics. Conversely, traditional research areas such as aircraft and engine performance have declined, likely due to advancements in these fields or shifting research priorities. The bubble chart analysis illustrates varying growth rates across research areas. Passenger demand forecasting remains the most studied topic but shows slower growth, reflecting its established status. In contrast, flight delay prediction and emerging areas like customer behaviour exhibit rapid growth, indicating increasing interest and potential for future research. Under-researched areas such as fuel and emissions forecasting present opportunities for innovation, particularly in light of sustainability and regulatory challenges.

RQ3. The evolution of research themes highlights a transition from traditional statistical methods to machine learning and big data-driven approaches. Early studies relied heavily on time series and regression analyses, but more advanced techniques have largely supplanted these methods. By 2021–2023, machine learning methods, including decision trees, random forests, and deep neural networks, became dominant, reflecting their superior accuracy and adaptability in managing aviation complexities. Key terms like "machine learning", "data mining", and "sentiment analysis" emerged post-2011 and have evolved to encompass more specialised methodologies. COVID-19 research themes emerged prominently after 2020, further underscoring the aviation industry's adaptive response to global crises. Continuous pricing and predictive maintenance also surfaced as critical areas, indicating a strategic shift towards real-time decision-making and dynamic operational adjustments.

RQ4. Topic mapping revealed distinct patterns in emerging and declining research areas. Emerging areas, like machine learning, customer behaviour analysis, and sentiment analysis, are poised to transform aviation forecasting by enabling more accurate demand forecasts and customer-centric strategies. Continuous pricing further highlights the shift towards dynamic pricing models for revenue optimisation. On the other hand, areas like air cargo forecasting and fleet scheduling have become more standardised, reducing the need for further research. These areas have established best practices, indicating industry maturity. However, underexplored topics such as fuel and emissions forecasting remain critical for addressing sustainability and regulatory compliance, presenting opportunities for future research and innovation. The COVID-19 pandemic catalysed a shift in research priorities, emphasising the importance of resilience, adaptability, and digital transformation in aviation recovery. Applying stochastic models and forecasting tools further underscores the industry's proactive approach to mitigating risks and enhancing operational efficiency. Future research should focus on deepening the understanding of predictive technologies and expanding studies on resilience and sustainability to navigate future disruptions effectively. Expanding this area could help the industry develop robust long-term strategies to withstand financial shocks and market disruptions, ensuring sustainable recovery and growth. Conversely, topic mapping highlights areas where research interest has stabilised or declined. Air cargo forecasting, fleet scheduling, and traditional statistical methods have become more standardised, reflecting their diminished relevance in addressing new challenges. The industry has largely developed best practices for these areas, reducing the need for further research. Practitioners in these domains should focus on applying proven solutions and optimising existing processes to maintain operational efficiency without significant innovation.

Despite topic diversification, certain themes remain under-researched, presenting key challenges in airline forecasting. However, fuel and emissions forecasting remains a niche research area critical for addressing sustainability and regulatory concerns. Expanding research in this direction presents an opportunity for innovation in sustainability initiatives and compliance strategies. Economic resilience in the face of financial shocks, particularly post-pandemics, is crucial for developing robust long-term operational strategies. The pandemic's impact has also highlighted the importance of economic resilience research, which remains under-explored. Expanding this area could help the industry develop robust long-term strategies to withstand financial shocks and market disruptions, ensuring sustainable recovery and growth. Addressing these challenges will be essential for the aviation sector to adapt to evolving complexities and achieve sustainable industry resilience.

## Conclusions

This study has successfully addressed its objectives and research questions (RQ1–RQ4) by comprehensively analysing aviation forecasting trends, thematic evolution, and emerging opportunities. Bibliometric analysis revealed significant shifts in publication trends from 1964 to 2023, driven by advancements in data analytics and aviation complexity. Passenger demand forecasting remains the most researched area due to its strategic importance for airline operations, while emerging topics in customer behaviour, financial forecasting, and competition analysis have gained prominence. Thematic evolution highlighted the transition from traditional statistical methods to machine learning and big data-driven approaches, underscoring aviation's increasing reliance on predictive analytics. Topic mapping emphasised emerging areas in sentiment analysis and dynamic pricing, reflecting evolving customer expectations and the need for real-time decision-making. Established areas like air cargo forecasting and fleet scheduling have become more standardised, indicating industry maturity and reduced research demand.

This work highlights key focus areas for airline managers to gain a competitive advantage by concentrating on emerging research trends and data-driven approaches. For stabilised research areas like fleet scheduling, practitioners can focus on applying established best practices to maintain operational efficiency and optimise decision-making processes. Future research should prioritise underexplored areas in fuel and emissions forecasting, economic resilience, and sustainability to address evolving regulatory and environmental challenges. Additionally, advancing research on real-time pricing, customer behaviour analysis, and predictive maintenance will ensure sustainable growth and resilience in the aviation industry as it adapts to new complexities and global disruptions.

The author gratefully acknowledges that the Final Project Recognition Grant, Universitas Gadjah Mada Number 5075/UN1.P.II/DitLit/PT.01.01/2023, provides financial support for this research. The author would also like to express gratitude to the Indonesian Education Scholarship (BPI), Center for Higher Education Funding and Assessment (BPPAPT) and the Indonesia Endowment Funds for Education (LPDP) from the Ministry of Finance of the Republic of Indonesia for granting the scholarship.

## Acknowledgments

The author gratefully acknowledges that the Final Project Recognition Grant, Universitas Gadjah Mada Number 5075/UN1.P.II/DitLit/PT.01.01/2023, provides financial support for this research. The author would also like to express gratitude to the Indonesian Education Scholarship (BPI), Center for Higher Education Funding and Assessment (PPAPT) and the Indonesia Endowment Funds for Education (LPDP) from the Ministry of Finance of the Republic of Indonesia for granting the scholarship.

## Literature

Abdella, J. A., Zaki, N. M., Shuaib, K., & Khan, F. (2021). Airline ticket price and demand prediction: A survey. *Journal of King Saud University - Computer and Information Sciences*, *33*(4), 375-391. doi: 10.1016/j.jksuci.2019.02.001

Andrés, E., González-Arribas, D., Soler, M., Kamgarpour, M., & Sanjurjo-Rivo, M. (2021). Informed scenario-based RRT∗ for aircraft trajectory planning under ensemble forecasting of thunderstorms. *Transportation Research Part C: Emerging Technologies*, *129*. doi: 10.1016/j.trc.2021.103232

Asif, S., Rafi, R., & Ali, A. (2020). A bibliometric analysis of revenue management in airline industry. *Journal of Revenue and Pricing Management*, *2002*. doi: 10.1057/s41272-020-00247-1

Banerjee, N., Morton, A., & Akartunalı, K. (2020). Passenger demand forecasting in scheduled transportation.

*European Journal of Operational Research*, *286*(3), 797-810. doi: 10.1016/j.ejor.2019.10.032

Baptista, M., Sankararaman, S., de Medeiros, I. P., Nascimento, C., Prendinger, H., & Henriques, E. M. P. (2018). Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Computers and Industrial Engineering*, *115*, 41-53. doi: 10.1016/j.cie.2017.10.033

Caiado, J., & Lúcio, F. (2023). Stock market forecasting accuracy of asymmetric GARCH models during the COVID-19 pandemic. *North American Journal of Economics and Finance*, *68*, 101971. doi: 10.1016/j.najef.2023.101971

Choi, S., Kim, Y. J., Briceno, S., & Mavris, D. (2016). Prediction of weather-induced airline delays based on machine learning algorithms. *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 1-6. doi: 10.1109/DASC.2016.7777956.

Colizza, V., Barrat, A., Barthélemy, M., & Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2015-2020. doi: 10.1073/pnas.0510525103

Czerny, A. I., Fu, X., Lei, Z., & Oum, T. H. (2021). Post pandemic aviation market recovery: Experience and lessons from China. *Journal of Air Transport Management*, *90*, 101971. doi: 10.1016/j.jairtraman.2020.101971

ElSaid, A. E. R., El Jamiy, F., Higgins, J., Wild, B., & Desell, T. (2018). Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration. *Applied Soft Computing Journal*, *73*, 969-991. doi: 10.1016/j.asoc.2018.09.013

Fan, W., Wu, X., Shi, X. Y., Zhang, C., Wai Hung, I., Kai Leung, Y., & Zneg, L. S. (2023). Support vector regression model for flight demand forecasting. *International Journal of Engineering Business Management*, *15*(2898). doi: 10.1177/18479790231174318

Gudmundsson, S. V., Cattaneo, M., & Redondi, R. (2021). Forecasting temporal world recovery in air transport markets in the presence of large economic shocks: The case of COVID-19. *Journal of Air Transport Management*, *91*, 102007. doi: 10.1016/j.jairtraman.2020.102007

Gui, G., Liu, F., Sun, J., Yang, J., Zhou, Z., & Zhao, D. (2020). Flight delay prediction based on aviation big data and machine learning. *IEEE Transactions on Vehicular Technology*, *69*(1), 140-150. doi: 10.1109/TVT.2019.2954094

Guimarães, M., Soares, C., & Ventura, R. (2022). Decision Support Models for Predicting and Explaining Airport Passenger Connectivity From Data. *IEEE Transactions on Intelligent Transportation Systems*, *23*(9), 16005-16015. doi: 10.1109/TITS.2022.3147155

Hadavandi, E., Shavandi, H., & Ghanbari, A. (2010). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems*, *23*(8), 800-808. doi: 10.1016/j.knosys.2010.05.004

Halpern, N., Mwesiumo, D., Suau-Sanchez, P., Budd, T., & Bråthen, S. (2021). Ready for digital transformation? The effect of organisational readiness, innovation, airport size and ownership on digital change at airports. *Journal of Air Transport Management*, *90*, 101949. doi: 10.1016/j.jairtraman.2020.101949

Huang, F., Zhang, T., Wang, Q., & Zhou, D. (2023). CO2 emission change in China's aviation industry: A fleet-wide index decomposition and scenario analysis. *Transportation Research Part D: Transport and Environment*, *119*, 1-16. doi: 10.1016/j.trd.2023.103743

IATA. (2016). *Forecasts passenger demand to double over 20 years*. Retrieved from https://airlines.iata.org/2016/11/23/passenger-numbers-double-2035

Iddrisu, A. M., Mensah, S., Boafo, F., Yeluripati, G. R., & Kudjo, P. (2023). A sentiment analysis framework to classify instances of sarcastic sentiments within the aviation sector. *International Journal of Information Management Data Insights*, *3*(2), 100180. doi: 10.1016/j.jjimei.2023.100180

İnan, T. T. (2022). Forecasting Recovery Period of the Airfreight Transportation from Covid-19 Pandemic by using Time Series Modelling. *Logistics Research*, *15*(1), 1-18. doi: 10.23773/2022_03

Kang, Z., Shang, J., Feng, Y., Zheng, L., Wang, Q., Sun, H., Qiang, B., & Liu, Z. (2021). A deep sequence-to-sequence method for accurate long landing prediction based on flight data. *IET Intelligent Transport Systems*, *15*(8), 1028-1042. doi: 10.1049/itr2.12078

Lamb, T. L., Winter, S. R., Rice, S., Ruskin, K. J., & Vaughn, A. (2020). Factors that predict passengers willingness to fly during and after the COVID-19 pandemic. *Journal of Air Transport Management*, *89*, 101897. doi: 10.1016/j.jairtraman.2020.101897

Lee, J., Marla, L., & Jacquillat, A. (2020). Dynamic Disruption Management in Airline Networks Under Airport Operating Uncertainty. *Transportation Science*, *54*(4), 973-997. doi: 10.1287/trsc.2020.0983

Li, Q., Guan, X., & Liu, J. (2023). A CNN-LSTM framework for flight delay prediction. *Expert Systems with Applications*, *227*, 120287. doi: 10.1016/j.eswa.2023.120287

Li, X., de Groot, M., & Bäck, T. (2023). Using forecasting to evaluate the impact of COVID-19 on passenger air transport demand. *Decision Sciences*, *54*(4), 394-409. doi: 10.1111/deci.12549

Linden, E. (2021). Pandemics and environmental shocks: What aviation managers should learn from COVID-19 for long-term planning. *Journal of Air Transport Management*, *90*, 101944. doi: 10.1016/j.jairtraman.2020.101944

Liu, J., Lei, F., Pan, C., Hu, D., & Zuo, H. (2021). Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion. *Reliability Engineering and System Safety*, *214*, 107807. doi: 10.1016/j.ress.2021.107807

McGill, J. I., & Van Ryzin, G. J. (1999). Revenue management: research overview and prospects. *Transportation Science*, *33*(2), 233–256. doi: 10.1287/trsc.33.2.233

Nieto, M. R., & Carmona-Benítez, R. B. (2021). An approach to measure the performance and the efficiency of future airport infrastructure. *Mathematics*, *9*(16). doi: 10.3390/math9161873

Oguntona, O. (2020). Longer - term aircraft fleet modelling: narrative review of tools and measures for mitigating carbon emissions from aircraft fleet. *CEAS Aeronautical Journal*, *11*(1), 13-31. doi: 10.1007/s13272-019-00424-y

Pérez-Acebo, H., Linares-Unamunzaga, A., Abejón, R., & Rojí, E. (2018). Research trends in pavement

management during the first years of the 21st century: A bibliometric analysis during the 2000-2013 Period. *Applied Sciences (Switzerland)*, *8*(7). doi: 10.3390/app8071041

Rahman, T., Zudhy Irawan, M., Noor Tajudin, A., Rizka Fahmi Amrozi, M., & Widyatmoko, I. (2023). Knowledge mapping of cool pavement technologies for urban heat island Mitigation: A Systematic bibliometric analysis. *Energy and Buildings*, *291*. doi: 10.1016/j.enbuild.2023.113133

Samli, R., Firat, M., & Yiltas-Kaplan, D. (2021). Forecasting air travel demand for selected destinations using machine learning methods. *Journal of Universal Computer Science*, *27*(6), 564-581. doi: 10.3897/JUCS.68185

Selçuk, A. M., & Avşar, Z. M. (2019). Dynamic pricing in airline revenue management. *Journal of Mathematical Analysis and Applications*, *478*(2), 1191-1217. doi: 10.1016/j.jmaa.2019.06.012

Suau-Sanchez, P., Voltes-Dorta, A., & Cugueró-Escofet, N. (2020). An early assessment of the impact of COVID-19 on air transport: Just another crisis or the end of aviation as we know it? *Journal of Transport Geography*, *86*, 102749. doi: 10.1016/j.jtrangeo.2020.102749

Sulistyowati, R., Kuswanto, H., Series, A. T., & Tsr, R. (2018). Hybrid Forecasting Model To Predict Air Passenger and Cargo in Indonesia. *International Conference on Information and Communications Technology*, 442-447.

Sznajder, M., Ratliff, R., & Kaya, C. (2023). A heuristic for incorporating ancillaries into air choice models with personalization (part 2: integrated multinomial logit and hedonic regression models). *Journal of Revenue and Pricing Management*, *22*(2), 140-151. doi: 10.1057/s41272-022-00400-y

Tang, Y. (2021). Airline Flight Delay Prediction Using Machine Learning Models. *ACM International Conference Proceeding Series*, 151-154. doi: 10.1145/3497701.3497725

Tascón, D. C., & Díaz Olariaga, O. (2021). Air traffic forecast and its impact on runway capacity. A System Dynamics approach. *Journal of Air Transport Management*, *90*. doi: 10.1016/j.jairtraman.2020.101946

Tseremoglou, I., Bombelli, A., & Santos, B. F. (2022). A combined forecasting and packing model for air cargo loading: A risk-averse framework. *Transportation Research Part E: Logistics and Transportation Review*, *158*, 102579. doi: 10.1016/j.tre.2021.102579

van Ryzin, G., & McGill, J. (2000). Revenue management without forecasting or optimization: an adaptive algorithm for determining airline seat protection levels. *Management Science*, *46*, 760-775.

Vinod, B. (2022). Airline revenue planning and the COVID-19 pandemic. *Journal of Tourism Futures*, *8*(2), 245-253. doi: 10.1108/JTF-02-2021-0055

Wang, J., Lu, X., He, F., & Ma, F. (2020). Which popular predictor is more useful to forecast international stock markets during the coronavirus pandemic: VIX vs EPU? *International Review of Financial Analysis*, *72*, 101596. doi: 10.1016/j.irfa.2020.101596

Wang, J., & Pan, W. (2022). Flight delay prediction based on ARIMA. *Proceedings - 2022 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI)*, 186-190. doi: 10.1109/ICCEAI55464.2022.00047

Wang, S., & Gao, Y. (2021). A literature review and citation analyses of air travel demand studies published between 2010 and 2020. *Journal of Air Transport Management*, *97*, 102135. doi: 10.1016/j.jairtraman.2021.102135

Wang, T., Zheng, Y., & Xu, H. (2022). A Review of Flight Delay Prediction Methods. *International Conference on Big Data Engineering and Education (BDEE)*, 135-141. doi: 10.1109/BDEE55929.2022.00029

Wang, Y. J., Zhu, Y. F., Zhu, C. P., Wu, F., Yang, H. J., Yan, Y. J., & Hu, C. K. (2019). Indicator of serious flight delays with the approach of time-delay stability. *Physica A: Statistical Mechanics and Its Applications*, *518*, 363-373. doi: 10.1016/j.physa.2018.11.038

Williams, J. K. (2014). Using random forests to diagnose aviation turbulence. *Machine Learning*, *95*(1), 51-70. doi: 10.1007/s10994-013-5346-7

Yan, Z., Zhongs, S., Lin, L., Cui, Z., & Zhao, M. (2022). A step parameters prediction model based on transfer process neural network for exhaust gas temperature estimation after washing aero-engines. *Chinese Journal of Aeronautics*, *35*(3), 98-111. doi: 10.1016/j.cja.2021.07.035

Yang, Z., Chen, Y., Hu, J., Song, Y., & Mao, Y. (2023). Departure delay prediction and analysis based on node sequence data of ground support services for transit flights. *Transportation Research Part C: Emerging Technologies*, *153*, 104217. doi: 10.1016/j.trc.2023.104217

Ye, Q., Zhou, R., & Asmi, F. (2023). Evaluating the Impact of the Pandemic Crisis on the Aviation Industry. *Transportation Research Record*, *2677*(3), 1551-1566. doi: 10.1177/03611981221125741

Zachariah, R. A., Sharma, S., & Kumar, V. (2023). Systematic review of passenger demand forecasting in aviation industry. *Multimedia Tools and Applications*, *82*(30), 46483-46519. doi: 10.1007/s11042-023-15552-1

Zhang, F., & Graham, D. J. (2020). Air transport and economic growth : a review of the impact mechanism and causal relationships. *Transport Reviews*, *40*(4), 506-528. doi: 10.1080/01441647.2020.1738587

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 9 December 2024
accepted: 10 May 2025

pages: 37-62



# Enhancing airport services: data-driven analysis of passenger satisfaction and service quality in Southeast Asia

Thitinan Pholsook
Sarawut Ramjan
Warit Wipulanusat

ABSTRACT

Airports encompass a range of service touchpoints that directly impact passenger satisfaction and, consequently, the likelihood of service recommendation. This study investigates the service quality of Southeast Asian airports by applying five supervised machine learning classification models — decision trees, random forests, support vector machines, neural networks, and gradient boosting machines — on passenger satisfaction data extracted from the Skytrax website. The dataset includes evaluations of various service dimensions, such as staff behaviour, queuing time, and overall experience. This study incorporates cross-validation and hyperparameter tuning to identify the most suitable model for classifying passenger satisfaction. Among the models tested, the random forest classifier achieved the highest accuracy (0.91), demonstrating strong robustness and interpretability. Model performance was assessed using confusion matrices, balanced accuracy, the Matthews correlation coefficient (MCC), and ROC curves. Furthermore, SHAP values were used to identify the most influential service touchpoints, highlighting airport staff performance and queue management as key factors. These findings align with existing literature emphasising the pivotal role of well-trained airport employees and efficient queuing systems in shaping positive passenger experiences. Studies have shown that courteous staff interactions, efficient conflict resolution, and reduced waiting times significantly contribute to customer satisfaction and loyalty. Additionally, the integration of smart technologies such as self-service kiosks, automated security systems, and touchless check-in and baggage solutions enhances operational efficiency and aligns with sustainability initiatives. This study offers a data-driven approach for airport managers to optimise service delivery, increase passenger experiences, and tailor improvements to specific airport environments.

KEY WORDS

**airport quality service, machine learning, data classification, hyperparameter tuning**

**10.2478/emj-2025-0011**

**Sarawut Ramjan**
Thammasat University,
Bangkok, 10200, Thailand
ORCID 0000-0002-6531-0419

Corresponding author:
e-mail: sarawut@citu.tu.ac.th

**Thitinan Pholsook**
Suranaree University of Technology,
Nakhon Ratchasima, 30000, Thailand
ORCID 0009-0005-8484-3156
e-mail: thitinan.p@g.sut.ac.th

**Warit Wipulanusat**
Thammasat University,
Pathumthani, 12120, Thailand
ORCID 0000-0003-1006-6540
e-mail: wwarit@engr.tu.ac.th

## Introduction

Air transportation continues to be a vital mode of transport, reaching a peak of 37.52 billion passengers in 2019. However, the outbreak of the COVID-19 pandemic led to a dramatic reduction in passenger numbers, plummeting to 15.57 billion in 2020 (World Bank Group, 2024), as illustrated in Fig. 1. The slow recovery of air travel in the Asia–Pacific region can be attributed to strict travel restrictions, slower vaccination rollouts, and prolonged lockdowns, collectively hindering its aviation sector. In contrast, North America and Europe exhibited recovery signs as early

Pholsook, T., Ramjan, S., & Wipulanusat, W. (2025). Enhancing airport services: data-driven analysis of passenger satisfaction and service quality in Southeast Asia. *Engineering Management in Production and Services, 17*(2), 37-62. doi: 10.2478/emj-2025-0011

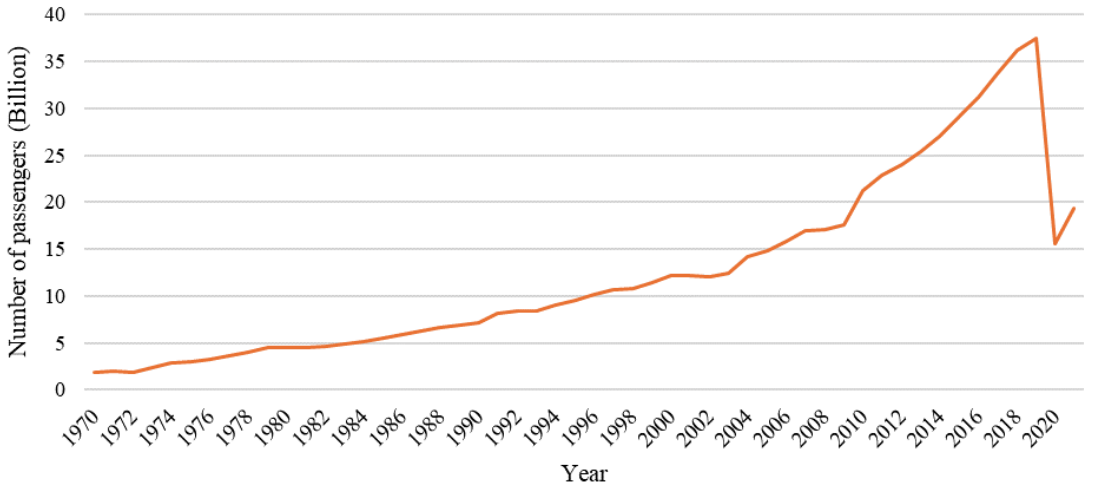



Fig. 1. Number of air passengers from 1970 to 2021

Source: World Bank Group, 2024. *Air transport, passengers carried*. Retrieved from https://data.worldbank.org/indicator/IS.AIR.PSGR

as the first half of 2021 due to robust immunisation campaigns, relaxed travel restrictions, and a swift resurgence in domestic travel demand. Despite these improvements, a full recovery to prepandemic levels is not anticipated until the latter half of 2024. This prolonged recovery period highlights the complexities associated with international travel, necessitating harmonised travel regulations, elevated global vaccination rates, and measures to rebuild passenger confidence. The aviation industry is projected to recover fully by 2025 (ACI, 2023).

According to the ACI World, enhancing the customer experience is the most effective strategy for increasing airport nonaviation revenue (NAR). An analysis of airport service quality (ASQ) data reveals that a 1% increase in the number of passengers is associated with a 0.7% to 1% increase in the NAR. Additionally, expanding the commercial area by 1% correlates with a 0.2% increase in the NAR, whereas a 1% improvement in overall passenger satisfaction results in an average 1.5% increase in the NAR (ACI, 2021b). As airports increasingly adopt a commercial management approach, the value of customer feedback has become substantially more prominent.

The primary aim of this study is to enhance the overall service quality of airports by establishing a secure, efficient, and passenger-centric environment that exceeds user expectations and contributes to an improved travel experience. (Elshafey et al., 2007). In 2006, airport service quality (ASQ) stood as the foremost global programme for benchmarking airport passenger services that evaluated customer satisfaction with airport facilities. ASQ provides objective metrics that assist airports in enhancing their performance and delivering competitive services to passengers. To help airports assess their performance in passenger services and comprehend passenger expectations (ACI, 2022).

Despite various challenges, digital solutions have been pivotal in addressing these issues. In 2020, investment in airports increased by 40% to improve capacity, operations, and the passenger experience (Zaharia & Pietreanu, 2018). Airports, which are complex environments offering a range of services, require effective tools for evaluating service quality. Given the competitive environment of the aviation industry, airport management must prioritise the enhancement of operations and passenger satisfaction. Consistent monitoring is essential to meet high service quality standards. The emergence of web-based platforms, such as Skytrax, Google Reviews, TripAdvisor, and social media, provides valuable traveller feedback. These platforms allow travellers to rate and review airport services, providing valuable insights into various service aspects. This feedback is a rich information source for travellers, service providers, and other stakeholders. Owing to the substantial volume of reviews available, machine learning is an effective tool for evaluating ASQ.

The competitive nature of the airport industry necessitates continuous improvement and innovation in service quality. This study leverages the Skytrax star rating system, a recognised measure of ASQ, which has improved through expertise in the field

(Skytrax, 2024). By analysing nine key variables — including queuing time, cleaning, seats, signs, food, retail, Wi-Fi, staff, and recommendations — this research focuses on the busiest airports in Southeast Asia, a critical region for aviation revenue. The aim is to develop a nuanced understanding of the factors influencing passengers' likelihood of recommending airports, particularly within the extended recovery timeline.

This study employs star ratings and recommendation data ("yes"/"no"), converting these into numerical formats suitable for analysis. It applies five machine learning algorithms — decision trees, random forests, support vector machines, neural networks, and gradient boosting machines — to the processed dataset. The objective is to devise a robust model for decision-making and strategic planning. This involves identifying factors that influence passengers' likelihood of recommending an airport and developing a classification model that aligns airport services with passenger expectations.

This research holds particular significance for ASEAN countries for two main reasons. First, there is a noticeable paucity of academic literature addressing airport service quality within the region. Second, the aviation sector in ASEAN countries has experienced substantial growth in recent years, underscoring the need for comprehensive service evaluations. The primary objective of this study is to develop a robust classification model designed to align airport services with passenger expectations. This model identifies and analyses the key determinants that significantly impact passenger satisfaction.

The structure of this paper is organised as follows: Section 2 reviews the relevant literature, followed by Section 3, which details the data collection process. Section 4 presents the experiments and results, and finally, Section 5 offers a conclusion.

## 1. Literature review

### 1.1. Airport service quality

Over the past two decades, the assessment of airport service quality (ASQ) has garnered significant attention within academic research. Studies by researchers such as Fodness and Murray (2007), Jiang and Zhang (2016), and Yeh and Kuo (2003) have predominantly utilised user surveys and expert evaluations to explore ASQ, employing questionnaires designed to capture passenger opinions. These investigations have illuminated critical discrepancies between passenger perceptions and their expectations of service quality. Enhancing passenger satisfaction has emerged as a key objective in airport management, directly influencing competitiveness and operational efficiency. The intricate relationship between airport aeronautical and non-aeronautical performance and passenger experiences has been well-documented in aviation scholarship (Wattanacharoensil et al., 2015). Notably, non-aeronautical revenue (NAR) has been identified as a vital contributor to the economic growth of the aviation sector. ASQ is widely acknowledged as a fundamental determinant of travellers' experiences and has a substantial impact on their journeys (Bezerra & Gomes, 2019; Del Chiappa et al., 2016; Freitas et al., 2021; Graham, 2009).

As the global benchmark established by the Airports Council International (ACI) for evaluating passenger satisfaction, ASQ underscores the aviation sector's dedication to not only meeting but exceeding passenger expectations (ACI, 2021a). A comprehensive evaluation approach incorporating both internal and external assessments is essential to effectively assess service levels and passenger satisfaction. Fodness and Murray (2007) emphasised the importance of systematically analysing passenger expectations to prioritise improvements across critical service dimensions, a sentiment echoed by Lubbe et al. (2011), who advocated for the integration of air travellers' perspectives in service-level assessments. In their analysis, Bezerra and Gomes (2020) employed structural equation modelling (SEM) to investigate the interplay of factors influencing passenger satisfaction, revealing insightful connections among passenger expectations, ASQ, switching costs, and loyalty dynamics. Similarly, Isa et al. (2020) conducted a study focused on the klia2 terminal of Kuala Lumpur International Airport, identifying eight key dimensions that contribute to overall passenger satisfaction through Partial Least Squares SEM (PLS-SEM).

Bulatović et al. (2023) utilised both ordinal regression and maximum likelihood structural equation modelling (ML SEM) to evaluate the efficacy of the Skytrax evaluation system, identifying facility comfort, wayfinding and signage, and restaurant outlets as significant indicators of airport service quality. The application of hybrid methodologies has been a notable trend in airport management research. For instance, Bakır et al. (2022) analysed online reviews of European airports using multiple regression analysis (MRA) and necessary condition analysis

(NCA), highlighting airport operators as pivotal influencers shaping passenger experiences. Additionally, Pholsook et al. (2023) conducted a comprehensive three-stage analysis to identify critical dimensions of ASQ, pinpointing airport facilities, wayfinding, and security as integral components essential for achieving overall passenger satisfaction.

With respect to digitisation, two prominent approaches have been adopted to obtain passenger sentiment: online questionnaires and data scraping from user-generated content (UGC). Digital and web-based online surveys enable airports to gather structured passenger opinions on various aspects of their experiences. Moreover, data scraping from passenger reviews such as Google Maps (Alanazi et al., 2024; Lee & Yu, 2018; Li et al., 2022), TripAdvisor (Nghiêm-Phú & Suter, 2018; Sezgen et al., 2019), Skytrax (Akan & Karataş, 2025; Alanazi, Li et al., 2024; Bakır et al., 2022; Brochado et al., 2024; Bulatović et al., 2023; Bunchongchit & Wattanacharoensil, 2021; Halpern & Mwesiumo, 2021; Homaid & Moulitsas, 2023; Wattanacharoensil et al., 2017; Yavuz et al., 2020), X (Barakat et al., 2021; Booranakittipinyo et al., 2024; Martin-Domingo et al., 2019), and airport online channels (Arasli et al., 2023) provides reliable information from passengers. Notably, Abouseada et al. (2023) conducted quantitative research using data from Skytrax, TripAdvisor, and other review sites, illustrating the value of scraping data to capture authentic passenger sentiments. Although several alternative platforms of customer's voice exist, Skytrax is a respected platform in aviation, offering valuable passenger review data for airport management research. The utilisation of platforms such as Skytrax to incorporate UGC highlights a significant trend, allowing airports to remain responsive to passenger sentiments and improving services to meet evolving customer expectations (Pholsook et al., 2024).

In recent years, there has been a significant increase in the use of machine learning and natural language processing techniques for evaluating ASQ, with UGC often serving as a key data source for analysing user satisfaction. Some studies have applied lexicon-based sentiment analysis to assess traveller sentiment (Kiliç & Çadirci, 2022; Song et al., 2020), whereas others have utilised traditional machine learning algorithms (Eswaran et al., 2019; Sulu et al., 2022). In research conducted by Rane and Kumar (2018), traveller sentiment was evaluated across six major U.S. airlines via seven classifiers, with the random forest classifier identified as the most effective for their dataset.

There are few studies that specifically apply machine learning to analyse Skytrax data concerning traveller sentiments and airport services. Gitto and Mancuso (2017) undertook the analysis of passenger satisfaction by analysing passenger reviews on Skytrax. Their analysis, which used the Semantria tool, classified reviews into positive, negative, or neutral sentiments. Moreover, Mizufune and Katsumata (2018) proposed a sentiment analysis approach employing latent Dirichlet allocation for the top 50 ranked airports, showing improvements in predicting review polarity. Recently, Pholsook et al. (2024) introduced a multimethod approach that focuses on leading Southeast Asian airports. Their findings identified queuing time and staff services as important factors affecting airport service quality.

### 1.2. Data mining models

In the context of evaluating passenger satisfaction with airport services in Southeast Asia, this research leverages several classification data mining models to derive meaningful insights from collected data. These models offer various methodologies for categorising and analysing passenger feedback, enhancing understanding and improving service delivery.

#### 1.2.1. Decision tree algorithm

A decision tree algorithm is utilised for both classification and regression tasks (Lively et al., 2024; Naveen et al., 2023). This method works by recursively dividing the dataset into smaller subsets on the basis of specific criteria, forming a tree-like structure. In this structure, internal nodes represent decisions on attributes, branches illustrate the outcomes of these decisions, and leaf nodes denote the final classifications. Therefore, the decision tree model uses the following equation to analyse the data:

- Splitting criteria:

  Entropy (H):
  $$H(S) = -\sum_{i=1}^{n} p_i \log_2(p_i) \quad (1)$$

  Information gain (IG):
  $$IG(S,A) = H(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} H(S_v)$$

- Tree construction:
  Recursive splitting:
  - Base case. If all samples within a node belong to a single class, generate a leaf node with that class label.

- Feature selection. Identify the feature that provides the highest information gain.
- Data splitting. The data are divided into subsets according to the values of the chosen feature.
- Iteration. Recursively implement the above steps on each subset until a stopping condition is reached.

Prediction:

- Tree traversal. For a new instance, the tree is navigated according to its feature values until it arrives at a leaf node.
- Classification. Assign the predicted class to the new instance on the basis of the majority class of the instances within that leaf node.

In this study, decision trees are applied to classify passenger satisfaction levels, offering insights into the various factors that impact customer experiences at airports in Southeast Asia.

### 1.2.2. RANDOM FOREST ALGORITHM

A random forest algorithm, known for its robustness and accuracy, is utilised to test passenger recommendations for airport services (Braikia et al., 2024). As an ensemble method, the random forest improves model reliability by aggregating predictions from multiple decision tree models to improve accuracy and robustness. A random forest performs a classification analysis as follows:

Given $M$ decision trees, the prediction of the random forest for a classification problem is typically the mode (majority vote) of the predictions made by each individual tree.

$$\hat{y} = mode(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_M) \quad (2)$$

where:
$\hat{y}$ is the final class predicted by the random forest.
$\hat{y}_i$ is the class predicted by the $i$-th decision tree.

This enables a comprehensive understanding of whether passengers are likely to recommend an airport service.

### 1.2.3. SUPPORT VECTOR MACHINE (SVM)

A support vector machine (SVM) is another supervised learning algorithm used in this research, primarily for classification tasks (Punne et al., 2024). Its capability extends to regression analysis and outlier detection. For classification, the support vector machine operates as follows:

$$f(x) = w \cdot \mathrm{x} + b \quad (3)$$

where:
x is the input feature vector.
w is the weight vector.
b is the bias term.

This study employs an SVM to explore classification outcomes related to passenger recommendations of airport services, aiming to identify influential factors.

### 1.2.4. NEURAL NETWORK

A neural network is a type of machine-learning model inspired by the structure and function of the human brain (Chinnarasri et al., 2012). It consists of interconnected layers of nodes (also called neurons) that process and transmit information. Neural networks are particularly well suited for tasks such as image recognition, speech processing, and complex pattern recognition. Therefore, the neural network operates as follows:

$$y = f(W \cdot x + b) \quad (4)$$

where:
$x$ is the input vector.
$W$ is the weight matrix.
$b$ is the bias vector.
where $f$ is the activation function applied to the linear combination of inputs.
where $y$ is the output of the neural network.

This experiment uses a neural network to classify how each satisfaction level from various features determines the recommendation class for airport services.

### 1.2.5. GRADIENT BOOSTING MACHINES

Gradient boosting machines (GBMs) are deployed, offering potent capabilities for classification and regression (Touzani et al., 2018). This ensemble technique enhances predictive accuracy by aggregating the outputs of several weak learners, predominantly decision trees, to formulate a robust model. Thus, the GBM operates according to the following equation:

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x) \quad (5)$$

where:
$F_m(x)$ is the model's prediction after $m$ iterations.

$F_{m-1}(x)$ is the prediction from the previous iteration.
$h_m(x)$ is the new weak learner (typically a decision tree) added at iteration $m$.
$\eta$ is the learning rate, controlling the contribution of each new model.

Therefore, this research uses the GBM as the final technique to experiment with predicting whether a passenger will recommend an airport service on the basis of their satisfaction level.

Collectively, these data mining models provide a robust framework for analysing passenger satisfaction, offering insights that can guide strategic improvements in service quality and customer experience at Southeast Asian airports. Through these methodologies, the research aims to identify factors that significantly influence passenger satisfaction and recommendations.

### 1.3. Evaluation of the binary classification model

To rigorously evaluate the performance of the binary classification model, this study incorporates several analytical tools and techniques, each offering a unique perspective on model efficiency and feature impact. The study specifically utilises the confusion matrix, ROC curve, and SHAP values to provide comprehensive insights into the model's capabilities and the influential features in passenger satisfaction classification.

#### 1.3.1. Confusion matrix

The confusion matrix is central to assessing classification model performance, particularly in binary or multiclass contexts (Sumiati et al., 2022). For a binary classification problem, the following metrics are derived from the confusion matrix:

- Accuracy. The proportion of correctly predicted instances (positive and negative) out of all predictions.
- Precision. The proportion of true positive predictions among all positive predictions made by the model.
- Recall (sensitivity or true positive rate). The proportion of actual positives that were correctly identified by the model.
- F1 score. The harmonic means of precision and recall, providing a balance between the two metrics.

#### 1.3.2. Balance accuracy

Balance accuracy is a metric used to evaluate the performance of a classification model, especially in situations where there is a class imbalance (one class has significantly more samples than the other) (Doo et al., 2024). It is calculated as the average of the recall obtained on each class, effectively giving equal weight to both classes regardless of their frequency. The formula for balanced accuracy is as follows:

$$Balanced\ Accuray = \\ = \frac{Sensitivity\ (Recall\ of\ Positive\ Class)+Specificity(Racall\ of\ Negative\ Class)}{2} \quad (6)$$

This metric is useful because it provides a more unbiased assessment of model performance across both classes, preventing the accuracy score from being skewed by class imbalance. In an experiment, a high balanced accuracy score indicates that the model performs well at classifying "recommend" and "not recommend" classes without bias toward either class.

#### 1.3.3. Matthews correlation coefficient (MCC)

The Matthews correlation coefficient (MCC) is a performance metric used to evaluate the quality of binary classifications, particularly in cases with class imbalance (Chicco et al., 2021). The MCC considers true and false positives and negatives, providing a comprehensive measure of the correlation between the predicted and actual classifications. It returns a value between -1 and +1, where:

+1 indicates a perfect prediction,

0 suggests a prediction no better than random chance,

-1 indicates total disagreement between the predictions and actual values.

The formula for the MCC is as follows:

$$MCC = \frac{(TP\times TN)-(FP\times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (7)$$

where:
TP = true positives
TN = true negatives
FP = false positives
FN = false negatives

The MCC is especially valuable because it provides a balanced measure that works well even if the classes are very different sizes. A high MCC score in your model indicates a strong correlation between

actual and predicted classes, suggesting reliable performance across "recommend" and "not recommend" categories.

#### 1.3.4. ROC curve

The ROC curve is a graphical plot used to evaluate the performance of a binary classification model. It illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1 — specificity) across different threshold settings (Chang & Newman, 2024). The key concepts of the ROC are as follows:

- True positive rate (TPR). Also known as sensitivity or recall, the TPR measures the proportion of actual positives that are correctly identified by the model.
- False positive rate (FPR) measures the proportion of actual negatives that are incorrectly identified as positive by the model.
- ROC curve plots the TPR (y-axis) against the FPR (x-axis) for various threshold values. Each point on the curve represents a different decision threshold. A model with a perfect classifier would have a point at the top left corner (TPR = 1, FPR = 0), representing 100 % sensitivity and 0 % false positives.
- Area under the curve (AUC) is a single scalar value that represents the overall performance of the model. It ranges from 0 to 1, where a model with an AUC of 1 indicates perfect performance and a model with an AUC of 0.5 indicates performance no better than random guessing.

#### 1.3.5. SHAP values

SHAP (SHapley Additive exPlanations) values provide a method to interpret the outputs of machine learning models by quantifying each feature's contribution to predictions (Mao et al., 2023). Rooted in game theory, SHAP values ensure a fair and consistent framework for assessing the importance of global and individual prediction features. As model-agnostic tools, they are applicable across various machine learning models, enhancing model transparency and trust. SHAP values are particularly valuable in identifying critical features that significantly influence the target label's classification and for debugging and improving model interpretability.

By integrating these sophisticated evaluation methods, this study analyses the binary classification model used for assessing passenger satisfaction with airport services in detail. The insights gained from these analyses aim not only to increase the accuracy and robustness of the classification model but also to uncover strategic service features crucial for satisfying passengers in Southeast Asia's airports.

## 2. Research methods

Skytrax serves as a global leader in the evaluation and certification of airline and airport services, providing a platform for passengers to share and rate their experiences with specific airport services. For this study, sentiment analysis was conducted on passenger feedback sourced from the Skytrax website (https://www.airlinequality.com/), with data extraction performed via the R programming language. This analysis focused on user-generated content (UGC) related to service ratings across nine airport attributes.

The scope of this study is limited to the leading airports in Southeast Asia (OAG, 2024), which include Changi Airport, Denpasar Ngurah Rai International Airport, Don Mueang International Airport, Hanoi Airport, Ho Chi Minh City Airport, Jakarta Soekarno-Hatta Airport, Kuala Lumpur International Airport, Manila Ninoy Aquino International Airport, and Suvarnabhumi International Airport.

The data collection process on Skytrax commenced with gathering demographic information from passengers, including details such as name, country of residence, airport visited, date of visit, and purpose of the trip. Subsequently, passengers were invited to rate eight dimensions of airport service: queuing time, clean, seats, signs, food, retail, Wi-Fi, and staff. These ratings form the basis of Skytrax's airport star ratings, which range from one to five stars. A rating of 1 indicates poor service, whereas a rating of 5 signifies excellent service. Furthermore, passenger reviews are categorised as "recommended" or "not recommended", providing an additional layer of qualitative data for analysis (Table 1).

This approach highlights the importance of voluntary UGC in understanding passenger sentiment and experiences at key Southeast Asian airports, facilitating a detailed analysis of service quality and customer satisfaction. By integrating comprehensive demographic data with service ratings, this study aims to provide insights into the factors influencing passenger recommendations and satisfaction with the region's busiest airports.

Tab. 1. Dataset characteristics

| No. | ASQ Dimensions | Characteristics | Data Type |
|---|---|---|---|
| 1 | Passenger satisfaction with queue management | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 2 | Airport cleaning service satisfaction | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 3 | Airport seating satisfaction | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 4 | Airport signage satisfaction | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 5 | Airport food service satisfaction by passengers | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 6 | Airport retail satisfaction by passengers | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 7 | Internet satisfaction of passengers | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 8 | Airport staff satisfaction by passengers | Satisfaction level 1 – 5 on a Likert scale | Ordinal Scale |
| 9 | Recommendation | "Yes" indicates a recommendation, and "No" indicates no recommendation. This is a targeting label. | Binary Data |

Tab. 2. Demographic information (n = 1061)

| Passenger Details | Category | | Frequency | Percentage |
|---|---|---|---|---|
| **Passenger experience** | Arrival and departure | | 425 | 40.06% |
| | Departure only | | 375 | 35.34% |
| | Transit | | 155 | 14.61% |
| | Arrival only | | 106 | 9.99% |
| **Passenger type** | Solo trip | | 376 | 35.45% |
| | Business trip | | 243 | 22.90% |
| | Family trip | | 234 | 22.05% |
| | Couple trip | | 208 | 19.60% |
| **Passenger by continent** | Asia | Southeast Asia | 471 | 44.39% |
| | | East Asia | 59 | 5.56% |
| | | South Asia | 31 | 2.92% |
| | | West Asia | 21 | 1.98% |
| | Oceania | Australia and New Zealand | 163 | 15.36% |
| | Europe | Northern Europe | 133 | 12.54% |
| | | Western Europe | 60 | 5.66% |
| | | Southern Europe | 9 | 0.85% |
| | | Eastern Europe | 8 | 0.75% |
| | North America | North America | 89 | 8.39% |
| | | Central America | 1 | 0.09% |
| | South America | South America | 2 | 0.19% |
| | Africa | Southern Africa | 1 | 0.09% |
| | Nonidentified | Nonidentified | 13 | 1.23% |

The dataset compiled for this study encompasses reviews from nine significant airports in Southeast Asia collected over a period extending from 2015 to June 2024. This dataset comprises 1,061 valid passenger reviews and offers a detailed portrayal of passenger demographics, as summarised in Table 2. This demographic information sheds light on various passenger attributes related to travel experiences and geographical origins.

Regarding the nature of passenger experience, the most common category, identified by 40.06% of respondents (n = 425), includes those who partici-

pated in both arrival and departure procedures. This group is closely followed by passengers who reported departing only, comprising 35.34% (n = 375) of the sample. Transit passengers represent 14.61% (n = 155) of all passengers, whereas those arriving only account for 9.99% (n = 106) of all passengers.

An analysis of the passenger types reveals that solo travellers constitute 35.45% (n = 376) of the dataset. Business travellers make up 22.90% (n = 243), whereas family travellers account for 22.05% (n = 234). Couples travelling together represented 19.60% (n = 208) of the respondents.

The data also provide insights into passenger origins by continent. Within Asia, Southeast Asian travellers predominated, accounting for 44.39% (n = 471) of the sample. This is followed by travellers from East Asia (5.56%, n = 59), South Asia (2.92%, n = 31), and West Asia (1.98%, n = 21). In Oceania, travellers from Australia and New Zealand contributed 15.36% (n = 163) to the dataset.

In the European context, Northern European passengers constitute a significant portion, accounting for 12.54% (n = 133) of the reviews, followed by Western Europe (5.66%, n = 60), Southern Europe (0.85%, n = 9), and Eastern Europe (0.75%, n = 8). North American passengers represented 8.39% (n = 89), with a minor contribution from Central America (0.09%, n = 1). Additionally, South America accounts for 0.19% (n = 2), whereas southern Africa contributes 0.09% (n = 1). There is also a small percentage (1.23%, n = 13) of passengers whose continental origins were not identified.

This diverse dataset not only captures a wide range of travel experiences and traveller types but also represents a broad spectrum of geographical origins, providing a comprehensive foundation for analysing passenger satisfaction and service quality at key Southeast Asian airports.

## 3. Research results

On the basis of the research objectives and literature review, this study selects supervised learning techniques, including decision trees, random forests, SVMs, neural networks, and GBMs, to develop a model that can identify whether a passenger will recommend an airport service in Southeast Asia based on their satisfaction levels across various dimensions of airport service. Consequently, this research develops an experimental model designed to address each research objective as follows.

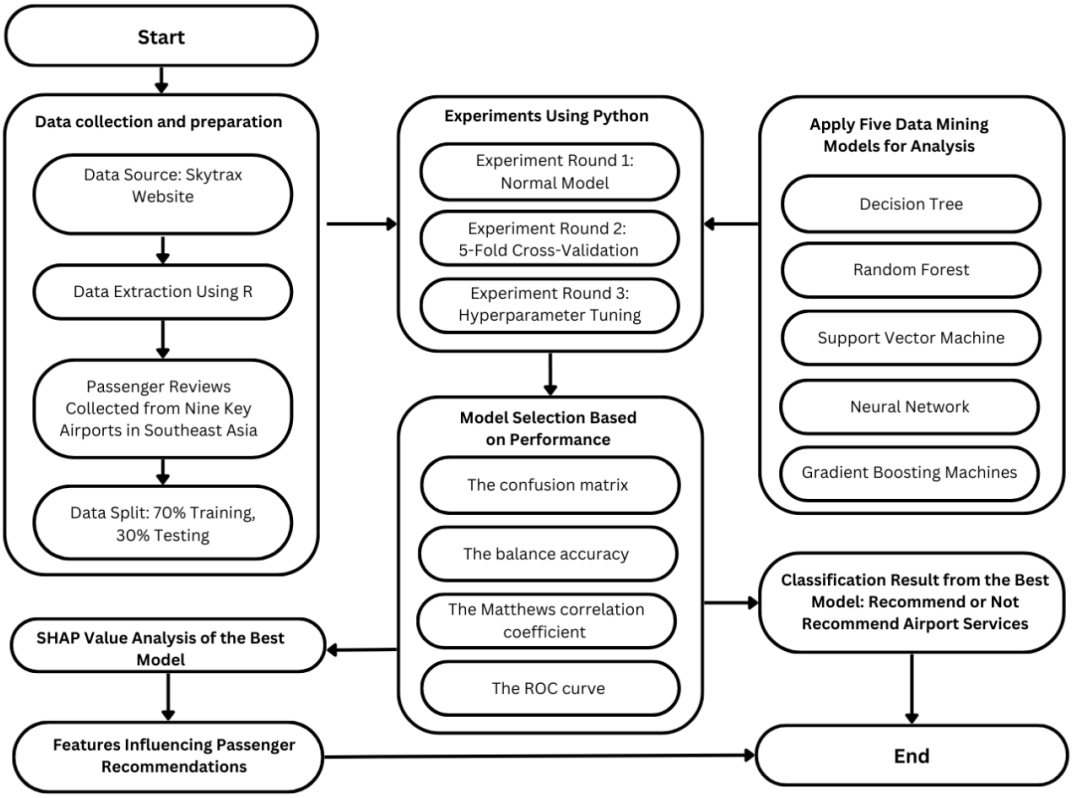



Fig. 2. Experimental model for enhancing airport services: data-driven analysis of passenger satisfaction and service quality in Southeast Asia

This research experiments with a dataset using five data mining techniques, as described by equations (1) through (5), and explains the results via confusion metrics, balance accuracy, the Matthews correlation coefficient (MCC), and the AUC, with the outcomes described next.

### 3.1. Decision tree experiment

The decision tree model performs well in classifying whether passengers would recommend airport services, with an accuracy of 85.58% and a balanced accuracy of 85.57%, indicating that it handles both recommendation options well. The Matthews correlation coefficient is 0.71, which shows a strong match between the classification results and the actual results. The precision and recall for "recommend" and "not recommend" are high, with F1 scores of approximately 0.85, indicating that the model is consistent. The confusion matrix shows that most classifications are correct, and the AUC score of 0.86 indicates that the model is good at distinguishing between passengers who would recommend and those who would not. This model provides useful insights to help improve airport services on the basis of passenger feedback.

To improve model performance via hyperparameter tuning, grid search tuning significantly improved the model performance, resulting in the following optimal parameters: criterion='entropy', max_depth=10, min_samples_leaf=1, and min_samples_split=20. These settings led to the best cross-validation score of 0.89. During 5-fold cross-validation, the model showed accuracy scores of 0.798, 0.948, 0.896, 0.821, and 0.892, with a mean accuracy of 87% and a low standard deviation of 0.05, indicating consistent performance across different data splits. On the test set, the tuned model achieved an accuracy of 86.83% and a balanced accuracy of 86.55%, suggesting that the model performs fairly well in classifying both recommendation classes.

The Matthews correlation coefficient (0.73) points to a strong positive correlation between classification and actual outcomes. The precision for "not recommend" was 0.87, and the recall was 0.90, indicating that the model correctly identifies passengers who would not recommend the service with high accuracy. For "recommend", the precision was 0.87, and the recall was 0.83, suggesting that the model also performs well in classifying recommendations. The overall F1 scores for both classes, 0.88 for "not recommend" and 0.85 for "recommend", reflect a good balance between precision and recall.

The confusion matrix further illustrates the model's performance, with 156 correct predictions for "not recommend" and 121 correct classifications for "recommend". The model had 18 false positives and 24 false negatives, which is an improvement over prior models and results in fewer misclassifications. The AUC score of 0.93 indicates that the model is highly capable of distinguishing between the two recommendation classes, demonstrating strong overall performance, as shown in Fig. 3.

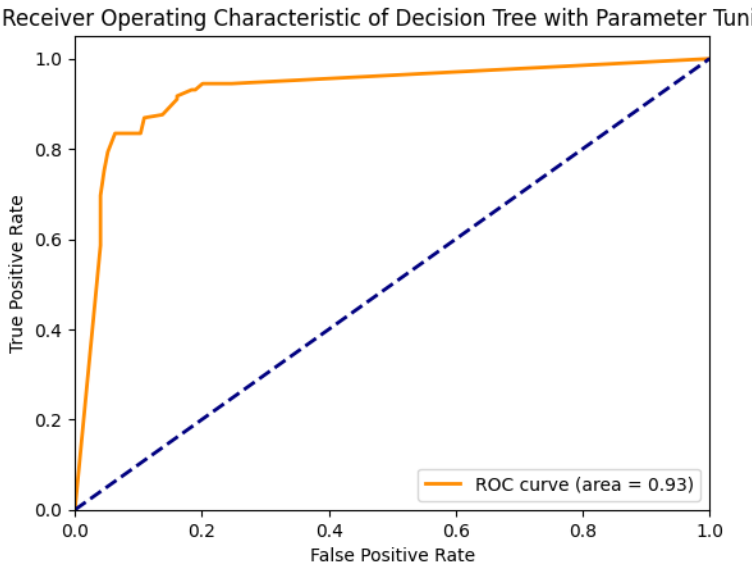



Fig. 3. Receiver operating characteristics of a decision tree with parameter tuning

According to Fig. 3, this high AUC score indicates that the model is robust for data classification and ensures that the model can be used to recommend taking a service at the airport on the basis of the features of the airport service.

To summarise the decision tree model experiment, the decision tree has strong accuracy and low variability due to good results from cross-validation, which is used to improve model reliability and reproducibility. The decision tree has balanced values between precision, recall, and F1 scores for both classes. This balance shows consistent and effective data classification performance. The model achieved a high AUC value, confirming its efficiency for class separability and its ability to classify unseen data.

### 3.2. Random forest

In the experiment, the random forest model demonstrated high performance, with an accuracy of 90.3%. It exhibited balanced precision (90% for both classes) and high recall rates, particularly for "not recommending taking an airport service" (91%), making it a robust model in this experiment. The balanced accuracy is 0.9046, which means that the random forest model correctly classified both classes: "not recommended" and "recommended" for taking an airport service. Additionally, the MCC of 0.8102 is high, indicating a strong positive correlation between the classified values and the actual data in the label feature. The precision is 90.9%, and the recall is 91.9%, whereas the "recommended" class has a precision value of 90.2% and a recall value of 88.97%. The macro F1 score is 90.5%, and the weighted average F1 score is 90.6%, indicating consistent performance across classes. The confusion matrix shows low false positives at 14 and false negatives at 16, reflecting the model's accuracy in classification. The AUC is 0.96, which means that the random forest model has high performance in classifying between "recommended" and "not recommended". Moreover, the model has a low tendency to produce false positives as well as false negatives.

To improve model performance, this research added cross-validation, finding that the random forest model achieved strong results, with a test accuracy of 89.66% and a mean cross-validation accuracy of 89.16%, showing consistent performance across data splits. The model maintained balanced accuracy, reaching 89.60% on the test set and 88.99% on average in cross-validation, which suggests that it can classify "recommended" and "not recommended" airport services effectively. The Matthews correlation coefficient (MCC) of 0.79 indicates a strong positive link between the classification results and actual results, indicating balanced performance. The classification report shows high precision, recall, and F1 scores for both classes, with precision values of 90.75% for "not recommended" and 88.36% for "recommended" and similar recall values of 90.23% and 88.97%, respectively, indicating accurate classification for both. An AUC of 0.95 suggests that the model can very effectively separate the classes. The confusion matrix confirms this, with only 17 false positives and 16 false negatives supporting the model's accuracy. Overall, these metrics show that the random forest model is robust and suitable for accurately classifying airport service recommendations. Note that the performances are lower than when the experiment splits the dataset into two sets: training and testing. This is because when cross-validation is used, which creates five opportunities to train a random forest model, it reveals a more realistic pattern of the data. Therefore, the experiment can better assess how the model will perform with unseen data.

In this experiment, the random forest model uses default settings because the focus is on evaluating performance with cross-validation, not on tuning specific parameters. By default, 100 trees — n_estimators — are used to make classifications, providing a balance between accuracy and computational efficiency. The maximum depth, max_depth, is set to None, allowing each tree to grow until all the leaves are pure, which helps capture complex patterns but can also lead to overfitting. The minimum number of samples split, min_samples_split, is 2, so a node must have at least 2 samples to split, which can create deeper trees. The minimum number of samples per leaf, min_samples_leaf, is 1, meaning that each leaf node must have at least one sample, which could lead to capturing very specific patterns and possibly increasing overfitting. Finally, the number of features at each split, max_features, is set to the square root of the total features (sqrt), a common choice in classification that adds randomness, helping to reduce correlations between trees and improve generalisation.

To improve performance, this research uses hyperparameter tuning with the grid search technique. The optimised random forest model shows strong performance, with an accuracy of 91.5% on the test set, meaning that it accurately classifies both "recommended" and "not recommended". The balanced accuracy is 91.4%, showing the model's ability to handle both classes well. The Matthews correlation

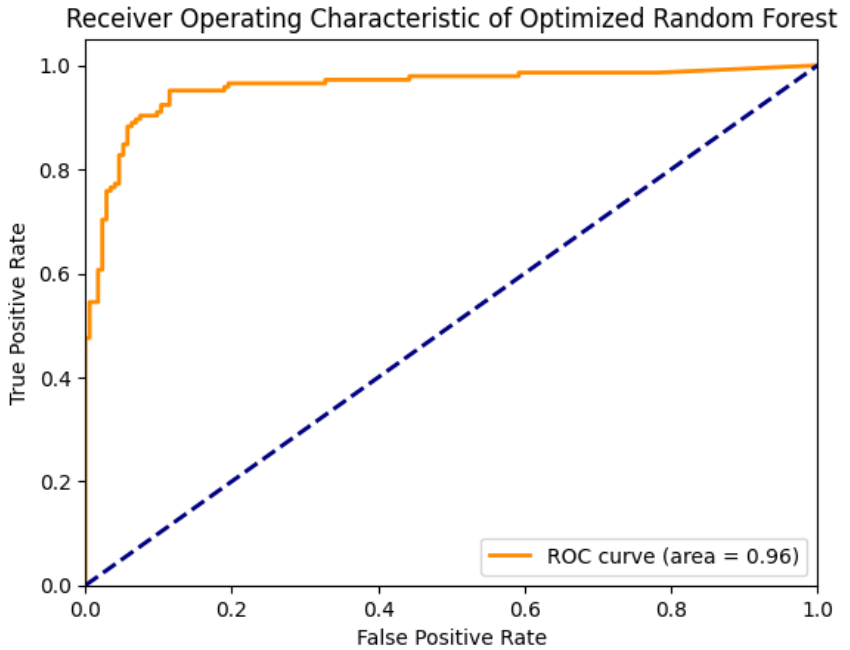



Fig. 4. Receiver operating characteristics of the optimised random forest

coefficient of 0.83 reflects a strong positive link between the model's classification and the actual outcomes, meaning that it performs reliably and balances both classes well. The classification report also shows high precision, recall, and F1 scores for each class: 92.0% precision and 92.5% recall for "not recommended" and 90.97% precision and 90.3% recall for "recommended". The macro- and weighted average F1 scores of 91.5% confirm consistent performance across classes, which is further supported by the confusion matrix showing 161 true negatives, 131 true positives, 13 false positives, and 14 false negatives. The AUC score is high at 0.96, indicating that the model has an excellent ability to separate the classes, as shown in Fig. 4.

For hyperparameter settings, these choices were made to improve the model's stability and repeatability. The number of trees, n_estimators, was set to 200 after several values, including 50, 100, and 200, were tested. The choice of 200 gave the best results, as more trees often improve accuracy by reducing variance. The maximum depth, max_depth, was set to 30 to find the best balance between bias and variance, allowing the model to capture patterns without overfitting. The cross-validation results revealed that deeper trees did not add much improvement, so 30 was ideal.

The minimum number of samples split, min_samples_split, of 5 and the minimum number of samples per leaf, min_samples_leaf, of 2, control the model complexity. They set the minimum number of samples needed to split a node or form a leaf. These settings help avoid overly specific rules that could cause overfitting, keeping the model general and effective. Finally, the number of features at each split, max_features, was set to "sqrt", meaning that the model randomly picks the square root of the total number of features at each split. This choice, which is common in random forests, reduces variance and prevents the model from depending too much on any one feature.

Overall, these hyperparameter choices, which are based on thorough cross-validation, make the random forest model accurate and robust while minimising overfitting. This helps ensure that the model is repeatable and reliable.

After hyperparameter tuning, the random forest model improved the accuracy, balance accuracy, and overall classification performance efficiency. This makes it suitable for use as a recommendation system for airport services. The research adjusts parameters such as the number of trees, maximum depth, minimum samples for splits and leaves, and number of features per split. This helps the model balance bias and variance better. Now, the model provides more accurate and reliable recommendations to passengers. This helps them make better choices about airport services with more information.

### 3.3. Support vector machine: SVM

The SVM model effectively classifies passenger recommendations regarding airport services, achieving high accuracy (90%), balanced accuracy (90%), and an AUC of 0.95. These metrics indicate that the model reliably differentiates between those who would and would not recommend the airport on the basis of features such as cleanliness, staff performance, and Wi-Fi access. With a high Matthews correlation coefficient (0.80), the model demonstrates robust predictive power, even when accounting for classification errors. Overall, the results affirm that SVM is a suitable and reliable approach for analysing customer satisfaction, offering valuable insights for improving airport services.

This research relies on cross-validation for the experiments. The cross-validation results confirm the SVM model's robust and consistent performance across multiple data splits, with a mean accuracy of 0.89 and a standard deviation of 0.04. This slight reduction from the single test split accuracy of 0.90 suggests minor variability in performance depending on the data split, yet the model remains stable. The balanced accuracy of 0.90 and MCC of 0.80 show that the model effectively handles "recommend" and "not recommend" classifications with minimal class-specific issues. Additionally, high precision, recall, and F1 scores in the classification report and a consistent AUC of 0.95 indicate excellent discriminatory power. Overall, cross-validation verifies the model's suitability for this task, confirming its generalisability and reliability across different data subsets.

Finally, this research aims to improve the data classification performance of the SVM model via parameter tuning, which is a grid search. After the experiments, the optimised SVM model demonstrated robust performance in predicting whether passengers would recommend airport services, achieving an accuracy of 89.3%, indicating that it correctly classified nearly 9 out of 10 cases. A balanced accuracy of 89.2% confirms the model's consistency across both recommendation classes, effectively handling any potential class imbalance. The Matthews correlation coefficient (MCC) of 0.78 shows a strong positive correlation between the actual and predicted values, underscoring the model's reliability in accurately balancing true and false predictions. The model also shows stability across various data subsets, as evidenced by a cross-validation mean accuracy of 89.5% with a low standard deviation of 3.3%. High precision (89.8% for "no" and 88.8% for "yes") and recall (90.8% for "no" and 87.6% for "yes") values, along with strong F1 scores for both classes, indicate that the model effectively minimises false positives and negatives. The AUC of 0.95 highlights its excellent discriminatory ability between recommendation categories, as shown in Fig. 5.

This optimised SVM model achieves high accuracy, balanced performance across both classes and excellent stability, as confirmed by cross-validation results. With high precision, recall, and AUC, the

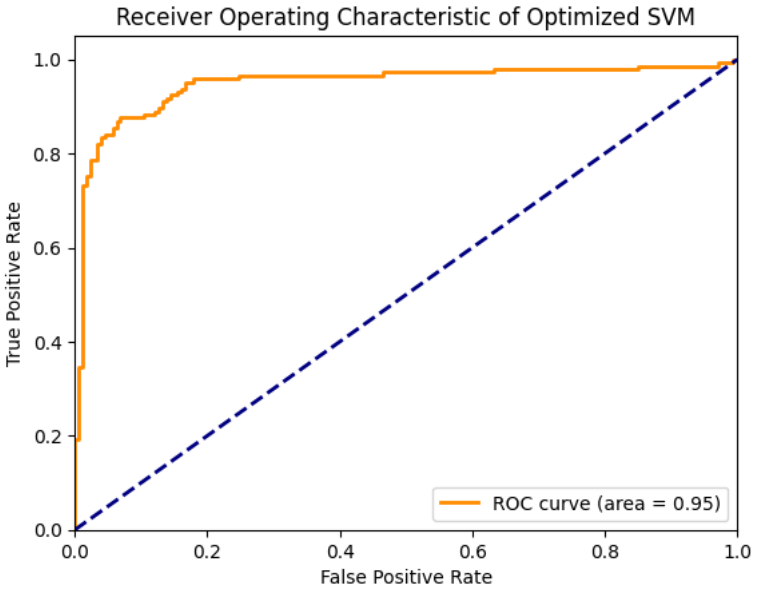



Fig. 5. Receiver operating characteristics of the optimised SVM

model is well-suited for accurately predicting passenger recommendations for airport services. The strong MCC further confirms its reliability in handling the classification task effectively.

### 3.4. Neural network

The neural network model effectively classifies whether passengers would recommend airport services, achieving a high overall accuracy of 90.3%. This performance highlights the model's ability to identify patterns in service-related factors, such as queue efficiency, cleanliness, seating quality, signage, food and beverage options, shopping facilities, Wi-Fi access, and staff performance. The network architecture consists of two hidden layers with 16 and 8 neurons, respectively, using the ReLU activation function to introduce nonlinearity, effectively preventing vanishing gradients in deeper networks. The output layer applies the sigmoid activation function to output probabilities suitable for binary classification.

The Adam optimiser is employed, utilising an adaptable learning rate of 0.001 for efficient training with sparse gradients. The model was trained with a batch size of 10 over 50 epochs, which allowed it to converge and generalise well without requiring additional regularisation techniques, such as dropout or L2 regularisation. This choice was validated by the model's balanced accuracy of 90.2%, which demonstrated consistent performance across both recommendation classes and effectively handled class imbalance.

In support of the model's reliability, the Matthews correlation coefficient (MCC) of 0.80 indicates a strong correlation between the true and predicted classifications, reinforcing the model's ability to minimise both false positives and negatives. The precision values (91.3% for "no" and 89.0% for "yes") and recall values (90.8% for "no" and 89.7% for "yes") confirm its robust classification capabilities, with F1 scores of 91.1% for "no" and 89.3% for "yes", indicating a balanced trade-off between precision and recall. Additionally, an AUC of 0.96 demonstrates the model's strong discriminatory power in distinguishing between recommendation categories. Overall, the model's reliable performance and robustness suggest that it is well suited for generating valuable insights into passenger satisfaction with airport services, providing actionable data-driven support for enhancing service quality and the overall passenger experience.

The neural network model was subsequently evaluated via cross-validation to confirm and enhance its performance, which revealed consistency across various subsets of data. Initially, a single train–test split evaluation reported an accuracy of 90.3%, whereas cross-validation yielded a mean accuracy of 88.7%, indicating the model's robust ability to generalise. The balanced accuracy remained high, with a mean of 88.6% in cross-validation, confirming the model's consistent handling of class imbalance across different folds.

In both evaluations, the Matthews correlation coefficient (MCC) was strong, with an initial value of 0.80 and a cross-validation MCC of 0.77, suggesting a reliable correlation between the true and predicted classifications. This reliability supports the model's consistent accuracy in identifying both the recommendation and non-recommendation classes, as further reflected in the confusion matrix. The aggregate confusion matrix revealed 523 true negatives, 64 false positives, 56 false negatives, and 418 true positives, indicating that the model effectively minimises both types of classification errors.

While precision and recall were not calculated directly for each cross-validation fold, the initial split provided high values, with precisions of 91.3% for "no" and 89.0% for "yes" and recalls of 90.8% for "no" and 89.7% for "yes". These metrics confirm the model's robust classification capabilities, further supported by a mean AUC of 0.96 across cross-validation, which is consistent with the initial split evaluation. This high AUC reaffirms the model's excellent discriminatory power in distinguishing between recommendation categories.

The neural network architecture, comprising two hidden layers with 16 and 8 neurons via ReLU activation, the Adam optimiser with a learning rate of 0.001, a batch size of 10, and 50 epochs, has proven effective, delivering strong performance without additional regularisation techniques. Overall, the cross-validation results validate the initial findings, demonstrating that the model performs consistently well across multiple data subsets, which supports its suitability for providing actionable insights into passenger satisfaction with airport services.

This experiment relies on hyperparameter tuning and sets an experimental design as follows:

#### 3.4.1. Network architecture

The neural network architecture was designed with one to three hidden layers, tuned via random search via a Keras tuner to balance model complexity

with performance. The final model selected consisted of two hidden layers.

Each hidden layer was tuned to have eight to 64 neurons in steps of eight, resulting in an optimal configuration of 16 neurons in the first hidden layer and eight neurons in the second hidden layer. These values were selected on the basis of cross-validation results, with higher neuron counts discarded to prevent overfitting without significant performance gains.

### 3.4.2. Activation functions

ReLU activation was used in each hidden layer. ReLU is effective in reducing vanishing gradient issues, which is particularly beneficial in deeper networks, as it allows faster convergence and more efficient learning of complex patterns.

The sigmoid activation function was applied to the output layer to produce probability estimates suitable for binary classification, mapping the output to values between 0 and 1.

### 3.4.3. Hyperparameter tuning

Hyperparameter tuning was conducted via a random search, allowing for efficient exploration of different configurations. This approach tested a range of values for the number of layers, neurons, dropout rates, learning rates, and optimiser options to identify the best-performing model setup.

### 3.4.4. Optimiser and parameters

The Adam optimiser was selected for its ability to adapt learning rates dynamically on the basis of gradient updates, which enhances performance on complex datasets.

The learning rate was tuned logarithmically between 10^(-4) and 10^(-2), with a final selection of 0.001, to ensure stable convergence without rapid oscillations.

Although SGD with momentum was considered, Adam's default configuration proved sufficient; thus, the momentum and decay parameters were not further tuned.

### 3.4.5. Batch size and epochs

The model was trained with a batch size of 10 over 50 epochs. Preliminary testing indicated that this batch size provided a balance between memory efficiency and gradient update frequency. Training over 50 epochs allowed sufficient learning time to capture patterns without excessive training cycles, which could lead to overfitting.

### 3.4.6. Regularisation techniques

Dropout was employed as a regularisation technique during hyperparameter tuning. Dropout rates between 0.1 and 0.5 were explored, with the final

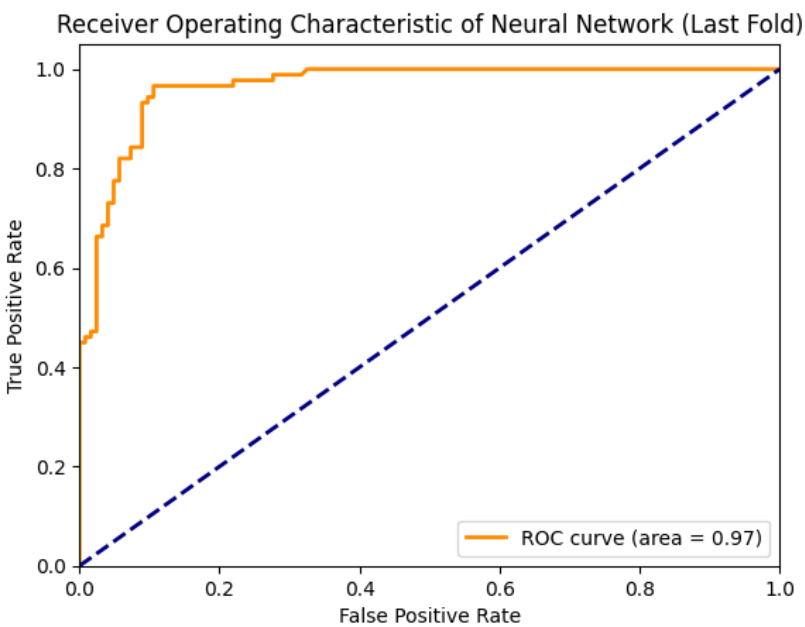



Fig. 6. Receiver operating characteristic (ROC) characteristics of the neural network (last fold).

model including a dropout rate of 0.2. This helps prevent overfitting by randomly deactivating neurons during training, compelling the model to learn more generalised features.

The experimental results indicate that the hyperparameter-tuned neural network model performs well, achieving a mean accuracy of 89.3%, a mean balanced accuracy of 89.1%, a Matthews correlation coefficient (MCC) of 0.78, and a mean AUC of 0.965 across cross-validation folds. The high accuracy and balanced accuracy confirm the model's consistency across the recommendation and non-recommendation classes. The MCC score indicates a reliable correlation between true and predicted classifications, whereas a high AUC indicates excellent discriminatory power, as shown in Fig. 6.

The aggregate confusion matrix further highlights this performance, with 534 true negatives, 53 false positives, 61 false negatives, and 413 true positives, indicating strong predictive accuracy and controlled misclassifications. The neural network model demonstrated effective performance in classifying passenger satisfaction with airport services, achieving an overall accuracy of 91.8% with high precision and recall values for "not recommended" (class 0) and "recommended" (class 1) services. Specifically, the model achieved a precision of 92.8% and a recall of 91.8% for predicting "not recommended" instances, along with a precision of 90.2% and a recall of 91.4% for "recommended" instances. These metrics indicate the model's balanced capability in identifying satisfied and dissatisfied passengers without bias toward either class, as reflected by the macro- and weighted F1 scores of 91.6%.

This hyperparameter-tuned neural network model achieved high performance through well-optimised architecture configuration, activation functions, and regularisation techniques. The cross-validation results confirm its stability across different data subsets, with highly balanced accuracy and MCC, indicating effective handling of class imbalance and low misclassification rates. This robust configuration is suitable for accurately predicting passenger recommendations and can offer valuable insights for enhancing airport service quality.

### 3.5. Gradient boosting model: GBM

The GBM model demonstrated strong performance in classifying passenger recommendations, achieving an accuracy of 90.60%. This high accuracy indicates that the classifier correctly identified 90.60% of the recommendations across both classes, highlighting its robust classification capabilities. Similarly, the balanced accuracy score was high at 90.52%, suggesting that the model performed equally well across both classes without bias. This metric is particularly valuable in cases of class imbalance, ensuring that the "recommend" and "not recommend" categories are classified with comparable accuracy.

The MCC value of 0.81 further demonstrates a strong correlation between the observed and predicted classifications, underscoring the model's robustness even when faced with potentially imbalanced class distributions. An MCC close to 1 suggests a high-quality prediction with minimal random error, making it a reliable measure of model performance. In terms of precision, the model achieved a value of 0.90 for the "recommend" class and 0.91 for the "not recommend" class, indicating accuracy in identifying true recommendations and non-recommendations. Both classes achieved recall values of approximately 0.90, demonstrating the model's effectiveness in capturing true instances for each class. The F1 scores were approximately 0.90 for both categories, indicating a good balance between precision and recall.

The model's confusion matrix revealed that it correctly classified 159 out of 174 instances as "not recommend" and 130 out of 145 instances as "recommend". With only 15 misclassifications in each class, the model's high precision and recall are further emphasised. Additionally, the AUC score of 0.95 reflects an excellent level of distinction between the "recommend" and "not recommend" classes, implying that the model effectively differentiates between passengers likely to recommend airport services and those who are not.

The experimental findings reveal that the features related to airport services (queue, cleanliness, seating, signage, F&B, shopping, Wi-Fi, and staff) are influential in determining passenger willingness to recommend these services to others. The model's high accuracy, balanced accuracy, and AUC highlight the potential of the GBM as a classifier in this context, while the MCC score reinforces the model's robustness under possibly imbalanced conditions. In conclusion, this experiment demonstrates that the gradient boosting machine is a highly effective tool for classifying passenger recommendations for airport services, confirming that the identified service features strongly correlate with passenger satisfaction and willingness to recommend. Future research could involve experimenting with alternative classification

models or integrating additional features to further improve the prediction accuracy.

To improve the GBM performance, this research relies on cross-validation. The updated experiment with 5-fold cross-validation and evaluation on the test set reveals consistent model performance, underscoring the robustness and reliability of the gradient boosting machine (GBM) classifier in predicting passenger recommendations for airport services.

The cross-validation results provide insights into how well the model performs across different splits of the training data. The mean cross-validation accuracy across the five folds was 88.68%, which was slightly lower than the test set accuracy of 90.91%. This indicates that the model generalises well to new data, although it performs slightly better on the test set, possibly owing to slight variations in the data distribution. Similarly, the mean cross-validation balanced accuracy was 88.60%, which closely aligns with the test set balanced accuracy of 90.75%. This metric reinforces that the model is capable of classifying "recommend" and "not recommend" classes with comparable accuracy across multiple data splits without favouring one class over the other. The mean cross-validation AUC, used here as a proxy for the Matthews correlation coefficient (MCC), was 0.959—nearly identical to the test set AUC of 0.96. This consistency in AUC values reflects the model's strong ability to distinguish between the two classes, demonstrating its reliability in effectively distinguishing between recommendations.

For the test set, the accuracy was 90.91%, showing a slight improvement over the cross-validation accuracy and confirming the model's high precision in identifying recommendations on a dedicated test split. This result aligns well with the original test accuracy, indicating a consistent performance level even when cross-validation is applied. The balanced accuracy on the test set was 90.75%, marginally higher than the cross-validation balanced accuracy, which shows that the model performs well across both classes, even on an unseen test set. The MCC on the test set was 0.82, which was slightly higher than the cross-validation proxy of the AUC (0.96), indicating a robust correlation between the predicted and actual classifications. An MCC close to 1 reflects a minimal random error in the model's performance, further supporting its reliability across classes.

The classification report shows that the precision and recall values were similar to those of previous results, with both classes achieving high values close to 0.90. Specifically, the precision for the "not recommend" class improved slightly to 0.91, whereas the recall for the "recommend" class remained consistent. This high precision and recall suggest that the model accurately captures true recommendations and non-recommendations, with minor variations likely due to the different training–testing splits. The confusion matrix shows that the model correctly classified 161 out of 174 instances as "not recommend" and 129 out of 145 instances as "recommend", with only minor misclassifications. This low misclassification rate aligns with previous results, confirming the model's precision and recall.

The AUC on the test set remains high at 0.96, indicating that the model consistently distinguishes between the "recommend" and "not recommend" classes with a high level of accuracy, as observed in both the cross-validation and test set results. This high AUC value suggests that the GBM model reliably identifies passenger recommendations on the basis of the provided service-related features.

In conclusion, the consistency between the cross-validation and test set results confirms that the gradient boosting machine is highly effective and generalisable well for this classification task. The slight variations between the cross-validation and test set metrics highlight that the model is stable across different data splits and remains unaffected by class imbalance. Overall, the addition of cross-validation reinforces the model's robustness, confirming that the GBM model is a reliable classifier for predicting airport service recommendations. Future experimentation with other model types or additional features could further enhance performance, but the current model configuration shows strong predictive power in this context.

Finally, this research aims to improve model performance by hyperparameter tuning and sets an experimental design as follows.

#### 3.5.1. The number of estimators

The number of estimators or the number of boosting rounds was set to 200, a choice that reflects a balance between underfitting and overfitting. The optimal number of estimators was determined through grid search cross-validation, with candidate values ranging from 100 to 200. A greater number of estimators often enables the model to learn more complex patterns, but this may increase overfitting risk. The selected value of 200 provided strong classification results without excessive model complexity.

### 3.5.2. Learning rate (shrinkage factor)

The learning rate controls the contribution of each tree to the overall model. In this experiment, a learning rate of 0.1 was selected after testing values ranging from 0.01 to 0.1. While reducing overfitting, lower learning rates often require a larger number of estimators for the model to converge effectively. A learning rate of 0.1 allowed faster convergence and good predictive accuracy without compromising the model's ability to generalise.

### 3.5.3. Maximum depth of trees

The maximum depth of the trees was set to five, allowing the model to capture complex patterns in the data. A deeper tree depth can model more intricate relationships but risks overfitting by learning specific details of the training set. Testing tree depths from three to five in the grid search revealed that a depth of five yielded the best balance, accurately capturing relationships without excessive complexity.

### 3.5.4. Subsampling rate and column sample rate

The subsampling rate was adjusted to 0.9, and the column sample rate was set to "log2" (a subset of features for each split). By using a subsampling rate of less than 1.0, the model introduces slight randomness to the training process, reducing variance and mitigating overfitting. The column sample rate ("log2") further enhances generalisation by limiting the number of features each tree evaluates, increasing diversity among trees. This strategy contributes to the model's robustness without sacrificing accuracy.

In the experiment, the best cross-validation accuracy from the grid search was 90.30%, demonstrating strong model performance. The cross-validation accuracy scores across folds were consistently high, with a mean of 89.22% and a balanced accuracy of 89.84%, indicating that the model performs well across the "recommended" and "not recommended" classes.

For the test set, the model achieved an accuracy of 89.03% and a balanced accuracy of 88.74%. The Matthews correlation coefficient (MCC) was 0.78, indicating a strong correlation between the predicted and actual classifications. The model's precision, recall, and F1 scores for each class were approximately 0.90, with minimal variations in the confusion matrix. Specifically, 160 out of 174 "not recommend" instances were correctly classified, and 124 out of 145 "recommend" instances were correctly classified. The test set AUC was 0.95, confirming the model's excellent ability to distinguish between recommendation categories, as shown in Fig. 7.

The thorough hyperparameter tuning of the GBM model confirmed its reliability for classifying passenger recommendations, with consistent results across the cross-validation and test sets. The selected hyperparameters balance complexity and generalisation, ensuring strong performance in predicting airport service recommendations. Future studies could explore additional regularisation techniques or alternative classifiers, but the current model setup demonstrates substantial predictive power for this dataset.

The experiments conducted with multiple machine learning models demonstrated consistent classification performance for predicting passenger recommendations regarding airport services. To obtain a clear result, this section aims to summarise the efficiency of each classification model after the model is trained via cross-validation and parameter tuning via an evaluation matrix, as shown in Table 3.

Table 3 shows that each model's performance, optimised through parameter tuning, was assessed via metrics such as accuracy, balanced accuracy, the Matthews correlation coefficient (MCC), and the area under the curve (AUC). The random forest and neural network emerged as the top-performing models, achieving the highest precision, recall, and F1 scores across both classes. Both demonstrated a balanced accuracy of approximately 91% and MCC values exceeding 0.80, indicating their effectiveness in distinguishing between the "recommend" and "not recommend" classes. Additionally, their high AUC score of 0.96 confirms their strong discriminatory power, further validating their reliability for classification tasks.

The decision tree and gradient boosting machine (GBM) models also performed well, although their balanced accuracy and F1 scores were slightly lower than those of the random forest and neural network models. The decision tree achieved a balanced accuracy of approximately 86%, with an MCC of 0.76, suggesting good performance, although with slightly less consistency in classification accuracy. The GBM yielded comparable metrics, with an AUC of 0.96, indicating that it is effective but slightly less consistent across different class distributions.

The support vector machine (SVM) displayed balanced accuracy, and MCC values close to 90% and 0.78, respectively, indicating stable and reliable per-

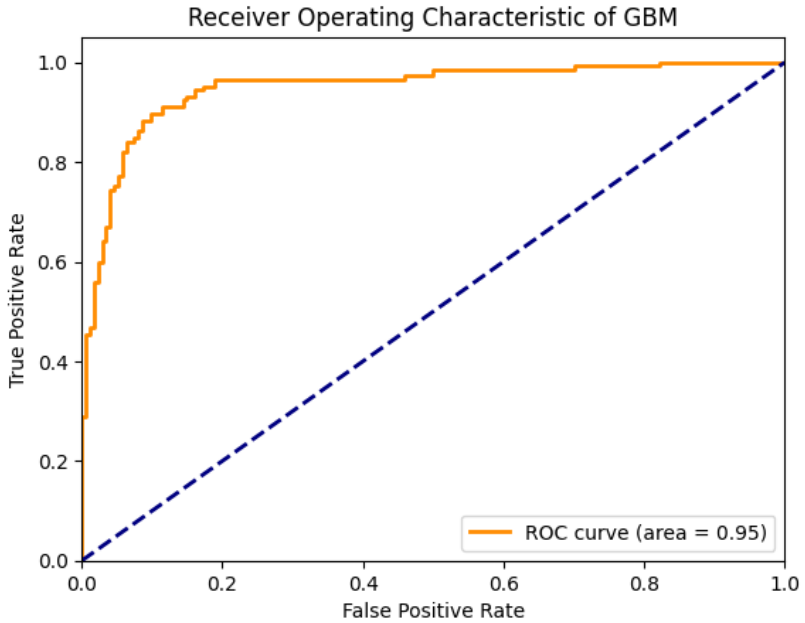



Fig. 7. Receiver operating characteristic (ROC) characteristics of GBM

Tab. 3. Evaluation matrices from all experimental models

| Models | Class Label | Classification Report with Optimised Model | | | | Accuracy | Balance Accuracy | MCC | AUC |
|---|---|---|---|---|---|---|---|---|---|
| | | Precision | Recall | F1-Score | Support | | | | |
| Decision tree | Not recommend | 0.86 | 0.89 | 0.88 | 174 | 0.86 | 0.86 | 0.76 | 0.93 |
| | Recommend | 0.87 | 0.83 | 0.85 | 145 | | | | |
| Random forest | Not recommend | 0.92 | 0.92 | 0.92 | 174 | 0.91 | 0.91 | 0.82 | 0.96 |
| | Recommend | 0.90 | 0.90 | 0.90 | 145 | | | | |
| SVM | Not recommend | 0.89 | 0.90 | 0.90 | 174 | 0.89 | 0.89 | 0.78 | 0.95 |
| | Recommend | 0.88 | 0.87 | 0.88 | 145 | | | | |
| Neural network | Not recommend | 0.92 | 0.91 | 0.92 | 117.4 | 0.91 | 0.91 | 0.83 | 0.90 |
| | Recommend | 0.90 | 0.91 | 0.90 | 94.8 | | | | |
| GBM | Not recommend | 0.89 | 0.90 | 0.90 | 174 | 0.89 | 0.88 | 0.77 | 0.96 |
| | Recommend | 0.88 | 0.86 | 0.87 | 145 | | | | |

Tab. 4. Confusion matrix of each classification model

| Models | | Predicted: No | Predicted: Yes |
|---|---|---|---|
| Decision tree | Actual: No | 156 | 18 |
| | Actual: Yes | 24 | 121 |
| Random forest | Actual: No | 161 | 13 |
| | Actual: Yes | 14 | 131 |
| SVM | Actual: No | 158 | 16 |
| | Actual: Yes | 18 | 127 |
| Neural network | Actual: No | 540 | 47 |
| | Actual: Yes | 41 | 433 |
| GBM | Actual: No | 158 | 16 |
| | Actual: Yes | 19 | 126 |

formance. However, the ability of SVM to distinguish between classes was marginally lower than that of random forest and neural networks, reflecting a slight limitation in its discriminatory power across class boundaries.

In addition, after the models are trained via cross-validation and the parameters are fine-tuned, a confusion matrix is analysed to provide a detailed breakdown of a classification model's performance by showing the counts of actual versus classifying outcomes across classes of each model, as shown in Table 4.

From Table 4, the confusion matrices reveal distinct prediction tendencies across the models. The random forest method demonstrated the fewest misclassifications, with only 13 false positives and 14 false negatives, underscoring its robustness in accurately distinguishing between recommendation classes with minimal errors. The neural network, while achieving high overall accuracy, had slightly higher misclassification rates, with 47 false positives and 41 false negatives, indicating some room for improvement in precise class separation despite its high AUC score. The decision tree, SVM, and GBM models also performed well but exhibited marginally higher misclassification rates, with GBM and SVM producing 16–18 false positives and 16–19 false negatives. Although these models are accurate, they are not as precise as the random forest in differentiating between classes.

The random forest model is the recommended choice for SHAP analysis because of its high accuracy, robustness, and interpretability. The random forest method demonstrated the fewest misclassifications, along with highly balanced accuracy, precision, and recall, indicating that a stable model effectively differentiates between the "recommend" and "not recommend" classes. This stability ensures that SHAP values provide a reliable interpretation of the importance of features. Additionally, the random forest's consistent performance across key metrics, including the MCC and AUC, reflects its reliability and strong discriminatory power, which enhances the trustworthiness of SHAP insights. Furthermore, random forest models are inherently interpretable, making them well suited for SHAP analysis, as SHAP values can clearly illustrate feature contributions across the ensemble of decision trees, providing a clear understanding of how each feature impacts predictions.

Fig. 8 shows that the SHAP value analysis of the random forest model reveals the relative importance of various features in predicting whether passengers will "recommend" airport services. The feature staff stands out as the most influential factor, indicating that the quality of service provided by airport staff significantly impacts passenger satisfaction and the likelihood of recommendation. This result suggests that passengers prioritise the quality of interactions with airport personnel, which can shape their overall experience and satisfaction.

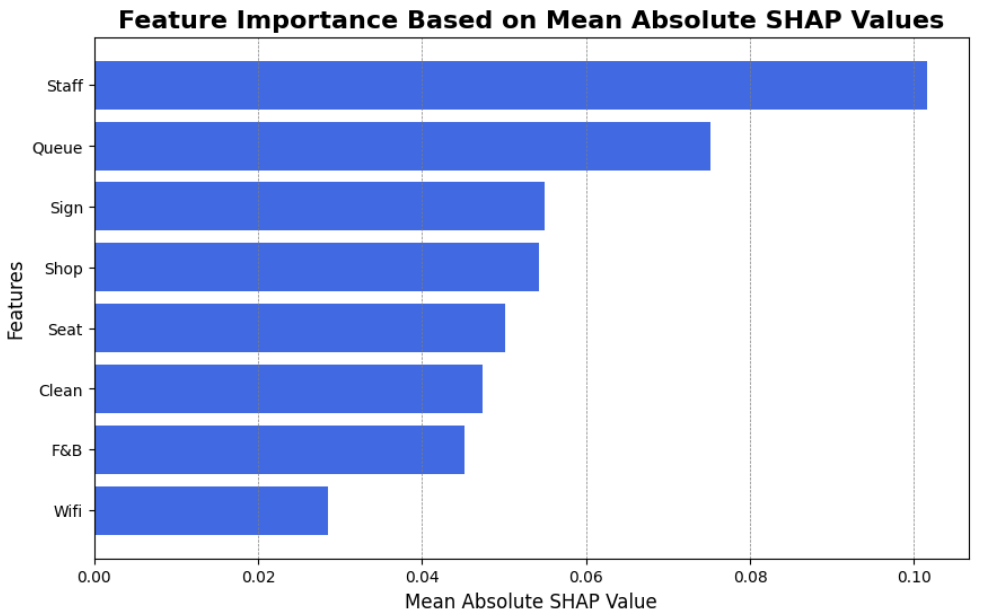



Fig. 8. Feature importance based on the mean absolute SHAP values

Queue management emerges as the second most important feature, underscoring that the efficiency of queue handling and wait times considerably influence passenger decisions. Passengers are likely to recommend airports if they experience minimal wait times and efficient service, highlighting the operational aspect of customer satisfaction.

Following queues, features such as signs and shops have comparable importance. The presence of clear signage contributes to ease of navigation within the airport, whereas the availability of shops adds convenience and variety to the passenger experience. Although these factors are not as influential as staff and queues are, they still play a notable role in enhancing passenger satisfaction.

"Seat" and "clean" exhibit moderate importance in influencing recommendations, suggesting that cleanliness and seating comfort are valued by passengers but are secondary to factors such as staff service and queue management. "F&B" (food and beverage), while still relevant, ranks lower in importance, indicating that passengers may not consider food services critical to their overall airport experience. "Wifi" has the lowest SHAP value, suggesting that internet connectivity, although expected by passengers, is not a primary factor driving satisfaction and recommendation.

In summary, the SHAP values indicate that staff and queue management are the most crucial features and are likely to increase passenger satisfaction and willingness to recommend airports. This insight implies that airports should focus on improving the quality of staff interactions and optimising queue management systems to reduce wait times. Secondary factors, such as signage, cleanliness, and seating, also contribute positively and should be maintained to create a well-rounded, pleasant experience. In terms of machine learning implications, this analysis demonstrates how SHAP values can offer a nuanced understanding of feature importance, enabling targeted improvements in service quality that align with customer priorities.

## 4. Discussion of the results

Among the passenger expectation changes and technology adoptions that could be used to enhance service quality, airports face many complaints from passengers. On the other hand, passengers suggest airport services to others. This research aims to alleviate the problem for airports regarding many variables that might affect passenger satisfaction. Moreover, airports cannot make certain decisions concerning service preparation that should exceed passenger expectations. Eventually, passengers might not recommend airports to others, such as family or friends, during their trip. These problems may vary across the different environments of each airport. To overcome these problems, a classification model was developed to match airports with services that meet passenger expectations and identify variables that might influence a passenger to recommend airports to others. Therefore, this research focuses on airports in Southeast Asia, which encompass many tourist destinations and support tourists from countries worldwide.

This study aims to analyse the factors influencing passenger intentions to recommend airport services in Southeast Asia. The research process began with data collection through data extraction from the Skytrax website using the R programming language. The resulting dataset comprises variables reflecting passenger satisfaction across various service dimensions, as well as their tendency to recommend the respective airports to others.

During the data preparation phase, the research team structured the data appropriately for processing with data mining techniques, splitting the dataset into two parts: 70% for building and training machine learning models (training dataset) and 30% for evaluating model performance (testing dataset).

The data analysis was conducted by applying five types of data mining models, including decision trees, random forests, support vector machines, neural networks, and gradient boosting machines. All the models were developed and tested using Python. The experimental process was divided into three rounds:

(1) Evaluating the performance of the five models via the basic dataset, (2) applying 5-fold cross-validation to enhance model learning performance, and (3) performing hyperparameter tuning to optimise model performance.

Model performance in each experimental round was evaluated using multiple metrics, including the confusion matrix, balanced accuracy, Matthew's correlation coefficient, and ROC curve. The comparison results indicated that the random forest model demonstrated the highest performance in classifying passenger tendencies to recommend airport services. This highlights the suitability of the random forest model for the regional context of airports in Southeast Asia. However, the results may vary if the model is applied to datasets from airports in other regions.

Additionally, the SHAP value analysis combined with the random forest model revealed that staff quality and queuing time were the most influential factors affecting passenger decisions to recommend airport services in Southeast Asia.

Overall, the findings not only present an effective model for predicting passenger behaviour but also provide deep insights into key factors influencing satisfaction and recommendation intentions in the context of airports in Southeast Asia. This research demonstrates the potential of applying data mining techniques and advanced analytics using Python to generate valuable knowledge for enhancing airport service quality and developing strategies to improve the passenger experience more effectively and in line with their needs.

## Discussion and theoretical implications

This research highlights the complex relationship between passenger satisfaction levels and the propensity to recommend airport services within the Southeast Asian context. By employing multiple data classification techniques — decision trees, random forests, support vector machines, neural networks, and gradient boosting machines — this study provides substantial empirical evidence linking various service touchpoints to passenger experiences.

The findings indicate that the random forest model stands out for its superior accuracy and interpretability and has emerged as the most effective tool for predicting passenger recommendations on the basis of a diverse set of service quality metrics. This model not only demonstrates its suitability for classification tasks but also offers a robust framework for analysing passenger satisfaction. The ability of random forests to accurately predict recommendations underscores the potential of machine learning as a valuable asset in the development of customer service strategies. Moreover, this research contributes significantly to the theoretical literature on airport service quality (ASQ) by integrating advanced machine learning methodologies with user-generated data sourced from platforms such as Skytrax. This approach not only enhances the understanding of key variables influencing passenger attitudes but also underscores the growing importance of data-driven decision-making in the aviation sector. The identification of critical factors such as staff service quality and queue management highlights key areas for targeted improvements in airport operations. The application of SHAP values for feature importance analysis offers practical insights, detailing which service components most significantly affect passenger satisfaction. Such granular detail is invaluable for airport authorities aiming to prioritise resource allocation and enhance service quality. Focusing efforts on improving staff interactions and optimising queue management can effectively address the most impactful aspects of the passenger experience, potentially leading to increased recommendation rates and a strengthened competitive advantage.

While this study explored multiple machine learning models, the superior performance of random forests compared with decision trees, SVMs, neural networks, and GBMs suggests a critical need for ongoing exploration of ensemble methods in similar contexts. Additionally, this research reinforces the notion that model selection should be guided by the specific characteristics of the dataset, emphasising the advantages of ensemble approaches when addressing complex classification challenges. By integrating machine learning techniques to analyse extensive and intricate datasets, this study enriches existing theories on service quality and customer satisfaction. The findings imply that traditional metrics of service quality can be effectively enhanced with data-driven insights, providing a more nuanced understanding of passenger expectations and satisfaction levels. The successful application of machine learning methods in this research highlights the vital role of technology in managing and advancing service quality. The ability to process and analyse user-generated content from platforms such as Skytrax opens new avenues for theoretical development regarding the utilisation of customer feedback in service enhancement. Furthermore, by identifying staff performance and queue management as pivotal determinants, this study provides empirical support for theories that stress the significance of personal interactions and operational efficiency within customer experience management. These insights not only align with but also expand upon prior research, suggesting that enhancing these aspects can lead to elevated levels of customer satisfaction and loyalty.

### 4.1. Managerial implications

From a managerial perspective, this study highlights the critical need for airport operators to prioritise improvements in both staff performance and queue management, as these factors are pivotal in

influencing passenger recommendations. This emphasis aligns with the findings of Akan and Karataş (2025), Brochado et al. (2024), Halpern and Mwesiumo (2021), and Pholsook et al. (2024), who also identify these factors as critical determinants of passenger satisfaction.

The analysis underscores the role of staff interactions as a primary determinant of passenger satisfaction, indicating that comprehensive training programmes focused on customer service skills are essential. Emphasising courteous and efficient staff behaviour is paramount, as prioritising human resource development can yield substantial improvements in passenger experiences and increase the likelihood of service recommendations. The findings suggest that targeted investments in staff training and the implementation of efficient operational practices can significantly increase customer satisfaction. By adopting a robust classification model, airport authorities are equipped to derive insights that inform strategic initiatives designed to improve the overall passenger experience. Such proactive measures are likely to foster increased passenger loyalty and ultimately contribute to enhanced nonaviation revenue streams. Research has consistently highlighted the pivotal role of airport employees in shaping passenger loyalty and satisfaction. Antwi et al. (2020) concluded that the helpfulness and communication skills of airport personnel significantly influence these factors. To support this view, Bakır et al. (2022) identified airport employees as the most critical determinants of passenger satisfaction. In a similar vein, Wattanacharoensil et al. (2017) noted that passengers often experience negative encounters at airports, primarily due to insufficient assistance from staff, including airline ground personnel, security officers, and immigration officials, as emphasised by Brochado et al. (2024). The contribution of staff to enhancing airport service quality is further underscored by Fodness and Murray (2007), who argued that well-trained employees, who are equipped with excellent service and communication skills, can increase airports' overall efficiency and productivity (Pantouvakis & Renzi, 2016). Passenger satisfaction within the airport context is influenced by a multitude of factors, with the quality of services rendered by staff being paramount. Creating a conducive service environment is essential for promoting customer satisfaction and fostering positive behavioural outcomes. Relógio and Tavares (2023) suggest that airport managers can bolster employee performance by refining their recruitment and training methodologies, providing incentives, and facilitating clear career progression paths. Additionally, the implementation of effective conflict resolution protocols, the establishment of enhanced feedback mechanisms, and the utilisation of observational techniques, such as mystery shopper assessments, are recommended to identify areas for improvement and increase service standards, as noted by Brochado et al. (2024).

Queuing time is also recognised as a critical factor in the overall airport experience, necessitating that managers map the customer journey to identify specific pain points associated with queues. To address these issues, the implementation of queue management systems is essential, along with continuous collaboration with airlines and security agencies, to optimise the flow of passengers (Brochado et al., 2024). Self-service check-in kiosks and automated security screening systems have the potential to significantly increase passenger satisfaction by streamlining processes. These technologies alleviate the need for traditional check-in counter queues, thereby saving time and providing a more efficient experience for travellers. Automated security screening systems, which incorporate advanced technologies such as millimetre-wave scanners and automated X-ray machines, further contribute to minimising waiting times and increasing overall convenience for passengers (Booranakittipinyo et al., 2024). In alignment with the Sustainable Development Report of the Airport of Thailand (AOT, 2022), initiatives such as the "Touchless Airport" concept have been introduced. This includes the deployment of common-use self-service (CUSS) kiosks and common-use bag drop (CUBD) technologies, both aimed at reducing touchpoints within the airport environment. Furthermore, encouraging early reservations for accommodations and restaurant services, along with promoting the use of digital payment systems, can significantly mitigate crowding and reduce queuing times for passengers.

While staff performance and queue efficiency are crucial, the significance of clear and effective signage should not be overlooked. An increase in airport wayfinding systems can alleviate passenger stress and improve the overall flow of individuals through airports, contributing to a more seamless travel experience. By understanding which specific features most influence passenger recommendations, airport managers can allocate resources more strategically. Investments should be concentrated in those areas that deliver the highest return in passenger satisfaction, such as staff quality and queue management systems, rather than being diffused across less impactful areas.

Integrating data analytics into service management practices allows for ongoing monitoring and continuous improvement of service quality. Managers should establish data-driven feedback loops, utilising insights generated from machine-learning models to inform strategic decision-making and adapt to evolving passenger needs and expectations. In addition to staff and queue management improvements, attention must also be given to other aspects of the passenger journey, such as the retail and dining experiences offered at the airport. Enhancing the variety and quality of shopping and dining options can add significant value to the passenger experience, making airports more attractive and recommendable destinations. Finally, encouraging real-time feedback from passengers is vital for identifying immediate areas for improvement and gauging satisfaction levels. The implementation of systems that facilitate swift passenger feedback can empower airports to be responsive and agile in addressing any service issues that may arise, thereby further increasing overall passenger satisfaction and loyalty.

### 4.2. Limitations and prospective research

While this study provides valuable insights into passenger satisfaction and service quality within Southeast Asian airports, it is essential to recognise its limitations. The dataset utilised in this research is confined to specific airports within Southeast Asia, which may restrict the generalisability of the findings to other regions or smaller airports. Additionally, the reliance on user-generated content may introduce biases related to self-selection, potentially skewing the overall satisfaction ratings and affecting the accuracy of the conclusions drawn.

Future research can address these limitations by expanding the dataset to encompass a broader range of airports globally. Such an expansion would enhance the robustness and applicability of the findings across different contexts. Moreover, longitudinal studies that monitor shifts in passenger expectations over time could provide critical insights into evolving preferences, particularly in the aftermath of the COVID-19 pandemic. Further exploration of advanced machine learning techniques, such as deep learning models, could yield richer insights into the factors influencing passenger satisfaction. Additionally, incorporating external variables — such as economic indicators and environmental conditions — may deepen the understanding of the complex dynamics affecting passenger experiences. Integrating natural language processing techniques into the analysis of passenger reviews could also prove beneficial, offering enhanced insight into the sentiments expressed by travellers. Moreover, examining alternative machine learning models and hybrid approaches may facilitate advancements in the accuracy of predictions related to passenger satisfaction.

## Conclusions

In conclusion, this study lays a solid foundation for continued exploration at the intersection of data analytics and airport service quality. By pursuing the recommended avenues for future research, the airport industry can refine service strategies that effectively resonate with passenger needs and expectations, thereby enhancing the overall traveller experience.

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This study was supported by Thammasat University Research Fund, Contract No TUFT0029/2567.

## Literature

Abouseada, A.-A., Hassan, T., Ibraheam Saleh, M., & Radwan, S. (2023). The power of airport branding in shaping tourist destination image: passenger commitment perspective. *GeoJournal of Tourism and Geosites, 47*, 440-449. doi: 10.30892/gtg.47210-1042

ACI. (2021a). *Airport Service Quality (ASQ)*. Retrieved from https://aci.aero/customer-experience-asq/

ACI. (2021b). *Airport Service Quality Customer Experience*. Retrieved from https://aci.aero/programs-and-services/asq/

ACI. (2022). *About ACI*. Retrieved from https://aci.aero/about-aci/

ACI. (2023). *Global passenger traffic expected to recover by 2024 and reach 9.4 billion passengers*. Retrieved from https://aci.aero/2023/09/27/global-passenger-traffic-expected-to-recover-by-2024-and-reach-9-4-billion-passengers/

Akan, S., & Karataş, H. (2025). Evaluating Service Quality of Europe's Leading Airports Using Skytrax Ratings After Istanbul Airport's Launch. In M. Bakır (Ed.), *Sustainable Marketing Practices in the Aviation Industry* (pp. 403-434). IGI Global. doi: 10.4018/979-8-3693-7215-9.ch012

Alanazi, M. S. M., Jenkins, K., & Li, J. (2024). Predicting passengers' feedback rate for airport service quality.

*Transportation Research Interdisciplinary Perspectives, 24*, 101046. doi: 10.1016/j.trip.2024.101046

Alanazi, M. S. M., Li, J., & Jenkins, K. W. (2024). Evaluating Airport Service Quality Based on the Statistical and Predictive Analysis of Skytrax Passenger Reviews. *Applied Sciences, 14*(20), 9472.

Antwi, C. O., Fan, C.-J., Nataliia, I., Aboagye, M. O., Xu, H., & Azamela, J. C. (2020). Do airport staff helpfulness and communication generate behavioral loyalty in transfer passengers? A conditional mediation analysis. *Journal of Retailing and Consumer Services, 54*, 102002. doi: 10.1016/j.jretconser.2019.102002

AOT. (2022). *Sustainable Development Report 2021*. Retrieved from https://www.airportthai.co.th/wp-content/uploads/2022/01/SDReport2021en.pdf

Arasli, H., Saydam, M. B., Jafari, K., & Arasli, F. (2023). Nordic Airports' service quality attributes: themes in online reviews. *Scandinavian Journal of Hospitality and Tourism*, 1-16. doi: 10.1080/15022250.2023.2259345

Bakır, M., Akan, Ş., Özdemir, E., Nguyen, P.-H., Tsai, J.-F., & Pham, H.-A. (2022). How to Achieve Passenger Satisfaction in the Airport? Findings from Regression Analysis and Necessary Condition Analysis Approaches through Online Airport Reviews. *Sustainability, 14*(4), 2151.

Barakat, H., Yeniterzi, R., & Martín-Domingo, L. (2021). Applying deep learning models to twitter data to detect airport service quality. *Journal of Air Transport Management, 91*, 102003.

Bezerra, G. C. L., & Gomes, C. F. (2019). Determinants of passenger loyalty in multi-airport regions: Implications for tourism destination. *Tourism Management Perspectives, 31*, 145-158. doi: 10.1016/j.tmp.2019.04.003

Bezerra, G. C. L., & Gomes, C. F. (2020). Antecedents and consequences of passenger satisfaction with the airport. *Journal of Air Transport Management, 83*, 101766. doi: 10.1016/j.jairtraman.2020.101766

Booranakittipinyo, A., Li, R. Y. M., & Phakdeephirot, N. (2024). Travelers' perception of smart airport facilities: An X (Twitter) sentiment analysis. *Journal of Air Transport Management, 118*, 102600. doi: 10.1016/j.jairtraman.2024.102600

Braikia, H., Hamida, S. B., & Rukoz, M. (2024). Random Forest Classifier for Marine Biodiversity Analysis. In International Conference on Intelligent Systems and Computer Vision (ISCV). Fez, Morocco.

Brochado, A., Veríssimo, J. M. C., & Lupu, C. (2024). Airport experience assessment based on Skytrax online ratings and importance-performance analysis: a segmentation approach. *Journal of Marketing Analytics*. https://doi.org/10.1057/s41270-024-00326-x

Bulatović, I., Dempere, J., & Papatheodorou, A. (2023). The explanatory power of the SKYTRAX's airport rating system: Implications for airport management. *Transport Economics and Management, 1*, 104-111. doi: 10.1016/j.team.2023.07.002

Bunchongchit, K., & Wattanacharoensil, W. (2021). Data Analytics of Skytrax's Airport Review and Ratings: Views of Airport Quality by Passengers Types. *Research in Transportation Business and Management*, *41*, 100688.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access, 9*, 78368-78381. doi: 10.1109/ACCESS.2021.3084050

Chinnarasri, C., Nonsawang, S., & Supharatid, S. (2012). Application of Artificial Neural Networks for River Stage Forecasting in Hatyai. *KMUTT Research and Development Journal, 26*(1), 125-137.

Del Chiappa, G., Martin, J. C., & Roman, C. (2016). Service quality of airports' food and beverage retailers. A fuzzy approach. *Journal of Air Transport Management, 53*, 105-113. doi: 10.1016/j.jairtraman.2016.02.002

Doo, F., Savani, D., Kanhere, A., Carlos, R., Joshi, A., Yi, P., & Parekh, V. (2024). Optimal Large Language Model Characteristics to Balance Accuracy and Energy Use for Sustainable Medical Applications. *Radiology, 312*, e240320. doi: 10.1148/radiol.240320

Elshafey, M., Rowlands, D., Contestabile, E., & Halim, A. (2007). Airport level of service perceptions before and after September 11: a neural network analysis. *WIT Transactions on The Built Environment, 94*, 337-345. doi: 10.2495/SAFE070331

Eswaran, P., Santhosh, M., Krishnan, A., & Kumar, T. (2019). Sentiment Analysis of US Airline Twitter Data using New Adaboost Approach. *International Journal of Engineering and Technical Research, 7*, 1-3.

Fodness, D., & Murray, B. (2007). Passengers' expectations of airport service quality. *Journal of Services Marketing, 21*(7), 492-506.

Freitas, P. T. C., Silva, L. M., Nascimento, M. V., & Borille, G. M. R. (2021). Passenger profile and its effects on satisfaction level in food and beverage establishments: Case study of major Brazilian airports. *Case Studies on Transport Policy, 9*(3), 1219-1224. doi: 10.1016/j.cstp.2021.06.009

Gitto, S., & Mancuso, P. (2017). Improving airport services using sentiment analysis of the websites. *Tourism Management Perspectives, 22*, 132-136. doi: 10.1016/j.tmp.2017.03.008

Graham, A. (2009). How important are commercial revenues to today's airports? *Journal of Air Transport Management, 15*, 106-111.

Halpern, N., & Mwesiumo, D. (2021). Airport service quality and passenger satisfaction: The impact of service failure on the likelihood of promoting an airport online. *Research in Transportation Business & Management*, 100667. doi: 10.1016/j.rtbm.2021.100667

Homaid, M. S., & Moulitsas, I. (2023). Measuring Airport Service Quality Using Machine Learning Algorithms. In *6th International Conference on Advances in Artificial Intelligence*. Birmingham, United Kingdom.

Isa, N. A. M., Ghaus, H., Hamid, N. A., & Tan, P.-L. (2020). Key drivers of passengers' overall satisfaction at klia2 terminal. *Journal of Air Transport Management, 87*, 101859.

Jiang, H., & Zhang, Y. (2016). An assessment of passenger experience at Melbourne Airport. *Journal of Air Transport Management, 54*, 88-92.

Kiliç, S., & Çadirci, T. O. (2022). An evaluation of airport service experience: An identification of service improvement opportunities based on topic modeling and sentiment analysis. *Research in Transportation Business & Management, 43*, 100744. doi: 10.1016/j.rtbm.2021.100744

Lee, K., & Yu, C. (2018). Assessment of airport service quality: A complementary approach to measure per-

ceived service quality based on Google reviews. *Journal of Air Transport Management, 71*, 28-44.

Li, L., Mao, Y., Wang, Y., & Ma, Z. (2022). How has airport service quality changed in the context of COVID-19: A data-driven crowdsourcing approach based on sentiment analysis. *Journal of Air Transport Management, 105*, 102298.

Lively, H., Valencia, A., & Shouse, R. (2024). Decision Trees Impact Learning: An Empirical Study Using Exchanges of Nonmonetary Assets. *Issues in Accounting Education, 39*, 1-24. doi: 10.2308/ISSUES-2021-024

Lubbe, B., Douglas, A., & Zambellis, J. (2011). An application of the airport service quality model in South Africa. *Journal of Air Transport Management, 17*(4), 224-227.

Mao, Q., Jiang, W., Sun, G., & Zeng, D. (2023). *XGBoost-Enhanced Prediction and Interpretation of Heart Disease Using SHAP Values*. In *4th International Conference on Computer, Big Data and Artificial Intelligence (ICCBD+ AI)* (pp. 738-742). Guiyang, China.

Martin-Domingo, L., Martín, J. C., & Mandsberg, G. (2019). Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). *Journal of Air Transport Management, 78*, 106-115.

Mizufune, K., & Katsumata, S. (2018, 17-20 Nov. 2018). Joint Classification Model of Topic and Polarity: Finding Satisfaction and Dissatisfaction Factors from Airport Service Review. In *IEEE International Conference on Data Mining Workshops (ICDMW)*. Abu Dhabi, United Arab Emirates.

Naveen, S., Upamanyu, M. S., Chakki, K., C, M., & Hariprasad, P. (2023). Air Quality Prediction Based on Decision Tree Using Machine Learning. In *International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*. Tumakuru, India.

Nghiêm-Phú, B., & Suter, J. R. (2018). Airport image: An exploratory study of McCarran International Airport. *Journal of Air Transport Management, 67*, 72-84.

OAG. (2024). *Southeast Asia aviation market*. Retrieved from https://www.oag.com/south-east-asia-aviation-flight-data

Pantouvakis, A., & Renzi, M. F. (2016). Exploring different nationality perceptions of airport service quality. *Journal of Air Transport Management, 52*, 90-98.

Pholsook, T., Wipulanusat, W., & Ratanavaraha, V. (2024). A Hybrid MRA-BN-NN Approach for Analyzing Airport Service Based on User-Generated Contents. *Sustainability, 16*, 1164. doi: 10.3390/su16031164

Pholsook, T., Wipulanusat, W., Thamsatitdej, P., Ramjan, S., Sunkpho, J., & Ratanavaraha, V. (2023). A Three-Stage Hybrid SEM-BN-ANN Approach for Analyzing Airport Service Quality. *Sustainability, 15*(11), 8885.

Punne, M., Indrabayu, & Nurtanio, I. (2024). *Mood Classification from Song Lyrics Using the Naive Bayes Algorithm, Support Vector Machine (SVM) and XGBoost*. In *2024 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)* (pp. 162-167). Bali, Indonesia.

Rane, A., & Kumar, A. (2018). *Sentiment Classification System of Twitter Data for US Airline Service Analysis*. In *IEEE 42nd annual computer software and applications conference (COMPSAC)* (pp. 769-773). Tokyo, Japan.

Relógio, A. T., & Tavares, F. O. (2023). An Evaluation of Passenger Satisfaction among Users of Huambo Airport in Angola. *Urban Science, 7*(2), 57.

Sezgen, E., Mason, K. J., & Mayer, R. (2019). Voice of airline passenger: A text mining approach to understand customer satisfaction. *Journal of Air Transport Management, 77*, 65-74. doi: 10.1016/j.jairtraman.2019.04.001

Skytrax. (2024). *Explaining Airport Star Rating levels*. Retrieved from https://skytraxratings.com/explaining-airport-star-ratings-levels

Song, C., Guo, J., & Zhuang, J. (2020). Analyzing passengers' emotions following flight delays- a 2011–2019 case study on SKYTRAX comments. *Journal of Air Transport Management, 89*, 101903. doi: 10.1016/j.jairtraman.2020.101903

Sulu, D., Arasli, H., & Saydam, M. B. (2022). Air-Travelers' Perceptions of Service Quality during the COVID-19 Pandemic: Evidence from Tripadvisor Sites. *Sustainability, 14*(1), 435.

Touzani, S., Granderson, J., & Fernandes, S. (2018). Gradient boosting machine for modeling the energy consumption of commercial buildings. *Energy and Buildings, 158*, 1533-1543. doi: 10.1016/j.enbuild.2017.11.039

Triayudi, A., & Kamelia, L. (2022). Application of Certainty Factor Method to Diagnose Venereal Diseases Using Confusion Matrix for Multi-Class Classification. In 1*6th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*. Bali, Indonesia.

Wattanacharoensil, W., Schuckert, M., & Graham, A. (2015). An Airport Experience Framework from a Tourism Perspective. *Transport Reviews*, *36*(3), 318-340.

Wattanacharoensil, W., Schuckert, M., Graham, A., & Dean, A. (2017). An analysis of the airport experience from an air traveler perspective. *Journal of Hospitality and Tourism Management, 32*, 124-135.

World Bank Group. (2024). *Air transport, passengers carried*. Retrieved from https://data.worldbank.org/indicator/IS.AIR.PSGR

Yavuz, N., Olgaç, S., Günay Aktaş, S., & Mert Kantar, Y. (2020). Passenger Satisfaction in European Airports. In I. O. Coşkun, N. Othman, M. Aslam & A. Lew (Ed.), *Travel and Tourism: Sustainability, Economics, and Management Issues: Proceedings of the Tourism Outlook Conferences* (pp. 223-237). Sri Lanka: Springer.

Yeh, C.-H., & Kuo, Y.-L. (2003). Evaluating passenger services of Asia-Pacific international airports. *Transportation Research Part E: Logistics and Transportation Review, 39*(1), 35-48.

Zaharia, S. E., & Pietreanu, C. V. (2018). Challenges in airport digital transformation. *Transportation Research Procedia, 35*, 90-99. doi: 10.1016/j.trpro.2018.12.016

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 1 February 2025
accepted: 1 June 2025

pages: 63-77



# Impact of global supply chain complexity on the individual performance of companies

Katarzyna Grondys
Michał Kot

ABSTRACT
Cooperation is a key element in building business relationships, particularly when these relationships are complex and implemented on an international scale. The process of globalisation has facilitated the creation of complex networks and the expansion into new markets. However, globalisation has also introduced vulnerabilities due to global crises that disrupt supply chains. In this context, this paper examines the role of global supply chain complexity, specifically focusing on the number of global partners and its impact on the performance outcomes of individual supply chain actors. A multivariate analysis of variance (MANOVA) was employed to test the significance of differences between the studied groups in the independent sample. This approach allows for the examination of how variations in the number of global partners influence multiple performance indicators within the supply chain context. The analysis revealed that a greater number of global partners positively affects several key performance outcomes for companies. Specifically, improvements were observed in lead time, overall supply chain efficiency, inventory and warehouse management, company revenue, and market share. These findings highlight the continued benefits of globalisation despite concerns about its associated risks. This study contributes to the ongoing debate about the advantages and risks of global supply chain complexity. It provides empirical evidence that, while globalisation introduces risks such as instability and reduced resilience, the benefits associated with a greater number of global partners remain significant. This research enhances the understanding of how global connectivity impacts supply chain performance and strategic decision-making for international businesses. The findings suggest that businesses should carefully assess the balance between expanding global networks and managing potential disruptions. Firms can capitalise on the advantages of global partnerships by enhancing key performance outcomes while simultaneously developing strategies to mitigate risks associated with instability. Furthermore, discussions surrounding the regionalisation and shortening of supply chains should consider how strategic global partnerships can complement efforts to build more resilient supply chains.

KEY WORDS
**global supply chain complexity, globalisation, supply chain performance, multivariate analysis (MANOVA)**

**10.2478/emj-2025-0012**

**Katarzyna Grondys**
Czestochowa University of Technology,
Armii Krajowej 19B,
42-201, Częstochowa, Poland
ORCID 0000-0001-6461-5475
Corresponding author:
email: katarzyna.grondys@pcz.pl

**Michał Kot**
Czestochowa University of Technology,
Armii Krajowej 19B,
42-201, Częstochowa, Poland
ORCID 0000-0002-5449-593X
e-mail: michal.kot@pcz.pl

## Introduction

Cooperation is a major key to success in business relationships. Each enterprise is an inseparable part of some supply chain, forced to cooperate at different levels and with different markets. Thus, the results achieved in the entire supply chain are the synergy of functioning individual links. Thus, partner selection affects the results produced in a particular entity. The number of direct and indirect suppliers or the industry can determine the results achieved at each level of cooperation.

Grondys, K., & Kot, M. (2025). Impact of global supply chain complexity on the individual performance of companies. *Engineering Management in Production and Services, 17*(2), 63-77. doi: 10.2478/emj-2025-0012

The widespread phenomenon of globalisation, which allows for creating a dense and wide network of cooperation between entities across the globe and the resulting increase in competitiveness and opportunities to conquer new markets, has contributed to new problems and constraints during the extreme global crisis. Recent global and international economic crises have revealed that the higher the complexity of the supply chain, the greater the impediments and risk of lowering the efficiency of the entire chain (Ivanov, 2021). Extensive global and international collaborations involving numerous players have been shown to cause exponential delays in producing goods and services. The interdependence of deliveries between supplies meant that each successive delay at one level caused even greater delays at the next. Ultimately, consumers received their products with significant delays.

All this has led to increasing talk of the need to shorten supply chains or regionalise them to reduce the risk of a decline in their performance, especially in a dynamically changing economic environment, as well as to increase their resilience to unusual environmental disruptions (Bielecki et al., 2024; Gereffi, 2020).

It turns out that in severe crises, simple local supply chains may be more likely to remain stable than global ones. This can lead to deglobalisation or relocation of production, yet goods will still have to change country several times as part of the production process before reaching the end customer (Cernat, 2022). Considering the role of the supply chain in the effects of its individual links, it is necessary to analyse the process of cooperation in supply chains conditioned by its complexity. Most of the research presented later in this article focuses on identifying challenges related to longer delivery times, the risk of disruptions, and the overall effects of the pandemic. However, there is a lack of in-depth, quantitative empirical research that directly analyses the impact of the number of links in the supply chain on the measurable economic effects of companies. This research gap is the subject of this article. To this end, a preliminary study was conducted to see how the number of partners in the supply chain affects the performance of individual entities in different areas of activities. The article consists of the following parts: the literature analysis part is designed to present arguments on the relationship between supply chain performance and its complexity. The empirical part presents the author's survey-based research conducted. To this end, it includes the characteristics of the research sample, the research methodology, the results of statistical analyses, their discussion, and a summary.

# 1. Literature review

## 1.1. Global supply chain management

Supply chain management aimed at increasing efficiency determines gaining and maintaining a competitive advantage more often than the product itself. Especially today, when advances in technology and civilization make it possible to realise services and deliveries on a global scale, the degree of interdependence between links and the processes implemented gain in value. Collaboration between companies in the supply chain brings numerous benefits, i.e., realisation of common market goals, effective formation and use of the supply chain structure, increasing global profitability, more effective response to demand, creation of potential beyond the capabilities of the company itself or reduction of inventory levels along the chain (Tsan-Ming Choi & Cheng, 2011).

In this context, globalisation is extremely important as it has significantly contributed to shaping the approach to supply chain management through the increasing opportunities to create connections between markets and the free flow of goods and information (Gereffi & Fernandez-Stark, 2016). At the same time, it has also posed new challenges especially for companies operating in the global market, especially in relation to advances in technology and communication. Ali, 2023 pays particular attention to the following benefits of globalisation in supply chain management:

- Greater market access. Globalisation has allowed increased access to new customers from different world regions, resulting in independence from one market and increased opportunities for sales and revenue sources.
- Reduction of production costs. As a result of globalisation, companies benefit from relocating production to cheaper regions where labour or input costs are lower, thus reducing and optimising operating costs and spreading fixed costs over a higher production level.
- Access to scarce resources and skills. Companies gain access to resources or skills unavailable in the domestic market, which positively influences business development and its innovation level.
- Business flexibility and scalability. As a result of global interconnectedness, companies gain the

ability to adjust the size of their production and distribution directions to fluctuations in demand and the type of needs and thus respond appropriately to changes in the environment.
- Building supply chain resilience. Geographic diversification of supply chains helps reduce risks from environmental, political or economic factors, minimising the impact of potential disruptions from the environment.
- Access to modern technologies. Globalisation allows the exchange of experience from different areas of the world, which in the supply chain increases innovation and efficiency of individual links;
- Use of outsourcing. Access to diverse markets expands access to specialised service providers in different regions of the world to carry out side activities, allowing the company to focus on its main objectives.
- Cooperation and knowledge transfer. Establishing closer and farther afield partnerships facilitates the exchange of knowledge and experience, provided there is well-developed communication and information flow.
- Minimising lead times. Efficient supply chain processes and the use of modern tools and methods to optimise logistics processes help maintain a competitive position.

However, as research shows (Chopra & Meindl, 2016), the number of links in global supply chains significantly affects their complexity and efficiency. A larger number of links can increase coordination costs and the risk of delays, which negatively affects the operating results of individual companies. In this context, this study analyses how the number of links in a supply chain affects the efficiency and profitability of companies.

However, it is increasingly observed that increasing economic freedom can also be a threat (Baldwin, 2016). The uneven distribution of benefits achieved from globalisation in different countries around the world undermines the perception of free trade only in terms of development. Moreover, according to Erixon, "globalisation has stalled — and global trade has become a casualty of increasing protectionism around the world" (Erixon, 2018). This is also confirmed by Marel, stating that globalisation is in increasing trouble and is entering a period of rapid deglobalisation, as evidenced by the slowdown in global trade in recent years (van der Mare, 2020).

Shortages such as delayed delivery of products to customers or the inability to meet consumer needs in the supply chain caused by the recent global crisis have mainly introduced numerous disruptions in the flow of goods (Laire, 2021). An adverse supply chain event in the last year was experienced by more than 75% of large global organisations (Bondar et al., 2023). Cernat (2022) also points out problems with various trade arrangements for processing activities, which hinder the integration of developing countries' global production chains. These regulations typically include a single perspective on a country's processing trade conditions and requirements, significantly limiting the chances for interoperability in complex global supply chains. These barriers, as indicated by UNCTAD (2023) research, increase the complexity of supply chains by adding additional links, which can negatively affect companies' financial performance, especially in developing countries. This forces business leaders to immediately adapt their processes to those of the supply chain and seek new partners to reduce and improve the efficiency of the entity's operations. Therefore, assessing how the number of links in the supply chain affects the performance of individual companies forms the basis of this study, which examines this issue in the context of global and regional supply chains through the analysis of survey responses.

### 1.2. Impact of the number of links in the supply chain on company performance

The selection of partners and the proper maintenance of relationships between them is essential to remain competitive in the global market. As the market has expanded, the possibility of expanding one's network has increased. The number of closer and farther partners directly involved in a given company's activities can have different effects on the efficiency and quality of supply chain operations. This can be observed in the long-term effects of globalisation, which has created opportunities to define new business models involving foreign capital, innovation and the field of research and development with greater intensity. Greater production capacity has determined a significant increase in specialised human capital (Erixon, 2018). Moreover, building global economic and technological connections is essential in maintaining a country's competitiveness and, thus, economic development. Therefore, they must promote sources of international cooperation and specialisation in value chains and actively make joint efforts with their partners (Bauer et al., 2023). There is also increasing talk that globalisation is shifting toward intangible and digital flows (van der Mare, 2020). This

should be seen as complementary to, not opposed to, trade and the role of the global supply chain in that exchange (Pepper et al., 2016).

Globalisation also contributes to changes in the structure of supply chains and affects the local community. The need to respond effectively to customer needs forces the realisation of production in different parts of the world, which, on the one hand, simplifies the manufacturing of finished products and, on the other hand, makes the process of sourcing raw materials and intermediate products more difficult (Bujak, 2012). At the same time, the complexity and level of sophistication of the supply chain complicate the flow of goods and information between its partners. An observed effect of globalisation is the lengthening of lead times. It causes global supply chain management to run the risk of delayed and even incomplete deliveries. Moreover, the lengthening of the supply chain also increases the risk of currency fluctuations and tends to lower the level of vigilance against unforeseen events and phenomena occurring in various economies around the world (Muzyczka, 2011). At the same time, it creates entirely new opportunities and possibilities for companies willing and able to adapt to the changing environment. Global supply chain management is thus the resultant subordination to disparate domestic and international factors. Global supply chains and the values of multinational companies are paving the way for a global market for trade and the fragmentation of chains (Bauer et al., 2016).

Despite its many benefits, globalisation has posed several challenges for supply chain management. These are primarily due to recent disruptions in the stability of international supply chains caused mainly by recent economic crises and the associated consequences, i.e., pandemics, natural disasters, operational risk, financial and geopolitical crises, cyber insecurity and logistical disruptions (Thuge Lund et al., 2023). This is an interesting subject area of much current research, which mainly refers to supply chain resilience to external factors and building strong relationships with specific suppliers.

The negative impact of recent disruptions could especially be observed in the automotive industry, where waiting times in the end market average between six months and a year (Belhadi, 2021). This is especially true for technologically advanced products, which, due to their equipment, involve longer production process times and the use of modern materials and products. Great concern could also be observed in the semiconductor industry, which, due to its global nature, is exposed to numerous supply chain disruptions (SIA, 2022). The semiconductor chain involves many stages, i.e., research and development, design, manufacturing and distribution. In doing so, it requires many different international stakeholders. A large semiconductor company may require the participation of as many as 16,000 suppliers (Bauer, 2023). This significantly increases the complexity of supply chain management, which, as shown in a McKinsey report (2020), leads to increased coordination costs and delivery delays. In contrast, the intensive cooperation with China over the past decade, due to lower labour costs and specialisation of competencies in different countries, has caused supply chains in multinational companies to become rather fragmented (Erixon & Messerlin, 2009).

Numerous global connections mean greater complexity and associated problems requiring a broader view of one's operations. This, in turn, brings a greater risk of supply chain disruptions related to the dispersion of suppliers and customers, global crises or the individual circumstances of doing business in a particular country (especially when different legal systems, traditions, and religions are considered). The global market brings new opportunities and new competitors who have access to the same suppliers and customers. In addition, the dispersion of the supply chain makes it difficult to collect and control data, including sensitive data (Champion, 2023).

### 1.3. Consequences of the number of links in the supply chain for individual companies

One of the important problems of the global supply chain is delivery delays or long transit times, which can be particularly acute in complex chains and networks with many links. Research by Sheffi (2015) indicates that each additional level in the supply chain increases the risk of delays by approximately 5–10%, directly affecting customer satisfaction and companies' financial results. Effective inventory management and establishing partnerships only with experienced suppliers of products and logistics services can somewhat minimise this problem (Vidrova, 2020). This is perfectly evident from observations of the 2008–2009 and 2019–2021 global crises, which caused significant disruptions in demand and supply sequentially in many industries, thus revealing the previously mentioned disadvantages of an elongated supply chain. Research by Christopher (2016) indicates that extended supply chains increase operational risk, negatively affecting the profitability of companies,

especially in industries with high demand volatility. To cope with the changing conditions and maintain efficiency, it is important to maintain the constant ability of supply chain links to adapt to new conditions.

The number of links in the supply chain directly impacts the operational efficiency and financial performance of individual companies. This study focuses on examining how the number of links in the supply chain affects key aspects of a company's operations, such as operating costs, order fulfilment time, resilience to disruptions, and profitability. In the context of company performance, the number of links in the supply chain is crucial.

Literature reviews indicate that a larger number of links in the supply chain increases the risk of disruptions and coordination costs. For example, Jasrotia et al. (2024) emphasise that each additional level in the supply chain increases management costs by approximately 3–5%, which can lead to a reduction in operating margins. In addition, companies with more extensive supply chains are more vulnerable to fluctuations in demand and supply, negatively affecting their financial liquidity (Alicke et al., 2024).

What is more, the number of links affects companies' ability to respond quickly to market changes. Research by Christopher and Peck (2004) shows that shorter supply chains with fewer links allow companies to adapt more quickly to changing conditions, which is crucial in highly volatile industries such as electronics and automotive. According to a Deloitte report (2022), companies with shorter supply chains (fewer links) report 15–20% lower operating costs and faster order fulfilment times compared to companies with more extensive chains. A larger number of links in the supply chain increases the risk of these disruptions, as each additional link introduces potential points of failure (Christopher, 2016). Research indicates that companies with more extensive supply chains report higher operating costs and longer order fulfilment times, negatively affecting their profitability (Mentzer et al., 2001).

### 1.4. Globalisation and regionalisation of the supply chain

Regionalisation of supply chains, which is understood as reducing the number of links and shortening their length, is increasingly seen as a strategy for improving business efficiency in response to the challenges of globalisation. In the literature and ongoing research, one can find different positions on the direction of globalisation and related cooperation in international supply chains. On the one hand, there is the approach that globalisation is less extensive than assumed. DHL's 2022 report shows interesting results regarding globalisation (Altman & Bastian, 2022). Most trade operations are more often domestic than international. Only about 20% of global economic output is exported (in value-added terms), and foreign investment flows account for only 6% of gross fixed capital formation. Moreover, international flows are already largely regionalised, with about half of the international trade and capital realised in the world's major regions. For example, in the ASEAN region, shorter supply chains with fewer links increase the operational efficiency of companies through lower transportation costs and faster order fulfilment times (ADB, 2023).

This means that if physical distances and differences between countries did not constrain flows, they would have a far wider scope. The political environment has also become less conducive to globalisation, which increases the risk of costly fragmentation of international flows or its decline. On the other hand, the value of post-pandemic international trade shows that it is quite resistant to global pandemic crises. During the COVID-19 pandemic, trade flows were carried out over greater distances, thus contradicting the ongoing phenomenon of regionalisation (Mena et al., 2022). Instead, the president of the European Central Bank mentions three key changes in post-pandemic world trade, among which he marks the apparent shift from globalisation to regionalisation (Lagarde, 2022). A key argument supporting the view that the pattern of flows may become more regionalised stems primarily from a change in the attitude of companies and governments, which are increasingly willing to support nearshoring and reshoring (European Parliament, 2021). The business model based on doing business with long-standing regional partners is also a major advantage of regionalisation. This is primarily due to the reduction of transportation and transit costs, the elimination of impediments caused by regulations and laws of other countries, or the reduction of the phenomenon of dependence on other countries (Altman & Bastian, 2022). Regionalisation is fostered by high levels of inflation and shortages of inputs, causing companies to seek local suppliers to ensure timely and reliable deliveries (Grazia et al., 2021). In turn, due to limited public budgets, only the most strategic supply chains can count on assistance in relocating production sites (Zhang et al., 2018). Banaszyk (2023) noted that the model of full globali-

sation would coexist with the model of a multilateral structure with a regional character centred around the main consumption and production centres.

Simplifying and shortening supply chains may be one of the future directions of change in supply chain performance management. According to Deloitte's Supply Chain Panel survey, more than 70% of entities "either have or will implement dual or multi-sourcing, coming from single sourcing" (Thuge Lund & Visgaard Duch, 2023). The speed of response to customer demand generates the need to produce close to the potential consumer.

A significant argument towards regionalising parts of supply chains is lengthening lead times due to globalisation. Regionalisation reduces the number of links and shortens delivery cycles, which increases supply chain resilience and improves companies' operating results (Sheffi, 2015). Regionalisation is considered within supply chain flexibility as an increase in the use of local suppliers and the associated reduction in cycle times, which is one factor in supply chain resilience (Thuge Lund & Visgaard Duch, 2023). Moreover, some companies are changing their operations to mitigate emerging efficiency problems by increasing procurement from local suppliers (Central Europe CFO Survey, 2023). Copacino and Byrnes observed that manufacturers are currently reducing the number of their suppliers, allowing them to capture a large market share. A McKinsey study, on the other hand, found that 45% of supply chain actors either do not have access to downstream performance in their supply chain or only see suppliers adjacent to the supply chain and in a direct relationship (Alicke et al., 2022).

The lack of visibility into downstream activities, resulting from many cells, often causes disruption and consequent inefficiency in supply chain management. This is borne out by the research of Binh An Thi Duong et al., which proved that the resonant influence mechanism in the supply chain increases the effectiveness of downstream activities and thus significantly affects the performance of individual entities in the chain (Duong et al., 2019). Meanwhile, it is observed that about half of the companies fail within the next five years of their operations; this is especially true for small entities (Zhou, 2024). The reasons for this situation can be traced to poor supply chain performance, which can contribute to business failure. According to a 2014 Deloitte survey, only 8% of companies with weaker supply chains experience above-average growth. In contrast, nearly 80% of companies that achieve above-average revenue growth in their industries operate in high-performance supply (Chew & Mohamed Zainal, 2024). Previous studies have focused mainly on general performance indicators, but the literature lacks more detailed analyses considering different categories of partner numbers. Therefore, this study aims to fill this gap and provide more detailed information on this topic. These observations are the starting point for this study, which aims to examine how the number of links affects the performance of companies in this chain.

## 2. Research methods

### 2.1. Research sample

This article aims to conduct a quantitative analysis of the impact of the number of links in the supply chain on the efficiency and profitability of companies operating under conditions of global disruption. The study is based on empirical data and aims to verify whether the number of entities in the chain translates into the performance of the companies that belong to it. Further empirical analysis primarily used data collected through a survey. The survey finally included 212 companies from four selected European countries: Poland (19%), Romania (33%), the United Kingdom (22%), and Turkey (26%). The selection of companies for the study was simple and random. To ensure diversity in the supply chain partnership, the research population included companies of various sizes cooperating with their suppliers or customers at the international level. In the end, about 40% of small, 26% of medium-sized, and 34% of large companies took part in the study. The study excluded micro-companies, which usually operate locally.

Countries selected for the study are economically interconnected in terms of trade, foreign investment and membership in international economic organisations. First, they have numerous bilateral agreements on economic cooperation, investment, avoidance of double taxation and investment protection, which is conducive to building mutual cooperation in the process of exchange of goods, including on a global scale. In addition, these countries, due to their numerous memberships in international organisations, have established strategic partnerships with many countries on a global scale. Therefore, the results will serve as a prelude to research on supply chain changes as a result of the global crisis.

Preliminary characterisation of the surveyed companies made it possible to observe that the most

frequent participants were those whose number of global supply chain partners was between one and three (nearly 37%), with one in three companies declaring that they work with three to ten global partners (nearly 30%) or with more than ten (nearly 30%). In the case of British companies, four to nine partners dominated (29%), while the range for the other countries surveyed was between one and three global partners (Poland with 21%, Romania with 36%, and Turkey with 31%). At the same time, Polish and Turkish companies most often declared the absence of a global partner in their direct supply chain (both 40%). Typically, the surveyed companies acted as manufacturers (33%) or logistics providers (28%) in the supply chain. The most common participants in the survey were companies that act as manufacturers (33%) or logistics providers (28%) in the supply chain.

### 2.2. Assumptions and research methods

The collected data were statistically analysed according to the assumptions of the selected statistical methods. It was assumed that there is a relationship between the performance of entities in the supply chain and the number of partners in the supply chain. Therefore, the following research hypotheses were set.

Hypothesis 0. The entity's average performance rating is similar for each category of the number of partners in the supply chain

Hypothesis 1. The entity's average performance rating varies by the category of the number of supply chain partners in at least one of the company's performance areas studied

To verify the research hypothesis, the dependent and independent variables were defined based on a review of literature and research on the impact of the supply chain on the performance of companies participating in it.

1. The multivariate dependent variable determines the evaluation of supply chain performance, i.e.:

- lead time,
- efficiency of the entire supply chain,
- logistics costs,
- efficiency in inventory, warehouse management,
- human resource management,
- fulfilment of customer requests,
- mitigation of risk in the supply chain,
- revenue of the company,
- market share — the company's share in the whole market in which it operates.

The scale is 1 to 5, where 1 is definitely worse than the previous years, 2 is worse than the previous year, 3 is comparable to the previous years, 4 is better than the previous years, 5 is definitely better than the previous year.

2. The multivariate independent variable specifies the number of global logistics partners included in three groups: Group 1 — from 0 to 3 partners; Group 2 — from 4 to 9 partners; and Group 3 — 10 or more partners.

The defined variables were subjected to statistical analysis using the multivariate analysis of variance MANOVA method for independent samples, which made it possible to examine simultaneously the associations of multiple independent variables of the qualitative type and multiple dependent variables of the quantitative type (Wątroba, 2009). The most liberal NIR post-hoc test was used to test for differences (Aranowska & Rytel, 2010).

## 3. Research results

### 3.1. Checking assumptions for MANOVA

According to the assumptions of MANOVA analysis, all explained variables must have a multivariate normal distribution in the population, and all independent populations must be determined by the levels of factors and their intersections (Todorov & Filzmoser, 2010). To check the validity of the assumption of a normal distribution, random deviations were made based on a histogram of the number of residuals (Chmielewski & Berczyński, 2001). The analysis assumes that the random deviations are independent, have zero expected value, have constant variance, and follow a normal distribution. If the proposed dependency model is adequate, then the residuals that are realisations of the random deviations should show trends that confirm the assumptions about the normality of the distribution. Moreover, when the marginal distributions of the explanatory variables can be considered normal, the multivariate distribution of these variables is often also a normal distribution (Gnanadesikan, 1977). Outlier observations were examined using the Mahalanobis distance measure for the distance of a case from the centre of gravity, determined by the independent variables (Ampula, 2015). It was determined that the value of the Mahalanobis distance was greater than the critical value of the chi test2 for each p level and for N=212. Using exploratory statistics, five extreme values were

identified, i.e., so-called outlier observations, which disturbed the achievement of a multivariate normal distribution of the studied population. Again, repeated analysis of the Mahalanobis distance on the reduced dataset yielded a maximum distance value of 24.679. For N=207, the critical value of the distribution Chi2 = 27.88 > the value of the Mahalanobis distance, for p =0.001 and df = 9 allows for assuming that the condition about the multivariate normality of the distribution in the population was met.

### 3.2. Testing for homogeneity of covariance

The basic assumption for the use of the MANOVA model is the assumption of equality of the covariance matrix in the compared samples verified by the Box test. This means that each explanatory variable has homogeneous variances in all independent subpopulations determined by factor levels and their intersections. The Box test results for all explanatory variables combined did not meet the assumption of equality of covariance, which may be due to the number of dependent variables. The more of them, the greater the probability of inequality of variances across groups (Tabachnick & Fidell, 2001). Since the equality and the size of individual subgroups can affect the results of testing, their distribution was examined for certainty (Table 1).

The number of observations for the study groups is similar but not identical, so the equality of the covariance matrix was checked using the Box M test, the results of which are shown in Table 2.

Box's M test value (F(90; 251)) < 0.001 theoretically does not indicate homogeneity of the covariance matrix across groups. Test results less than 0.001 suggest possible bias in the results for the alpha level. However, according to the approach of Tabachnick and Fidell (2001), the condition of analysis can be considered satisfied if the sample sizes are not equal. In this case, the sample sizes are not equal, so the result was considered significant, allowing the MANOVA model to be applied to the new dataset.

### 3.3. Testing for differences between variances in compared groups

The next step was to test the assumption of homogeneity of variance, that is, that the results are similarly dispersed around the mean in each subgroup. The null hypothesis of this test is that the variances in the subgroups are identical. For this purpose, the Levene's test was applied.

Table 3 shows the result of Levene's test based on the mean for each dependent variable; the difference between the variances in the compared groups is statistically insignificant (p > 0.05). This means that it can be assumed that the variances of the subgroups are equal or close to each other.

### 3.4. MANOVA testing results

The multivariate analysis of variance (MANOVA) procedure provides regression analysis and analysis of variance for the dependent variables studied according to a single factor of the number of global supply chain partners, categorised into three groups. The hypothesis that there is one mean difference or more between the studied categories of the number of partners and business performance was tested. Table 4 presents the results of the multivariate analysis of variance.

Regarding the main effect of the number of partners based on the above data, three of the four tests obtained the same F-test value and significance. The results of the Pillai's Trace test F(18.392) = 1.78;

Tab. 1. Frequency distribution of the population by study subgroups

| Number of partners in the global supply chain | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Percentage | Percentage of valid | Cumulative percentage |
| Number of partners | 0-3 | 86 | 41.5 | 41.5 | 41.5 |
| | 4-9 | 62 | 30.0 | 30.0 | 71.5 |
| | 10< | 59 | 28.5 | 28.5 | 100.0 |
| | Total | 207 | 100.0 | 100.0 | |

Tab. 2. Box's homogeneity of the covariance matrices test

| $\chi^2$ | df | p |
|---|---|---|
| 251 | 90 | < .001 |

Tab. 3. Tests for homogeneity of variance

| | Levene's test | df1 | df2 | Relevance |
|---|---|---|---|---|
| Lead time | 1.692 | 2 | 209 | 0.187 |
| Efficiency of the entire supply chain | 3.033 | 2 | 209 | 0.050 |
| Logistics costs | 1.795 | 2 | 207 | 0.169 |
| Efficiency in inventory, warehouse management | 0.447 | 2 | 209 | 0.640 |
| Human resource management | 0.810 | 2 | 205 | 0.446 |
| Fulfilment of customer requests | 0.875 | 2 | 209 | 0.418 |
| Mitigation of risk in the supply chain | 1.112 | 2 | 209 | 0.331 |
| Revenue of the company | 2.559 | 2 | 209 | 0.080 |
| Market share: the share of the company in the whole market on which it operates | 0.398 | 2 | 209 | 0.672 |

Tab. 4. Multivariate tests

| | | value | F | df1 | df2 | p |
|---|---|---|---|---|---|---|
| Number of partners | Pillai's Trace | 0.151 | 1.78 | 18 | 392 | 0.026 |
| | Wilks' Lambda | 0.854 | 1.78 | 18 | 390 | 0.026 |
| | Hotelling' Trace | 0.165 | 1.77 | 18 | 388 | 0.026 |
| | Roy's Largest Root | 0.107 | 2.33 | 9 | 196 | 0.017 |

$p = 0.026$ mean that the main effect is statistically significant for $\alpha = 0.05$. The null hypothesis should be rejected in favour of the alternative hypothesis, which assumes that at least one outcome factor has a different mean value for the compared populations. This means that the effect of a factor is different for different levels of another factor.

### 3.5. Explain the nature of the main effect

Once the significance of the main effect of the factor under study was established, an analysis was conducted to explain the significance of the effect. For this purpose, post hoc tests/contrast analysis was used. Due to the lack of an interobjective factor, checking which averages are statistically significantly different. A simple main effect analysis of the number of partners was used. That is, the significance of differences between supply chain effects were compared separately in each group due to the number of partners in the chain. This analysis aims to answer the following questions:

- Whether the effects of the chain, as measured by the Likert scale, are different in the group where the number of partners in the chain is less than or equal to three.
- Whether the effects of the chain, as measured by the Likert scale, are different among the groups where the number of partners in the chain is greater than three but less than nine.
- Whether the effects of the chain as measured by the Likert scale are different among the groups where the number of partners in the chain is equal to or greater than ten.

The NIR post-hoc test, which is the most liberal test, was used to test for differences (Table 5).

The results from the table allow for concluding that there is a significant statistical difference between:

- Lead time and the number of partners in the chain. These results differ in the group of up to ten or more partners in the chain.
- Efficiency of the entire supply chain. These results differ among the three groups studied.
- Efficiency in inventory and warehouse management. These results differ between two groups, i.e.,

Tab. 5. Post-hoc NIR test results

| Dependent variable | (I) number_partne 4. number of partners | (J) number_partne 4. number of partners | Difference in averages (I-J) | Standard error | Relevance | 95% confidence interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower limit | Upper limit |
| Lead time | 0–3 | 4–9 | -.26026 | .17606 | .141 | -.6073 | .0868 |
| | | 10< | -.55059 | .17606 | .002 | -.8977 | -.2035 |
| | 4–9 | 0–3 | .26026 | .17606 | .141 | -.0868 | .6073 |
| | | 10< | -.29032 | .19071 | .129 | -.6663 | .0856 |
| | 10< | 0–3 | .55059 | .17606 | .002 | .2035 | .8977 |
| | | 4–9 | .29032 | .19071 | .129 | -.0856 | .6663 |
| Efficiency of the entire supply chain | 0–3 | 4–9 | -.33798 | .16529 | .042 | -.6638 | -.0121 |
| | | 10< | -.37023 | .16529 | .026 | -.6961 | -.0444 |
| | 4–9 | 0–3 | .33798 | .16529 | .042 | .0121 | .6638 |
| | | 10< | -.03226 | .17905 | .857 | -.3852 | .3207 |
| | 10< | 0–3 | .37023 | .16529 | .026 | .0444 | .6961 |
| | | 4–9 | .03226 | .17905 | .857 | -.3207 | .3852 |
| Logistics costs | 0–3 | 4–9 | .03812 | .18510 | .837 | -.3268 | .4030 |
| | | 10< | -.15758 | .18690 | .400 | -.5261 | .2109 |
| | 4–9 | 0–3 | -.03812 | .18510 | .837 | -.4030 | .3268 |
| | | 10< | -.19570 | .20217 | .334 | -.5943 | .2029 |
| | 10< | 0–3 | .15758 | .18690 | .400 | -.2109 | .5261 |
| | | 4–9 | .19570 | .20217 | .334 | -.2029 | .5943 |
| Efficiency in inventory, warehouse management | 0–3 | 4–9 | .11510 | .17945 | .522 | -.2387 | .4689 |
| | | 10< | -.30425 | .17945 | .091 | -.6580 | .0495 |
| | 4–9 | 0–3 | -.11510 | .17945 | .522 | -.4689 | .2387 |
| | | 10< | -.41935 | .19438 | .032 | -.8026 | -.0362 |
| | 10< | 0–3 | .30425 | .17945 | .091 | -.0495 | .6580 |
| | | 4–9 | .41935 | .19438 | .032 | .0362 | .8026 |
| Human resource management | 0–3 | 4–9 | .17132 | .19030 | .369 | -.2039 | .5465 |
| | | 10< | .02401 | .18849 | .899 | -.3476 | .3956 |
| | 4–9 | 0–3 | -.17132 | .19030 | .369 | -.5465 | .2039 |
| | | 10< | -.14731 | .20488 | .473 | -.5513 | .2566 |
| | 10< | 0–3 | -.02401 | .18849 | .899 | -.3956 | .3476 |
| | | 4–9 | .14731 | .20488 | .473 | -.2566 | .5513 |
| Fulfilment of customer requests | 0–3 | 4–9 | .12023 | .18014 | .505 | -.2349 | .4754 |
| | | 10< | -.13783 | .18014 | .445 | -.4930 | .2173 |
| | 4–9 | 0–3 | -.12023 | .18014 | .505 | -.4754 | .2349 |
| | | 10< | -.25806 | .19513 | .187 | -.6427 | .1266 |
| | 10< | 0–3 | .13783 | .18014 | .445 | -.2173 | .4930 |
| | | 4–9 | .25806 | .19513 | .187 | -.1266 | .6427 |
| Mitigation of risk in the supply chain | 0–3 | 4–9 | .04692 | .16832 | .781 | -.2849 | .3788 |
| | | 10< | -.30792 | .16832 | .069 | -.6397 | .0239 |
| | 4–9 | 0–3 | -.04692 | .16832 | .781 | -.3788 | .2849 |
| | | 10< | -.35484 | .18233 | .053 | -.7143 | .0046 |
| | 10< | 0–3 | .30792 | .16832 | .069 | -.0239 | .6397 |
| | | 4–9 | .35484 | .18233 | .053 | -.0046 | .7143 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Revenue of the company | 0–3 | 4–9 | -.09164 | .18057 | .612 | -.4476 | .2643 |
| | | 10< | -.47874 | .18057 | .009 | -.8347 | -.1228 |
| | 4–9 | 0–3 | .09164 | .18057 | .612 | -.2643 | .4476 |
| | | 10< | -.38710 | .19559 | .049 | -.7727 | -.0015 |
| | 10< | 0–3 | .47874 | .18057 | .009 | .1228 | .8347 |
| | | 4–9 | .38710 | .19559 | .049 | .0015 | .7727 |
| Market share: the share of the company in the whole market on which it operates | 0–3 | 4–9 | .04692 | .16346 | .774 | -.2753 | .3692 |
| | | 10< | -.50147 | .16346 | .002 | -.8237 | -.1792 |
| | 4–9 | 0–3 | -.04692 | .16346 | .774 | -.3692 | .2753 |
| | | 10< | -.54839 | .17707 | .002 | -.8975 | -.1993 |
| | 10< | 0–3 | .50147 | .16346 | .002 | .1792 | .8237 |
| | | 4–9 | .54839 | .17707 | .002 | .1993 | .8975 |

- 4–9 and above ten partners.
- Revenue of the company. These results differ in the group of up to ten or more partners in the chain.
- Market share is the share of the company in the whole market on which it operates. These results differ among up to ten or more chain partners.

No statistically significant differences were shown between the studied groups, i.e., logistics costs, human resource management, fulfilment of customer requests, and mitigation of risk in the supply chain. This means that these supply chain results do not differ due to the number of partners in the chain. Thus, the number of partners in the supply chain does not affect the performance of companies.

The results (Table 6) show that the greater the number of partners in the supply chain, the better the company's performance in the current year compared to the previous year. At the same time, these results were most often significantly better when the number of partners in the supply chain was greater than ten.

The improved results can be traced to the pandemic-related economic crisis that preceded the year of the survey. Nonetheless, a surprising result indicates that shortening supply chains is not necessarily good for business performance. Because the surveyed companies have more cooperators, they improved their performance in terms of lead time — operation time from point a to b; efficiency of the entire supply chain; efficiency in inventory and warehouse management; revenue of the company and market share — the share of the company in the whole market on which it operates. A big difference in improvement is observed, especially in the efficiency of the entire supply chain indicator, where each additional link in the chain can result in better performance for the company operating in it.

For completeness, the percentages of grades for the indicators studied were also analysed for the trend of change, where Grades 1 and 2 indicate a deterioration in performance compared to the previous year, Grade 3 shows no change, and Grades 4 and 5 indicate an improvement in performance compared to the previous year (Table 7).

In the case of the lead time indicator, the relationship is very clear: the higher the number of partners, the more often companies observe an improvement in performance, and conversely, a lower number of partners results in a deterioration of the company's performance. Companies with an average number of partners observe the highest percentage of no change in lead time results. A similar trend applies to the relevant results for efficiency in inventory and warehouse management, where the results improve most often in companies with more than ten partners than in companies with 4–9 partners. Also, in this group, the indicator level remains the same. The indicator of revenue of the company is significantly more often indicated as the number of partners in the chain increases. At the same time, it is in the second group that the deterioration in performance is more often observed than in the other groups. In the group of up to four partners, it is also significantly more often observed that there is no change in the level of revenue compared to the previous period.

Efficiency of the entire supply chain indicates that the most frequent improvement in performance was observed in the middle group of the surveyed population, which at the same time is the least likely to indicate a deterioration in the performance of their companies. A similar positive trend also extends to companies in Group 3. Most often, the lack of improvement in performance is typical of companies that work with the least number of partners in the

Tab. 6. Average values of statistically significant results in the supply chain

| Significant dependent variables | Average values | Number of partners | | |
|---|---|---|---|---|
| | | 1–3 | 4–9 | >10 |
| | | Average rating | | |
| Lead time | 3.3962 | | 3.1591 3.4194 | 3.7097 |
| Efficiency of the entire supply chain | 3.3208 | 3.1136 | 3.4516 | 3.4839 |
| Efficiency in inventory, warehouse management | 3.3962 | - | 3.2258 | 3.6452 |
| Revenue of the company | 3.4623 | 3.2955 | 3.3871 | 3.7742 |
| Market share | 3.4057 | 3.2727 | 3.2258 | 3.7742 |

Tab. 7. Percentage of indications

| Range of evaluation of changes | Number of partners | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Group 1: 1–3 | | | Group 2: 4–9 | | | Group 3: >10 | | |
| | | | | Share% | | | | | |
| | 1-2 | 3 | 4-5 | 1-2 | 3 | 4-5 | 1-2 | 3 | 4-5 |
| Lead time | 29.6 | 29.5 | 40.9 | 16.1 | 38.7 | 45.1 | 9.7 | 32.2 | 58.1 |
| Efficiency of the entire supply chain | 29.5 | 31.8 | 38.6 | 12.9 | 35.5 | 51.7 | 19.4 | 35.5 | 45.2 |
| Efficiency in inventory and warehouse management | 18.2 | 36.4 | 45.4 | 29 | 25.8 | 45.2 | 9.7 | 29 | 61.3 |
| Revenue of the company | 20.4 | 36.4 | 43.1 | 25.9 | 22.6 | 51.7 | 16.1 | 12.9 | 71 |
| Market share | 15.9 | 43.2 | 40.9 | 29 | 32.2 | 38.7 | 9.7 | 29 | 61.3 |

Where:

Positive trend  Neutral trend  Negative trend

supply chain. The trend of changes in the market share indicator is slightly different. It is significantly more often better for Groups 1 and 3. In contrast, considering the lack of change, Group 1 dominates. Group 2, on the other hand, has significantly more often worsened performance than other groups.

## 4. Discussion of the results

In response to the growing geopolitical tensions and dynamic changes of the past few years, there has been increasing talk of a move towards national self-reliance and sovereignty, which could provide some degree of trade independence. This discussion is accompanied by the idea of a shift from globalisation to regionalisation, which could halt or even reverse the progressive process of globalisation and related trade on the world stage. At the same time, the international nature of the activities of most domestic companies may be at odds with this perception of economic reality. The results of the research presented in this article confirm that the complexity of the supply chain and the resulting cooperation with foreign partners positively impact the performance of companies. A detailed analysis of data from companies operating in the manufacturing and service sectors showed that the number of partners in the supply chain significantly impacts key operational and financial indicators. The results indicate that a larger number of partners in the supply chain contributes to improving key performance indicators of the companies participating in it. Similar conclusions were

reached by Saengchai and Jermsittiparserta (2019), who demonstrated a statistically significant correlation between supply chain performance and companies' financial results but did not analyse the level of operational indicators in the context of the number of partners.

This is proven by the results of the research conducted in this article, indicating that the complexity of the supply chain and the associated cooperation with foreign partners affects the use of the results generated in a given entity. This, in turn, means that globalisation continues to benefit and is not fading away. This is also confirmed by the latest DHL report data on international flows of goods, services and capital, the levels of which have returned to pre-pandemic levels. This is especially true for those susceptible to changing macroeconomic conditions and the political environment (Altman & Bastian, 2022). In this case, regionalisation yields better results. During the crisis, this was mainly the case for the tourism industry, but regionalisation can also prove to be much more beneficial for agriculture, industry or energy. In the case of industrial and energy sectors, global supply chains ensure greater stability and access to resources, as confirmed by research conducted by Kano et al. (2020). Companies involved in global cooperation networks gain an advantage through broader access to knowledge and technology, which increases their innovation and operational efficiency, in particular, improving lead times by 18% and inventory efficiency by 22% compared to regional companies, which is consistent with our research results.

Also, studies conducted by Altman and Bastian (2023) suggest that despite the trend towards regionalisation, globalisation continues to play a key role in the world economy. According to them, the overall level of international trade and interconnectedness remains significant, negating large-scale moves toward regionalisation, even if some supply chains have succumbed to it. The study presented in this article complements these findings by showing that companies with more partners in their supply chains achieve better business performance, confirming the importance of global linkages, particularly in highly complex sectors.

An interesting perspective is presented by O'Neil, who believes that the actions of a group of countries that are increasing their exports, like China, India, and Mexico, can have a large impact on global flows (O'Neil, 2023). This study partially confirms the observation, indicating that companies actively participating in global supply chains report greater flexibility in responding to changes in demand. This is also confirmed by research by Kim and Chai (2017), who showed that multi-level supply chains consisting of several links increase operational flexibility by 20% and improve financial performance by 12% compared to simpler regional structures. This research and the results of other studies in this area show that in some sectors, globalisation and global supply chains will remain the dominant arena for trade and cooperation, providing further benefits for all participants.

## Conclusions

The results showed that cooperation with global logistics partners in the supply chain benefits its individual participants. A larger number of partners in the supply chain, i.e., a larger number of links, positively impacts key performance indicators of the company, such as lead time, efficiency of the entire supply chain, efficiency in inventory and warehouse management, company revenue, and market share. Based on the results, a basic recommendation can be made for companies seeking to improve their performance in the global supply chain: they should actively engage in extensive and complex cooperation networks with a larger number of international partners. This approach increases opportunities to operate in international markets, even during periods of temporary crisis.

This study has certain limitations due to its preliminary nature. It relied only on the number of international partners for supply chain complexity. For more accurate results, future research should focus on considering other relevant characteristics of supply chain complexity due to globalisation and regionalisation processes. No analysis of sectoral differences was conducted, which limits the applicability of the conclusions to specific industries. The distance of individual partners in the supply chain is also an important aspect with regard to the study of their impact on the performance of a single company, which was not considered in this study. In addition, business performance was assessed on a five-point scale, which also limits the accuracy of the results and includes the risk of subjectivity in the respondent assessments. Therefore, in future research, it is advisable to collect specific quantitative data on economic and performance indicators based on the actual scale.

Future research should consider an in-depth aspect of the impact of supply chain complexity on company performance depending on the industry

and scope of activity. Various other dimensions of supply chain complexity related to its horizontal and vertical structure, horizontal technological integration between partners, and the dynamics of cooperation should be considered.

## Literature


Ali, H. (2023). Challenges and Solutions in Supply Chain Management. *Journal of technological information, management & engineering sciences*, *1*(02), 68-74.

Alicke, K., Barriball, E., Foster, T., Mauhourat, J., & Trautwein, V. (2022). *Taking the pulse of shifting supply chains*. Retrieved from: https://www.mckinsey.com/capabilities/operations/our-insights/taking-the-pulse-of-shifting-supply-chains#/

Alicke, K., Foster, T., & Trautwein, V. (2024). *Supply chains: Still vulnerable. McKinsey's & Company*. Retrieved from https://www.mckinsey.com/capabilities/operations/our-insights/supply-chain-risk-survey

Altman, S. A., & Bastian, C. R. (2023). DHL global connectedness index 2022. *An in-depth report on the state of globalization*, *301*.

Altman, S. A., & Bastian, C. R. (2023). The State of Globalization in 2023: Data disproves the notion that the world has become more regionalized in recent years. *Harvard Business Review Digital Articles*, 1-14.

Ampuła, D. (2015). Wartości resztowe w procesie regresji [Residuum values in the regress proces]. *Problemy Techniki Uzbrojenia*, *44*, 81-102.

Aranowska, E. & Rytel, J. (2010). Wielowymiarowa analiza wariancji – MANOVA [Multivariate Analysis of Variance – MANOVA]. *Psychologia Społeczna*, *5*(14), 117-141.

Asian Development Bank. (2023). *Trade, Investment, and climate change in Asia and the Pacific. Asian Economic Integration Report 2023*. Retrieved from https://aric.adb.org/pdf/aeir/AEIR2023_highlights.pdf.

Attinasi, M. G., Balatti, M., Mancini, M., & Metelli, L. (2022). Supply chain disruptions and the effects on the global economy. *Economic Bulletin Boxes*, *8*.

Baldwin, R. (2016). *The Great Convergence: Information Technology and the New Globalization*. Cambridge, United States of America: Harvard University Press.

Baller, S., Dutta, S., & Lanvin, B. (2016). *Global information technology report 2016*. Retrieved from https://www3.weforum.org/docs/GITR2016/WEF_GITR_Full_Report.pdf

Banaszyk, P. (2023). Reshoring and friendshoring as factors in changing the geography of international supply chains. *Engineering Management in Production and Services*, *15*(4), 25-33. doi: 10.2478/emj-2023-0026

Bauer, M., & Erixon, F. (2016). The EU's trade with emerging markets: Climbing the value-added chain and growing IP intensity? *ECIPE Occasional Paper*, *1*(2016).

Bauer, M., Zilli, R., & Pandya, D. (2023). The imperative of international cooperation for EU competitiveness and resilience in technology-driven industries. *ECIPE Policy Brief*, *14*(2023).

Belhadi, A., Kamble, S., Jabbour, C. J. C., Gunasekaran, A., Ndubisi, N. O., & Venkatesh, M. (2021). Manufacturing and service supply chain resilience to the COVID-19 outbreak: Lessons learned from the automobile and airline industries. *Technological forecasting and social change*, *163*, 120447. doi: 10.1016/j.techfore.2020.120447

Bielecki, M., Wisniewski, Z., Dukic, G., & Dujak, D. (2024). Supply Chain 4.0: what the supply chains of the future might look like. *Engineering Management in Production and Services*, *16*(1), 77-92. doi: 10.2478/emj-2024-0006

Bondar, M., Mussomeli, A., Cascone, J. & Graeff, K., (2023). *Is your supply chain trustworthy?* Retrieved from https://www2.deloitte.com/us/en/insights/focus/supply-chain/issues-in-global-supply-chain.html

Bujak, A. (2012). *Współczesne zarządzanie (globalnym) łańcuchem dostaw* [*Modern (Global) Supply Chain Management*]. Retrieved from https://polskiprzemysl.com.pl/transport-spedycja-i-logistyka/wspolczesne-zarzadzanie-globalnym-lancuchem-dostaw/

Cernat, L. (2022). Processing trade and global supply chains: Towards a resilient" GVC 2.0" approach. *ECIPE Policy Brief*, *06*(2022).

Champion, A. (2023). *How Globalization Impacts Supply Chain Management*. Retrieved from https://flow.space/blog/globalization-supply-chain/

Chew, Y. C., & Mohamed Zainal, S. R. (2024). A Sustainable Collaborative Talent Management Through Collaborative Intelligence Mindset Theory: A Systematic Review. *SAGE Open*, *14*(2). doi: 10.1177/21582440241261851

Chmielewski, K., & Berczyński, S. (2002). *Statystyka matematyczna. Ćwiczenia laboratoryjne z wykorzystaniem pakietu STATISTICA PL* [*Mathematical Statistics. Practical exercises from the STATISTICA PL package*]. Szczecin, Poland: Wydawnictwo Uczelniane Politechniki Szczecińskiej.

Choi, T. M., & Cheng, T. E. (2011). *Supply chain coordination under uncertainty*. Berlin, Germany: Springer Science & Business Media.

Chopra, S., & Meindl, P. (2016). *Supply Chain Management: Strategy, Planning, and Operation (6th ed.)*. London, England: Pearson.

Christopher, M. (2016). *Logistics and Supply Chain Management. 5th Edition*. London, England: Pearson.

Christopher, M., Peck, H. (2004). Building the Resilient Supply Chain. *The International Journal of Logistics Management*, *15*, 1-14. doi: 10.1108/09574090410700275

Copacino, W. C., & Byrnes, J. L. (2001). How to Become a SUPPLY CHAIN MASTER. *Supply Chain Management Review*, *5*(5), 24-24.

Deloitte Report. (2022). Supply Chain resilience. This report sheds light on resilience in supply chains. Retrieved from https://www.deloitte.com/dk/en/services/consulting/research/supply-chain-resilience-report-2022.html

Denkowska, S., Fijorek, K., Salamaga, M., & Sokołowski, A. (2009). Empiryczna ocena mocy testów dla wielu wariancji [Empirical power rating for multiple variants]. *Przegląd Statystyczny*, *56*(3-4), 26-39. doi: 10.59139/ps.2009.03-04.2

Duong, B. A. T., Truong, H. Q., Sameiro, M., Sampaio, P., Fernandes, A. C., Vilhena, E., & Yadohisa, H. (2019). Supply chain management and organizational performance: the resonant influence. *International Journal of Quality & Reliability Management*, *36*(7), 1053-1077. doi: 10.5267/j.uscm.2021.8.006

Erixon, F. (2018). The economic benefits of globalization for business and consumers. Retrieved from https://ecipe.org/wp-content/uploads/2018/01/Globalization-paper-final.pdf

Erixon, F., & Messerlin, P. (2009). Containing Sino-European Protectionism. *Economic Affairs*, *29*(1), 83-85. doi: 10.1111/j.1468-0270.2009.01874.x

Gereffi, G. (2020). What Does the COVID-19 Pandemic Teach Us About Global Value Chains? The Case of Medical Supplies. *Journal of International Business Policy*, *3*(3), 287-301. doi: 10.1057/s42214-020-00062-w

Gereffi, G., & Fernandez-Stark, K. (2016). *Global Value Chain Analysis: A Primer, 2nd Edition*. Retrieved from https://hdl.handle.net/10161/12488.

Gnanadesikan, R. (2011). *Methods for statistical data analysis of multivariate observations*. Retrieved from http://ndl.ethernet.edu.et/bitstream/123456789/40685/1/Walter%20A.%20Shewhart.pdf

Ivanov, D. (2012). *Management, Modelling, Technology*. Berlin, Germany: Springer.

Jasrotia, S.S., Rai, S.S., Rai, S., & Giri, S. (2024). Stage-wise green supply chain management and environmental performance: Impact of blockchain technology. *International Journal of Information Management Data Insights*, *4*(2), 100241-100250. doi: 10.1016/j.jjimei.2024.100241

Kano, L., Tsang, E. W. K., & Yeung, H. W. C. (2020). Global value chains: A review of the multidisciplinary literature. *Journal of International Business Studies*, *51*, 577-622. doi: 10.1057/s41267-020-00304-2

Kim, Y., & Chai, S. (2017). The impact of supply chain complexity on firm performance: The moderating role of supply chain resilience. *International Journal of Production Economics*, *185*, 18-25. doi: 10.1108/IJRDM-06-2015-0076

Lagarde, Ch. (2022). *A new global map: European resilience in a changing world*. Retrieved from https://www.ecb.europa.eu/press/key/date/2022/html/ecb.sp220422~c43af3db20.en.html

Laire T. (2022). *The Effects of the 2021 Supply Chain Crisis on Manufacturing and Construction and a Forecast for 2022*. Retrieved from https://www.lairedigital.com/blog/the-effects-of-the-supply-chain-crisis-on-manufacturing-and-construction?&referral_domain=www.google.comNaNhttps://www.google.com/

Mena, C., Karatzas, A., & Hansen, C. (2022). International trade resilience and the Covid-19 pandemic. *Journal of Business Research*, *138*, 77-91. doi: 10.1016/j.jbusres.2021.08.064

Mentzer, J. D. (2001). Defining Supply Chain Management. *Journal of Business Logistics*, *22*, 1-25.

Muzyczka, R. (2011). Wpływ globalizacji na zarządzanie łańcuchem dostaw w przemyśle motoryzacyjnym [The impact of globalization on automotive Supply Chain Management]. *International Journal of Management and Economics*, *31*.

O'Neil, S. K. (2023). *It's Not Deglobalization, It's Regionalization. Council Foreign Relations*. Retrieved from: https://www. cfr. org.

Saengchai, S., & Jermsittiparsert, K. (2019). Supply chain and firm performance: examining the moderating role of supply chain information strategy. *Polish Journal of Management Studies*, *20*(2), 454-465.

Semiconductor Industry Association. (2022). *Fact Book*. Retrieved from https://www.semiconductors.org/wp-content/uploads/2022/05/SIA-2022- Factbook_May-2022.pdf.

Sheffi, Y. (2015). Preparing for Disruptions Through Early Detection. *MIT Sloan Management Review*, *57*(1), 30-42.

Tabacnick, B. G., & Fidell, L. S. (2001). Using Multivariate Statistics. Fourty Edition Allyn and Bacon. *Inc*., 996p., New York.

Thuge Lund, C., Visgaard Duch E. (2023). *Report Deloitte, Supply Chain resilience*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/dk/Documents/Consulting/Resilience-Report-Marts-2023.pdf

Todorov, V., & Filzmoser, P. (2010). Robust Statistic for the One-way MANOVA. *Computational Statistics & Data Analysis*, *54*(1), 37-48. doi: 0.1016/j.csda.2009.08.015

UNCTAD. (2023). Development prospects in a fractured world: Global disorder and regional responses. Geneva, Switzerland: United Nations.

Van der Marel, E. (2020). Globalization isn't in decline: It's changing. *ECIPE Policy Brief*, 6(2020).

Vidrova, Z. (2020). Supply chain management in the aspect of globalization. *SHS Web of Conferences*, *74*. doi: 10.1051/shsconf/20207404031

Wątroba, J. (2012). Jak poprawnie zastosować wielowymiarową analizę wariancji (MANOVA) – pokonać trudności i osiągnąć korzyści. Retrieved from https://media.statsoft.pl/_old_dnn/update/downloads/jak_poprawnie_zastosowac_wielowymiarowa_analize_wariancji.pdf

Zhang, Y., Zhao, C., & Pang, B. (2018). Budget allocation in coping with supply chain disruption risks. *International Journal of Production Research*, *56*(12), 4152-4167. doi: 10.1080/00207543.2018.1430905

Zhou, L. (2024). *Small Business Statistics: The Ultimate List in 2024*. Retrieved from https://www.luisazhou.com/blog/small-business-statistics/

Volume 17 • Issue 2 • 2025

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 31 January 2025
accepted: 5 June 2025

pages: 78-89



# System quality, information quality, perceived usefulness, and ease of use as determinants of intention to continue using a Digital Accounting System among Jordanian SMEs

Moh'd Alsqour
Hani Attar
Mohammad Haider Alibraheem
Sameh Alsaqoor
Enas Alsaleem

ABSTRACT

This study aims to investigate the determinants of the intention to continue using digital accounting systems (ICU-DAS) among Jordanian small and medium-sized enterprises (SMEs). This study used a quantitative approach with questionnaires distributed to o 440 Jordanian SMEs. The sample included key decision-makers such as managers and owners, primarily from SMEs. A total of 318 sets of questionnaires were returned with complete responses and were examined further. The research model was analysed using partial least squares (PLS) with the Smart-PLS software. The results of this investigation revealed that system quality, information quality, perceived usefulness, and perceived ease of use significantly influence the intention to continue using digital accounting systems. This study bridges the gap between theory and practice, offering academic and practical insights into the factors influencing the continued use of digital accounting systems in SMEs. This research contributes significantly to understanding technology continuance, particularly in the context of SMEs in Jordan. Integrating ISSM, TAM, and ECM provides insights into how these systems are adopted and sustained in less-developed regions. The study provides actionable recommendations for SMEs and policymakers. SMEs can enhance SQ, IQ, and user satisfaction to improve adoption and long-term use by understanding the key factors driving DAS continuance. Policymakers can leverage these findings to create policies that encourage SMEs to adopt and maintain DAS, contributing to the region's economic growth.

KEY WORDS

**system quality, information quality, perceived usefulness, perceived ease of use, digital accounting systems**

**10.2478/emj-2025-0013**

**Enas Alsaleem**
Department of Accounting,
Jerash Private University, Jerash-Jordan
ORCID 0000-0003-2539-407X

Corresponding author:
e-mail: enas.alsalim@yahoo.com

**Moh'd Alsqour**
Department of Accounting,
Irbid National University, Irbid, Jordan
ORCID 0000-0003-0167-3429

**Hani Attar**
Energy Engineering, Department,
Faculty of Engineering Technology,
Zarqa University, Zarqa, Jordan
College of Engineering, University of Business and Technology, Jeddah, 21448, Saudi Arabia
ORCID 0000-0001-8028-7918

**Mohammad Haider Alibraheem**
Accounting Information System Department,
Faculty of Administrative and Financial Sciences, Irbid National University, Jordan
ORCID 0000-0001-6850-0480

**Sameh Alsaqoor**
Tafila Technical University
PO Box 179, 66110 Tafila, Jordan
ORCID 0000-0002-0552-8926

## Introduction

Nearly every industry, including accounting, endured change caused by information technology. As a result, the importance of information technology and systems in company operations has been growing. Consequently, businesses have been relying more and more on information systems and technologies to stay competitive and swiftly adjust to changes (Ritchi et al., 2019; Jaklič et al., 2018; Almaqtari et al., 2023; Al-Hattami et al., 2021c). Employees, managers, and legislators must make numerous decisions in the modern business environment all the time, usually making

Alsqour, M., Attar, H., Alibraheem, M. H., Alsaqoor, S., & Alsaleem, E. (2025). System quality, information quality, perceived usefulness, and ease of use as determinants of intention to continue using a Digital Accounting System among Jordanian SMEs. *Engineering Management in Production and Services, 17*(2), 78-89. doi: 10.2478/emj-2025-0013

decisions supported by adequate and trustworthy information and facts (Harris, 2018; Kowalczyk & Buxmann, 2015; Al-Hattami & Kabra, 2022; Jaklič et al., 2018). Digital accounting systems (DAS) are essential decision-support tools that improve corporate decisions by delivering accurate, timely, and substantial information regarding finances and accounting (Alawaqleh & Al-Sohaimat, 2017; Romney & Steinbart, 2016).

DAS are now widely recognised as essential instruments for data storing, financial transaction management, and decision-making facilitation. DAS are becoming increasingly more popular among small and medium-sized businesses (SMEs) aiming to improve their efficiency and consolidate financial processes (Al-Hattami, 2022; Al-Hattami et al., 2021b). Research indicates that implementing DAS benefits these companies by enhancing their organisational performance (Al-Hattami et al., 2021b; Grande et al., 2011). In particular, studies show that decision-making using DAS results in better quality decisions (Al-Hattami, 2022). DAS is seen as a crucial resource for business rivalry and leadership success, and this view incentivises companies to increase their efforts in the field (Mohammad, 2018; Ritchi et al., 2019; Zain & Hussin, 2019; Al-Hatami et al., 2022). Nevertheless, SMEs' intentions to continue employing these systems are influenced by several elements necessary for the effective development and ongoing use of these platforms. According to the research, a company's ability to effectively apply information technology and systems depends more on persistence than on adoption or its use at first, which could be met with resistance (Mishra et al., 2023; Bhattacherjee, 2001; Zain & Hussin, 2019; Yan et al., 2021; Hayat et al., 2020). Although DAS adoption can help with practical decision-making (Al-Hattami, 2021; Al-Hattami et al., 2022b; Akrong et al., 2022; Al-Hattami, 2022), more research is required to identify the factors contributing to user intentions to continue using DAS (ICU-DAS) following adoption. Based on the authors' best knowledge, user-friendly ICU-DAS have received relatively little attention thus far, especially for organisational uses.

Essam et al. (2024) provided an in-depth survey of quality techniques for question analysis in Arabic question answering that helped improve the search quality in the Arabic language. The available literature often provides extensive research into user acceptance along with sustained usage of information technology systems (IT/systems) (Hou, 2016; Venkatesh et al., 2012; Ashfaq et al., 2020; Chiu et al., 2021; Al-Hattami et al., 2021; Qutaishat et al., 2023; Al-Debei et al., 2022). However, behaviour intention to continue using DAS has not been well studied (Li & Wang, 2021). Remarkably, there has been relatively little research on the intention to continue using DAS, especially among SMEs in less developed nations. Based on studies, the COVID-19 pandemic resulted in numerous obstacles and problems for SMEs (Adam & Alarifi, 2021). According to researchers' predictions (Guo et al., 2020; Al-Hatami et al., 2022), SME strategies and activities should centre on the exploitation and execution of digital technology, particularly accounting systems. Regretfully, no research has yet examined the factors determining the users' ICU-DAS during this timeframe or setting. This research attempts to address the research gaps by identifying the primary motivators for organisational users to continue adopting DAS. To be more precise, this study supported the information systems success model (ISSM), expectation confirmation model (ECM), and the technology acceptance model (TAM) theoretical frameworks to look at how the perceived usefulness (PU), perceived ease of use (PEU), system quality (SQ), and information quality (IQ) affect SMEs' intentions to continue using DAS. Based on the authors' best knowledge, no such practical research has been conducted in Jordan. The results showed that compared to ISSM, ECM, and TAM, the synthesised paradigm by Al-Hattami and Almaqtari (2023) had a greater explanation capacity, where the results also suggested that important variables in ICU-DAS in SMEs include PU, SQ, PEU, IQ, and satisfaction. Moreover, the significance of these study findings lies in their ability to shed light on the changing patterns of DAS uptake and persistence in the world of SMEs. The study's results may help SMEs, legislators, technology suppliers, and the whole economy by promoting sustainable digitalisation initiatives and enabling better decision-making. Thus, this empirical study aims to investigate the factors that influence Jordanian SMEs' intentions to continue using digital accounting systems, namely IQ, SQ, PU, and PEU. The review of relevant literature, conceptual framework, formulation of methodology, hypotheses, and the analytical part are covered in the following section.

## 1. Literature review

Most enterprises in Jordan, an Arab Peninsula nation, are small and medium-sized enterprises (SMEs). In fact, 97% of all privately registered enterprises in Jordan are SMEs (ASI, 2013). Al-Hattami

and Almaqtari (2023) defined the Jordanian enterprises as companies with up to 50 employees. In Jordan, SMEs are critical to reducing poverty, fostering economic expansion, and creating jobs. In particular, SMEs are thought to be a driver of economic expansion, a significant contributor to equitable development, and especially a source of the fastest job creation (Al-Hattami, 2022). Furthermore, given the volume of capital spent, their impact on the total GDP is noteworthy, confirming their critical role in development that extends well beyond large enterprises (OCHA, 2021). However, Jordanian SMEs face obstacles and problems, including the detrimental effects of COVID-19 and the nation's political unrest (Saleh & Manjunath, 2020; OCHA, 2021). These businesses have the power to bring about socio-economic change in Jordan, notwithstanding these obstacles. SMEs may increase their reach, enhance their efficiency, and aid in the recovery of the economy by utilising technology, particularly DAS (Al-Hattami, 2022; Abdullah et al., 2018; Al-Hattami et al., 2022a). Al-Hattami et al. (2021b) stated that many Jordanian SMEs are still ignorant of the significance and function of such systems in boosting their success. Several SMEs stopped utilising these systems after a while. The factors influencing the Jordanian SMEs' acceptance and utilisation of DAS have been the subject of several research efforts (Al-Hattami, 2022; Al-Hattami et al., 2022a, b; Al-Hattami & Kabra, 2022). It is crucial to research the factors encouraging users to continue using DAS because the success of businesses utilising information technology/systems relies on the ongoing use compared with the implementation or first use, which might be met with resistance (Mishra et al., 2023; Bhattacherjee, 2001; Yan et al., 2021; Zain & Hussin, 2019; Yan et al., 2021). However, to the authors' best knowledge, no previous research has looked at the factors encouraging DAS users in Jordanian SMEs to keep using them. This literature gap emphasises the need for more studies to improve the comprehension of the variables impacting the adoption and continued use of DAS in this particular setting. In Jordan's SME industry, where efficient financial management may have a major impact on company performance and economic growth, it is imperative to investigate these factors not just for academic purposes but also to guide policy and practice (Nair et al., 2020; Al-Hattami & Kabra, 2022).

Over time, the business world actors (managers, accountants, and finance specialists) are likelier to remain with a particular digital accounting system. Factors that could affect continuation include system quality, accuracy of the information, ease of use, overall user experience, and perceived usefulness (Li & Wang, 2021; Bhattacherjee, 2001; Cheng, 2020, Ujawry-Gil et al., 2025). The decision to keep using a digital accounting system impacts business operations because it enhances decision-making, helps with adherence to regulations, and increases the precision and efficacy of financial procedures (Cheng, 2020; Veeramootoo et al., 2018; Almaqtari et al., 2023).

### 1.1. Underpinning theory / Theoretical framework

DeLone and McLean (1992) utilised the ISSM, which is currently widely used as a framework to measure the efficacy and achievement of IS, to deliver a significant literature contribution to measuring the outcomes of IS (Mehta et al., 2022). Three categories comprise ISSM: quality, use, and impacts. The first category includes two structures (SQ and IQ) as well as the second (use and satisfaction), and the third category comprises two constructs (individual impact and organisational impact) (McLean & DeLone, 1992). For instance, depending on the investigation's goal, a certain set of scales or constructs may be employed (Floropoulos et al., 2010; Fadelelmoula, 2018). Additional aspects or constructs independently of or subservient to additional models, like ECM, may be added to the model (Cheng, 2019; Khayer et al., 2020; Veeramootoo et al., 2018; Franque et al., 2021; Akrong et al., 2022). This provided the model with the adaptability to measure various IS precisely in a variety of settings; additionally, this allowed for the approach to be applied and used at any assessment stage that the investigator felt was most appropriate (Petter et al., 2008; Dalloul et al., 2023; Al-Hattami, 2022). This aspect of ISSM is determined by the reliability dimension in conjunction with the functionality dimension since the current study gauges user intentions to stick with DAS. Furthermore, the work combines two well-known models (TAM and ECM) using ISSM to evaluate the continuity of intention within the realm of DAS.

TAM theory of IS, which explores how people embrace and employ technology, was first introduced by Davis (1989). Moreover, TAM typically refers to the ultimate decision made by users to buy, install, or utilise a system. It considers two essential components, PU and PEU, that emerge in most major examples of technology adoption (Verkasalo, 2008; Davis, 1989). These concepts work together to ensure the user's intention to utilise the investigated technol-

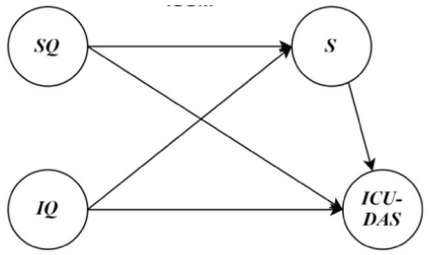



Fig. 1. ISSM model

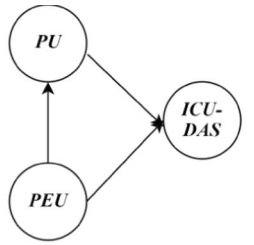



Fig. 2. TAM model

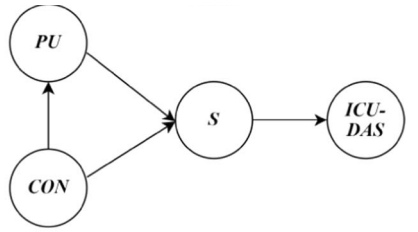



Fig. 3. ECM model

ogy, which ultimately results in the system being used (Schöpfel et al., 2019; Charness & Boot, 2016). To be more precise, TAM proposes that PU influences behavioural intent in both directions, while PEU influences PU, and PU affects behavioural intent in both directions (Davis, 1989). TAM employs the "attitude" of a mediator, but "satisfaction" may rather be employed because satisfaction at the post-usage phase evaluates pre-usage attitudes, use intentions, and plans to continue using information technology or systems, all of which are related constructs (Cheng, 2020; Bhattacherjee, 2001). Scientific investigations reaffirmed the importance of incorporating technology acceptance and usage of TAM with other aspects related to use continuation, and TAM investigation has shed light on these areas (Hou, 2016; Ashfaq et al., 2020; Zhou et al., 2018). Aiming to ascertain the ongoing use of DAS, the current study extended the TAM by ISSM along with ECM, following earlier research regarding continuous use.

The enterprise continuity paradigm (ECM) is widely used in various fields concerning the sustainability of information technology and systems (Veeramootoo et al., 2018; Mishra et al., 2023; Cheng, 2020). ECM assumes that a user's choice to continue depends on the level of post-adoption satisfaction, PU, and confirmed information technology/systems (Bhattacherjee, 2001; Cheng, 2020). However, behavioural intention towards continued use of IS may also be influenced by other constructs, such as SQ, IQ, and ease of use. Thus, to create a robust IS consistency model within the DAS setting, this work explores the feasibility of combining several elements from additional models, such as ISSM and TAM. Some investigators (e.g., Ashfaq et al., 2020) have backed such synthesis, albeit in a different setting and culture than

DAS. Therefore, applying the components of these models to the DAS would provide a broader viewpoint for the IS continuation study.

### 1.2. Research conceptual framework

Based on the three models mentioned above, the proposed investigation model has five components, as depicted in Fig. 1, which are connected around the ICU-DAS objective. The concepts shown in Fig. 1 are interrelated and aid in comprehending ICU-DAS. Enhanced SQ may positively impact other model elements since happy users are more satisfied with well-functioning systems. Additionally, reliable evidence for making decisions is made possible by good IQ, which particularly raises customers' perceptions of value and satisfaction (Al-Hattami, 2022; Floropoulos et al., 2010). Easy-to-use technologies can potentially increase user satisfaction and PU in a similar environment. Furthermore, affirmative validation enhances contentment and PU, but adverse validation could lead to discontent and a reduced propensity to employ the system going forward (Mishra et al., 2023; Bhattacherjee, 2001). In conclusion, ICU-DAS displays user preparedness to implement the digital accounting software at a later time. It might be affected by every construct that had come prior to it. When users think a system is useful for their work, easy to use, provides accurate information, and is of high quality, they are also inclined to express a positive desire to continue using it. This model helps academics and practitioners comprehend the complexities that affect ICU-DAS by providing a thorough view of the relationships among these aspects.

### 1.3. Hypothesis development

This work presents four hypotheses to undergo empirical tests, which are described below, reflecting prior research and the aforementioned underlying theories.

According to ISSM, satisfaction and usage intention are impacted by SQ (DeLone & McLean, 2003; Petter et al., 2008; DeLone & McLean, 1992), and system quality (SQ) stands for DAS's technological ability to provide users with reliable and fast access to information with adequate security (Belfo &Trigo, 2013; Al-Hattami, 2021a). Numerous studies show that SQ influences continuing intentions along with satisfaction favourably in various situations. Veeramootoo et al. (2018) discovered, for instance, a strong correlation between SQ with satisfaction and continued intention of utilising e-filing. Sharma and Sharma (2019) discovered that SQ had no effect on customer satisfaction or the propensity to keep utilising mobile banking facilities. Zheng et al. (2013) concluded that, in a different situation, SQ influences satisfaction directly and intent to continue indirectly. Al-Hattami et al. (2022b), Kabra (2022), and Al-Hattami (2022) discovered that SQ favourably impacted satisfaction and actual usage within the setting of DAS. Li and Wang (2021) found that contentment is positively impacted by SQ. Nevertheless, there was no research on the connection between SQ and ICU-DAS, especially with regard to Jordanian SMEs. Thus, the following pertinent pathways are outlined:

H1: SQ positively affects ICU-DAS.

The impact of IQ on user satisfaction and intention to utilise is a topic covered by ISSM (DeLone & McLean, 2003; Petter et al., 2008; DeLone

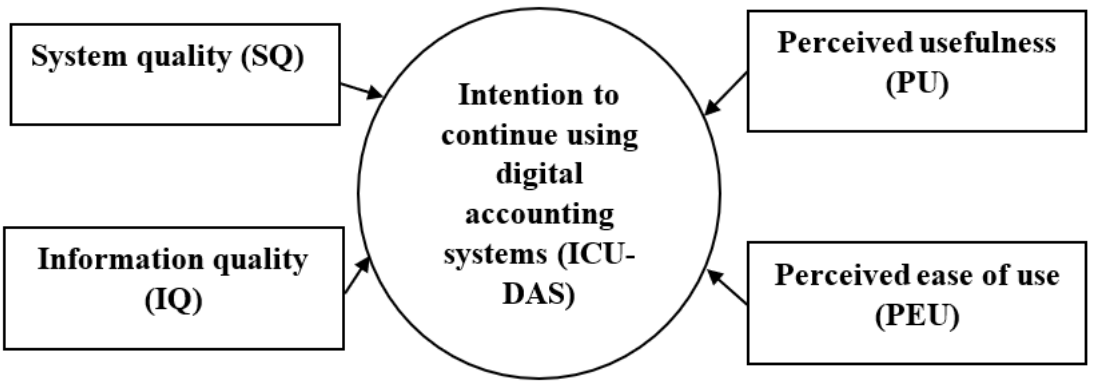



Fig. 4. Research framework

& McLean, 1992). IQ is a measure of suitability for use in institutions, according to Xu et al. (2003). According to Al-Hatami (2022), a decision-making support system (DAS) is deemed appropriate if it produces comprehensive, pertinent, timely, and accurate information. Beneficial effects were seen at the DAS degree when IQ was considered for satisfaction and real usage (Al-Hattami & Kabra, 2022; Al-Hattami, 2022; Al-Hattami et al., 2022b). However, Li and Wang (2021) suggested that IQ influences satisfaction directly while indirectly influencing the will to continue. Sharma and Sharma (2019) discovered that SQ had a favourable impact on satisfaction and intention to continue in a different setting. Nevertheless, no research has been done on this effect in the context of SMEs within Jordan. Consequently, this research suggests:

H2: IQ positively affects ICU-DAS.

Among independent constructs within TAM and ECM, PU is regarded as usefulness (Fig. 1). Thus, behavioural intention and satisfaction are directly impacted by PU (Bhattacherjee, 2001; Davis, 1989). In the fields of internet banking (Rahi et al., 2022), internet-based shopping (Al-Hattami, 2021b), and online learning (Suzianti & Paramadini, 2021), the importance of PU has been widely acknowledged. They define PU as the arbitrary potential that a user's performance on a given task would be enhanced by the application of technology. They stated that PU is one important factor influencing behavioural intention and pleasure. In the case of DAS, PU may positively affect satisfaction and ICU-DAS. Likewise, individuals who consider that using DAS could boost their job performance tend to be content with it while also obtaining an extra ideal continuation intention of DAS use (Floropoulos et al., 2010; Cheng, 2020). Consequently, the following theories are put forth throughout this paper:

H3: PU positively affects ICU-DAS.

Users gain more from simpler technology (Mishra et al., 2023). According to TAM, PU and behavioural intention are highly impacted by PEU (Cheng, 2020). Ramayah et al. (2012) claimed that the primary determinant of behavioural intention was the ease of use or accessibility of a system. Cheng (2019) adds that user happiness and their views of DAS effectiveness significantly improve when they feel the technology is easy to use and requires minimal effort and time. In a similar vein, Cheng (2020) found that PEU increases continuation intention and improves satisfaction. Thus, it is presumed that:

H4: PEU positively affects ICU-DAS.

# 2. Research methods

## 2.1. Measurement of variables

The model for investigating the intention to continue using DAS focused on four key constructs: "system quality", "information quality", "perceived usefulness", and "perceived ease of use". Each construct was measured using scales adapted from the study by Al-Hattami and Almaqtari (2023) and customised to reflect the context of DAS in Jordanian SMEs. A total of 15 items were used to assess these five constructs, with each item rated on a five-point Likert scale.

## 2.2. Data collection

For this research, a survey was distributed to 440 Jordanian SMEs, and 318 responses were received, yielding a response rate of 72.3%. Non-response bias was not a significant concern, as the high response rate helped ensure representativeness. The confidentiality of respondents' answers was emphasised, contributing to the high participation rate. After data cleaning, 308 valid responses were used for analysis. The sample included key decision-makers such as managers and owners, primarily from SMEs employing up to 50 people. These individuals with an accounting background provided critical insights into DAS usage and its future prospects. The respondents were from diverse sectors, including manufacturing, services, and commerce, ensuring a broad range of perspectives. Table 5 shows the factor loadings, Cronbach's alpha (CA), and composite reliability (CR), which were all satisfactory, ensuring the reliability and validity of the measurement items.

## 2.3. Statistical tool

The research model was analysed using partial least squares (PLS) with the Smart-PLS software. Smart-PLS was chosen over traditional regression and covariance-based SEM (CB-SEM) due to its advantages in handling complex models with multiple constructs. It is well-suited for non-normally distributed data and small-to-medium sample sizes, making it an ideal choice for this study. Bootstrapping techniques were used to minimise estimation bias and improve the robustness of the analysis. PLS is frequently utilised in technology acceptance studies due to its flexibility and efficiency in processing intricate models (Hair et al., 2017; Henseler et al., 2016).

### 2.4. Data analysis and measurement model

The main objective of the measurement model is to filter the data, which is to assess and confirm the constructs' validity and reliability before establishing the goodness of measures. The data were examined through the indicator's reliability, whereas 0.4 is accepted. For internal consistency, using composite reliability, 0.7 is considered the accepted level. Convergent validity using average variance extracted (AVE), which must be 0.5 and above (Chin, 1998). For discriminant validity using factor loading, any item loading on the other construct higher than their loadings should be deleted (Chin, 1998; Hair, 2010). Hence, it is resolved that the instrument adapted in this study is reliable since none of the items is with less than 0.4.

Tab. 1. Factor loading

| Items | Factor loadings | Composite reliability | AVE |
|---|---|---|---|
| SQ1 | 0.790 | | |
| SQ2 | 0.698 | 0.781 | 0.544 |
| SQ3 | 0.722 | | |
| IQ1 | 0.667 | | |
| IQ2 | 0.826 | 0.821 | 0.606 |
| IQ3 | 0.831 | | |
| PU1 | 0.789 | | |
| PU2 | 0.794 | 0.826 | 0.548 |
| PU3 | 0.557 | | |
| PEU1 | 0.793 | | |
| PEU2 | 0.719 | 0.735 | 0.567 |
| PEU3 | 0.775 | | |
| ICU-DAS1 | 0.813 | | |
| ICU-DAS2 | 0.569 | 0.896 | 0.742 |
| ICU-DAS3 | 0.688 | | |

Tab. 2. Discriminant validity

| Constructs | SQ | IQ | PU | PEU | ICU-DAS |
|---|---|---|---|---|---|
| SQ | **0.871** | | | | |
| IQ | 0.594 | **0.871** | | | |
| PU | 0.298 | 0.515 | **0.814** | | |
| PEU | 0.002 | 0.237 | 0.522 | **0.753** | |
| ICU-DAS | 0.002 | 0.237 | 0.522 | 0.645 | **0.861** |

All items loaded on their respective construct range from 0.536 to 0.831, which is deemed acceptable since all values range above the cut-off point of 0.4 (Chin, 1998; Hair, Ringle & Sarstedt, 2011). Equally, the values of composite reliability range from 0.781 to 0.941, which are also greater than the recommended value of 0.7 (Hair et al., 2011). AVE was used to determine the convergence validity. The AVE values range from 0.518 to 0.606, which are above the minimum cut-off point of 0.5 (Hair et al., 2011). Lastly, to determine the discriminant validity, the AVE is compared to the correlation squared of the interrelated variables of concerned constructs, which indicates adequate discriminant validity. Table 1 presents factor loading, and Table 2 presents the discriminant validity.

## 3. Research results

The next step in the data analysis process was testing the research model and its hypothesised relationships. This study specifically focused on evaluating the role of SQ, IQ, PU, and PEU in influencing the ICU-DAS among Jordanian SMEs. The SEM approach was employed to test the hypothesised paths and determine the explanatory power of the model.

The results indicated that all four constructs significantly contributed to explaining the intention to continue using DAS. Specifically, SQ, IQ, and PU had a direct positive impact on and positively influenced ICU-DAS. PEU and PU were also significant predictors of ICU-DAS. Based on Table 3, the statistical analysis proved that H1 is supported where SQ is significantly and positively related to ICU-DAS ($\beta$=.115; t=2.507). There is a significant positive influence of IQ on ICU-DAS ($\beta$=.213; and t=3.739), so H2 is supported. H3 and H4 were also supported, indicating a positive significant influence of PU and PEU on ICU-DAS ($\beta$=.250; t=4.585) and ($\beta$=.192; t=3.228), respectively.

Overall, the model explained 63.6% of the variance in the intention to continue using DAS, demonstrating a strong fit to the data, which suggests that SQ, IQ, PU, and PEU are significant factors in determining the continued adoption of digital accounting systems among Jordanian SMEs. Additionally, the model's explanatory power was higher than that of previous models, confirming the importance of integrating these four factors for a comprehensive understanding of DAS continuance.

In conclusion, this study highlights the critical role of a system and IQ, PU, and PEU in the decision

Tab. 3. Hypothesis testing results

| Hypotheses | Relationships | Beta | SE | T Statistics |
|---|---|---|---|---|
| H1 | SQ -> ICU-DAS | 0.115* | 0.046 | 2.507 |
| H2 | IQ -> ICU-DAS | 0.213*** | 0.057 | 3.739 |
| H3 | PU -> ICU-DAS | 0.250*** | 0.055 | 4.585 |
| H4 | PEU -> ICU-DAS | 0.192** | 0.059 | 3.228 |

of SMEs to continue using digital accounting systems. These findings offer valuable insights for SMEs and accounting professionals aiming to enhance the usability and effectiveness of digital accounting technologies.

## 4. Discussion of the results

The ICU-DAS is critical in predicting user behaviour post-adoption, as it reflects the decision to persist with system usage after initial implementation (Bhattacherjee, 2001; Chen et al., 2021). This study focused on the key factors influencing ICU-DAS among Jordanian SMEs, specifically examining SQ, IQ, PU, and PEU.

The findings supported the proposed hypotheses, revealing that SQ positively influences satisfaction and ICU-DAS, which is consistent with previous research (Al-Hattami, 2022; Li & Wang, 2021). This indicates that when the system provides flexibility, accessibility, and reliability, users are more likely to continue using it. IQ also had a significant positive effect on satisfaction and ICU-DAS, confirming prior studies (Floropoulos et al., 2010; Al-Hattami & Kabra, 2022), which found that relevant, timely, and accurate information leads to higher user satisfaction and continued system use.

PU was another significant determinant of both satisfaction and ICU-DAS. In the context of mandatory use, such as Jordanian SMEs, users are more likely to continue using a system if they perceive it as useful for improving work efficiency (Floropoulos et al., 2010; Fadelelmoula, 2018). This study's findings align with this conclusion, showing that a system perceived as useful leads to greater satisfaction and a higher likelihood of continued use (Cheng, 2020).

The PEU was equally important, as simpler systems reduce cognitive load and allow users to complete tasks more efficiently (Mishra et al., 2023). The study confirmed that PEU positively influences PU and ICU-DAS, with users more likely to adopt and continue using systems that are intuitive and easy to operate (Kumar & Natarajan, 2020).

Satisfaction emerged as a strong predictor of ICU-DAS, supporting the expectation confirmation theory proposed by Bhattacherjee (2001). Users who are satisfied with the system are more likely to continue using it, which is in line with previous studies on technology adoption and continued use (Cheng, 2020; Hou, 2016; Zhou et al., 2018).

In summary, SQ, IQ, PU, and PEU are crucial in influencing user satisfaction and the intention to continue using digital accounting systems among Jordanian SMEs. High-quality systems reduce errors, ensure data accuracy, and offer a seamless user experience while accurate and relevant financial data support decision-making and compliance (Al-Hattami & Kabra, 2022). Systems that are easy to use and perceived as useful enhance user motivation, leading to continued usage and long-term adoption.

## Conclusions

This research focused on the determinants of the intention to continue using digital accounting systems. The study employed several models, such as ISSM, TAM, and ECM, to find that SQ, IQ, PU, and PEU provided the most comprehensive explanation for continuance intention, accounting for 63.6% of the variance. Key findings revealed that SQ, IQ, PU, and PEU significantly influence the intention to continue using DAS. Additionally, satisfaction was influenced by SQ, IQ, and PU.

High SQ ensures a seamless user experience by reducing technical issues and instilling confidence in the system's performance. IQ guarantees that financial data is accurate and relevant, improving user interaction and satisfaction. Users are likelier to continue using DAS if they perceive that the system enhances productivity, PU, and PEU. A system that is complex or difficult to use may lead to user frustration, reducing the likelihood of continued usage.

Furthermore, satisfaction plays a crucial role, as users who are satisfied with DAS are more inclined to continue using it.

This research contributes significantly to the understanding of technology continuance, particularly in the context of SMEs in Jordan. Integrating ISSM, TAM, and ECM provides insights into how these systems are adopted and sustained in less-developed regions. The findings serve as a foundation for future research on technology adoption, offering a framework for examining continuance intention across different sectors and regions.

On a practical level, the study provides actionable recommendations for SMEs and policymakers. SMEs can enhance SQ, IQ, and user satisfaction by understanding the key factors driving DAS continuance to improve adoption and long-term use. Policymakers can leverage these findings to create policies that encourage SMEs to adopt and maintain DAS, contributing to the economic growth of the region.

In summary, this study bridges the gap between theory and practice, offering academic and practical insights into the factors influencing the continued use of digital accounting systems in SMEs.

Lastly, the limitations of the study should be noted. Future research should consider extended models, examine demographic variables, and explore longitudinal effects as DAS continues to evolve. Additionally, the lack of a direct relationship between confirmation and satisfaction requires further investigation, and the applicability of these findings to other regions and industries should be explored.

## Literature

Abdullah, A., Thomas, B., Murphy, L., & Plant, E. (2018). An investigation of the benefits and barriers of e-business adoption activities in Yemeni SMEs. *Strategic Change*, *27*(3), 95-208.

Adam, N. A., & Alarifi, G. (2021). Innovation practices for survival of small and medium enterprises (SMEs) in the COVID-19 times: the role of external support. *Journal of Innovation and Entrepreneurship*, *10*(1), 1-22.

Afthanorhan, W. M. A. B. W. (2013). A comparison of partial least square structural equation modeling (PLS-SEM) and covariance based structural equation modeling (CB-SEM) for confirmatory factor analysis. *International Journal of Engineering Science and Innovative Technology*, *2*(5), 98-205.

Alawaqleh, Q., & Al-Sohaimat, M. (2017). The relationship between accounting information systems and making investment decisions in the industrial companies listed in the Saudi Stock market. *International Business Research*, *10*(6), 199-211. doi: 10.5539/ibr.v10n6p199

Al-Debei, M. M., Dwivedi, Y. K., & Hujran, O. (2022). Why would telecom customers continue to use mobile value-added services? *Journal of Innovation & Knowledge*, *7*(4), 100242.

Al-Hattami, H. M. (2021). University accounting curriculum, IT, and job market demands: Evidence from Yemen. *Sage Open*, *11*(2), 1-14. doi: 10. 1177/21582440211007111

Al-Hattami, H. M. (2021a). Validation of the D&M IS success model in the context of accounting information system of the banking sector in the least developed countries. *Journal of Management Control*, *32*(1), 127-153.

Al-Hattami, H. M. (2021b). Determinants of intention to continue usage of online shopping under a pandemic: COVID-19. *Cogent Business & Management*, *8*(1), 1936368.

Al-Hattami, H. M. (2022). Impact of AIS success on decision-making effectiveness among SMEs in less developed countries. *Information Technology for Development*, 1-21. doi: 10.1080/02681102.2022.2073325

Al-Hattami, H. M., & Almaqtari, F. A. (2023). What determines digital accounting systems' continuance intention? An empirical investigation in SMEs. *Humanities and Social Sciences Communications*, *10*(1), 1-13.

Al-Hattami, H. M., & Kabra, J. D. (2022). The influence of accounting information system on management control effectiveness: The perspective of SMEs in Yemen. *Information Development*, 02666669221087184. doi: 10. 1177/02666669221087184

Al-Hattami, H. M., Abdullah, A. A. A. H., Kabra, J. D., Alsoufi, M. A. Z., Gaber, M. M. A., & Shuraim, A. M. A. (2022b). Effect of AIS on planning process effectiveness: a case of SMEs in a less developed nation. *The Bottom Line*, *35*(2/3), 33-52. doi: 10. 1108/BL-01-2022-0001

Al-Hattami, H. M., Abdullah, A. A. H., & Khamis, A. A. A. (2021a). Determinants of intention to continue using internet banking: Indian context. *Innovative Marketing*, *17*(1), 40–52. doi: 10.21511/im.17(1).2021.04

Al-Hattami, H. M., Hashed, A. A., & Kabra, J. D. (2021b). Effect of AIS success on per- formance measures of SMEs: evidence from Yemen. *International Journal of Business Information Systems*, *36*(1), 144-164. doi: 10.1504/IJBIS. 2021.112399

Al-Hattami, H. M., Hashed, A. A., Alnuzaili, K. M., Alsoufi, M. A., Alnakeeb, A. A., & Rageh, H. (2021c). Effect of risk of using computerized AIS on external auditor's work quality in Yemen. *International Journal of Advanced and Applied Sciences*, *8*(1), 75-81. doi: 10.21833/ijaas.2021.01.010

Al-Hattami, H. M., Senan, N. A. M., Al-Hakimi, M. A., & Azharuddin, S. (2022a). An empirical examination of AIS success at the organizational level in the era of COVID- 19 pandemic. *Global Knowledge, Memory and Communication*, *73*(3), 312-330. doi: 10.1108/GKMC-04-2022-0094

Al-Mamary, Y. H. S., Al-Nashmi, M. M., Shamsuddin, A., & Abdulrab, M. (2019). Devel- opment of an integrated model for successful adoption of management information systems in Yemeni telecommunica-

tion organizations. *International Journal of Scientific & Technology Research*, *8*(11), 39123939.

Al-Mamary, Y. H., Abubakar, A. A., & Abdulrab, M. (2023). The effects of the expectation confirmation model (ECM) and the technology acceptance model (TAM) on learning management systems (LMS) in sub-saharan Africa. *Interactive Learning Environments*, 1-17. doi: 10.1080/ 10494820.2023.2191272

Almaqtari, F. A., Farhan, N. H., Al-Hattami, H. M., & Elsheikh, T. (2023). The moderating role of information technology governance in the relationship between board characteristics and continuity management during the Covid-19 pandemic in an emerging economy. *Humanities and Social Sciences Communications*, *10*(1), 1-16.

Al-Okaily, M., Alghazzawi, R., Alkhwaldi, A. F., & Al-Okaily, A. (2022). The effect of digital accounting systems on the decision-making quality in the banking industry sector: a mediated-moderated model. *Global Knowledge, Memory and Communication*, *72*(8/9), 882-901. doi: 10.1108/GKMC-01-2022-0015

Alzahrani, A. I., Mahmud, I., Ramayah, T., Alfarraj, O., & Alalwan, N. (2019). Modelling digital library success using the DeLone and McLean information system success model. *Journal of Librarianship and Information Science*, *51*(2), 291-306.

Andoh-Baidoo, F. K. (2016). Organizational information and communication tech- nologies for development. *Information Technology for Development*, *22*(2), 193-204. doi: 10.1080/02681102.2016.1160703

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determi- nants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 101473.

ASI. (2013). *Conducting a study of micro, small, and medium enterprises in Yemen-Adam Smith International*. Retrieved from https://adamsmithinternational. com/projects/conducting-a-study-of-micro-small-and-medium-enterprises-in- yemen/#

Belfo, F., & Trigo, A. (2013). Accounting information systems: Tradition and future directions. *Procedia Technology*, *9*, 536-546.

Bhattacherjee, A. (2001) Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, *25*(3), 351-370.

Bhattacherjee, A., Perols, J., & Sanford, C. (2008). Information technology continuance: A theoretic extension and empirical test. *Journal of Computer Information Systems*, *49*(1), 17-26.

Charness, N., & Boot, W. R. (2016). Technology, gaming, and social networking. In K. Warner Schaie, S. L. Willis, *Handbook of the Psychology of Aging*, (pp. 389-407). Oxford, England: Academic Press.

Chen, J. V., Chung, H. Y., Widjaja, A. E., & Ha, Q. A. (2021). An Empirical Investigation of Users' Continuance Intention to Use Cloud Storage Service for Organi- zational Uses. *International Journal of Business Information Systems*, *45*(2). doi: 10.1504/IJBIS.2021.10038640

Cheng, Y-M. (2019). A hybrid model for exploring the antecedents of cloud ERP continuance: Roles of quality determinants and task-technology fit. *Inter- national Journal of Web Information Systems*, *15*(2), 215-235. doi: 10.1108/IJWIS-07-2018-0056

Cheng, Y-M. (2020). Understanding cloud ERP continuance intention and indivi- dual performance: a TTF-driven perspective. *Benchmarking: An International Journal*, *27*(4), 1591-1614. doi: 10.1108/BIJ-05-2019-0208

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295-336.

Chin, W. W. (2010). How to write up and report PLS analyses. In *Handbook of partial least squares* (pp. 655–690). Berlin, Germany: Springer.

Chin, W. W., & Newsted, P. R. (1999). Structural equation modeling analysis with small samples using partial least squares. *Statistical Strategies for Small Sample Research*, *2*, 307-342.

Chiu, W., Cho, H., & Chi, C. G. (2021). Consumers' continuance intention to use fitness and health apps: an integration of the expectation–confirmation model and investment model. *Information Technology & People*, *34*(3), 978-998. doi: 10.1108/ITP-09-2019-0463

Dalloul, M. H. M., Binti Ibrahim, Z., & Urus, S. T. (2023). The Impact of Quality Dimensions of Accounting Information System Success on the Effectiveness of During- Financial Crisis Management: The Mediating Role of System Usage in a Gov- ernment Sector Context. *Asian Economic and Financial Review*, *13*(1), 18-48.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, *19*(4), 9-30. doi: 10.1080/07421222.2003.11045748

Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfac- tion. *MIS quarterly*, *12*(2), 259-274.

Essam, M., Deif, M. A., Attar, H., Alrosan, A., Kanan, M. A., & Elgohary, K. (2024). Decoding Queries: An In-Depth Survey of Quality Techniques for Question Analysis in Arabic Question Answering Systems. *IEEE Access*, *12*, 135241-135264. doi: 10.1109/ACCESS.2024.3458466.

Fadelelmoula, A. A. (2018). The impacts of the quality dimensions of the ERP system on the realization of the fundamental business objectives and perceived usefulness. *International Journal of Enterprise Information Systems*, *14*(4), 89-107.

Floropoulos, J., Spathis, C., Halvatzis, D., & Tsipouridou, M. (2010). Measuring the success of the Greek taxation information system. *International Journal of Information Management*, *30*(1), 47-56. doi: 10.1016/j.ijinfomgt.2009. 03.013

Franque, F. B., Oliveira, T., & Tam, C. (2021). Understanding the factors of mobile payment continuance intention: empirical test in an African context. *Heliyon*, *7*(8).

Gefen, D., Straub, D., & Boudreau, M. C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, *4*(7), 1-77. doi: 10.17705/1CAIS. 00407

Grande, E. U., Estébanez, R. P., & Colomina, C. M. (2011). The impact of Accounting Information Systems (AIS) on performance measures: empirical evidence in Spanish SMEs. *The international Journal of Digital Accounting Research*, *11*(1), 25-43.

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572-2593.

Guo, H., Yang, Z., Huang, R., & Guo, A. (2020). The digitalization and public crisis responses of small and medium enterprises: Implications from a COVID-19 survey. *Frontiers of Business Research in China*, *14*, 1-25.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, USA: Sage.

Halilovic, S., & Cicic, M. (2013). Understanding determinants of information systems users' behaviour: A comparison of two models in the context of integrated accounting and budgeting software. *Behaviour & Information Technology*, *32*(12), 1280-1291.

Harris, E. (Ed.). (2018). *The Routledge Companion to Performance Management and Control*. New York, USA: Routledge

Hayat, A., Akhmad, B., Budiman A. (2020). Integrating people and technology in accrual accounting management to support quality financial reporting. *Polish Journal of Management Studies*, *22*(2), 158-172. doi: 10.17512/pjms.2020.22.2.11

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, *116*(1), 2-20. doi: 10.1108/IMDS-09-2015-0382

Hou, C. K. (2016). Understanding business intelligence system continuance intention: An empirical study of Taiwan's electronics industry. *Information Development*, *32*(5), 1359-1371.

Ismail, N. A., Abdullah, S. N., & Tayib, M. (2003). Computer-based accounting systems: The case of manufacturing-based small and medium enterprises in the northern region of peninsular Malaysia. *Jurnal Teknologi*, *39(*1), 19-36.

Jaklič, J., Grublješič, T., & Popovič, A. (2018). The role of compatibility in predicting business intelligence and analytics use intentions. *International Journal of Information Management*, *43*, 305-318.

Khayer, A., Bao, Y., & Nguyen, B. (2020). Understanding cloud computing success and its impact on firm performance: an integrated approach. *Industrial Management & Data Systems*, *120*(5), 963-985. doi: 10.1108/IMDS-06-2019-0327

Kumar, K. A., & Natarajan, S. (2020). An extension of the Expectation Confirmation Model (ECM) to study continuance behavior in using e-Health services. *Innovative Marketing*, *16*(2), 15-28.

Li, Y., & Wang, J. (2021). Evaluating the Impact of Information System Quality on Continuance Intention Toward Cloud Financial Information System. *Frontiers in Psychology*, *12*(12), 713353.

Mehta, N., Chauhan, S., & Kaur, I. (2022). Extending the story of IS success: A meta- analytic investigation of contingency factors at individual and organisational levels. *European Journal of Information Systems*, *31*(5), 617-640.

Mishra, A., Shukla, A., Rana, N. P., Currie, W. L., & Dwivedi, Y. K. (2023). Re-examining post- acceptance model of information systems continuance: A revised theoretical model using MASEM approach. *International Journal of Information Management*, *68*, 102571.

Mohammad, A. A. (2018). An exploration of accounting information system's role in SMEs failure triangular. *International Journal of Agile Systems and Management*, *11*(2), 155-178. doi: 10.1504/IJASM.2018.092546

Monteiro, A. P., Vale, J., Leite, E., Lis, M., & Kurowska-Pysz, J. (2022). The impact of information systems and non-financial information on company success. *International Journal of Accounting Information Systems*, *45*, 100557.

Nair, S., Radman, O., & Ahamad, S. (2020). The budgetary process and its effects on financial performance: a study on small and medium enterprises in Yemen. *International Journal of Innovation. Creativity and Change*, *14*(4), 816-834.

OCHA, (2021). *Yemen Socio-Economic Update*. Retrieved from https://reliefweb.int/report/yemen/yemen-socio-economic- update-issue-66-november-2021-enar

Petter, S., DeLone, W., & McLean, E. (2008). Measuring information systems success: models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, *17*(3), 236-263. doi: 10.1057/ejis.2008.15

Qutaishat, F., Abushakra, A., Anaya, L. & Al-Omari, M. (2023). Investigating the factors affecting the intention to adopt cloud-based ERP systems during the COVID-19 era: evidence from Jordan, *Business Process Management Journal*, *29*(3), 653-670. doi: 10.1108/BPMJ-09-2022-0462

Rahi, S., Alghizzawi, M., & Ngah, A. H. (2022). Factors influence user's intention to continue use of e-banking during COVID-19 pandemic: the nexus between self-determination and expectation confirmation model, *EuroMed Journal of Business*, *18*(3), 380-396. doi: 10.1108/EMJB-12-2021-0194

Ramayah, T., Ahmad, N. H., & Hong, T. S. (2012). An assessment of e-training effectiveness in multinational companies in Malaysia. *Journal of Educational Technology & Society*, *15*(2), 125-137.

Ramdani, B., Raja, S., & Kayumova, M. (2022). Digital innovation in SMEs: a systematic review, synthesis and research agenda. *Information Technology for Development*, *28*(1), 56-80.

Ritchi, H., Azis, Y., Adrianto, Z., Setiono, K., & Sanjaya, S. (2019). In-app controls for small business accounting information system: a study of domain understanding. *Journal of Small Business and Enterprise Development*, *27*(1), 31-51. doi: 10.1108/JSBED-12-2018-0372

Romney, M. B., & Steinbart, P. J. (2016). *Accounting information systems*. London, England: Pearson.

Rubin, A., & Babbie, E. R. (2016). *Empowerment series: Research methods for social work*. Boston, USA: Cengage Learning.

Rubino, M., & Vitolla, F. (2014). Internal control over financial reporting: opportunities using the COBIT framework. *Managerial Auditing Journal*, *29*(8), 736-771.

Saleh, M. A. K., & Manjunath, K. R. (2020). Small and medium enterprises in Yemen: Navigating

through additional obstacles during Covid-19. In *11th International Conference on Shifting Paradigm in Business, Economy and Society: Vision* (Vol. 2050, pp. 1-12).

Schöpfel, J., Azeroual, O., & Saake, G. (2019). Implementation and user acceptance of research information systems: An empirical survey of German universities and research organisations. *Data Technologies and Applications*, *54*(1), 1-15.

Sharma, S. K., & Sharma, M. (2019). Examining the role of trust and quality dimensions in the actual usage of mobile banking services: An empirical investigation. *International Journal of Information Management*, *44*, 65-75.

Suzianti, A., & Paramadini, S. A. (2021). Continuance intention of e-learning: The con- dition and its connection with open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(1), 97.

Teo, T. S., Srivastava, S. C., & Jiang, L. I. (2008). Trust and electronic government success: An empirical study. *Journal of management information systems*, *25*(3), 99-132.

Ujwary-Gil, A., & Florek-Paszkowska, A. (2025). Introduction to the Innovability Index: Beyond the fusion of innovation and sustainability. *Journal of Entrepreneurship, Management, and Innovation, 21*(2), 5-14. doi: 10.7341/20252121

Uzialko, A. (2022). *Choosing the Right Small Business Accounting Software*. Retrieved from https://www.businessnewsdaily.com/ 7542-choosing-accounting-software.html

Veeramootoo, N., Nunkoo, R., & Dwivedi, Y. K. (2018). What determines success of an e-government service? Validation of an integrative model of e-filing continuance usage. *Government Information Quarterly*, *35*(2), 161-174.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178.

Verkasalo, H. (2008). Dynamics of mobile service adoption. *International Journal of E-Business Research*, *4*(3), 40-63.

Williamson, K., & Johanson, G. (Eds.). (2017). *Research methods: Information, systems, and contexts*. Colchester, England: Chandos Publishing.

Xu, H., Horn Nord, J., Daryl Nord, G., & Lin, B. (2003). Key issues of accounting information quality management: Australian case studies. *Industrial Management & Data Systems*, *103*(7), 461-470. doi: 10.1108/02635570310489160

Yan, M., Filieri, R., & Gorton, M. (2021). Continuance intention of online technologies: A systematic literature review. *International Journal of Information Management*, *58*, 102315.

Volume 17 • Issue 2 • 2025

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 20 January 2025
accepted: 1 June 2025

pages: 90-105



# Multi-project buffer setting and dynamic monitoring of a critical chain based on comprehensive factors

Zhuzhen Bu
Hu Meng

ABSTRACT

The paper introduces a new multi-factor critical chain buffer estimation model and designs a dynamic monitoring method based on the project elements. A literature analysis determined a research gap and a research problem. It was found that the existing methods offer scarce collaborative studies on buffer setting and monitoring and insufficient research on buffer setting considering project economic indicators. However, these topics are often given priority consideration in practical engineering applications. Therefore, the study proposes a multi-factor critical chain buffer setting and its dynamic monitoring method. The planning stage analyses the impact of income, resources, and probability of success on buffer size setting and defines the calculation model of capacity constraint buffer. The execution stage dynamically sets buffer monitoring points according to the progress of project implementation, monitors the remaining buffer amount at the completion of each activity on the critical chain, and takes corresponding actions to ensure that the progress is controllable. The method was applied in a multi-project of a Chinese software enterprise. To further verify the effectiveness of this research, the method is compared with the traditional static buffer monitoring method (TBMM) and the relative buffer monitoring method (RBMM), and the construction period of the real project is simulated through the computer program for analysis. Results show that the research method can reduce unreasonable buffer settings, enhance the robustness of a buffer against complex environments, and reduce the probability of false warnings in the monitoring process.

KEY WORDS

**multi-project management, critical chain method, buffer setting, dynamic monitoring**

**10.2478/emj-2025-0014**

**Hu Meng**
College of Biomass Science and Engineering,
Sichuan University, 24 South Section 1, 1st Ring Road, Chengdu, China
ORCID 0000-0002-2145-5448

Corresponding author:
e-mail: menghu@scu.edu.cn

**Zhuzhen Bu**
Shanghai Yangtze Delta Innovation Institute,
14-15 Yunfei Building, No. 257 Xiangke Road, Pudong New Area, Shanghai, China
ORCID 0009-0003-4165-2443
e-mail: buzhuzhen@163.com

## Introduction

The developing economy produces a more competitive environment and greater uncertainty for the software industry. Many software enterprises encounter resource limitations and schedule control issues in multi-project implementation (An et al., 2024; Peng & Peng, 2022). The critical chain, a method developed from the constraint theory, has been regarded as an effective approach to solving these dilemmas (Tukel et al., 2006; Zhao et al., 2024). It focuses on the critical factors limiting the achievement of project objectives,

Bu, Z., & Meng, H. (2025). Multi-project buffer setting and dynamic monitoring of a critical chain based on comprehensive factors. *Engineering Management in Production and Services, 17*(2), 90-105. doi: 10.2478/emj-2025-0014

considers the relationship between activity constraints and resource constraints from the perspective of global optimisation, and finds the critical chain that determines the overall progress based on the discrimination principle that activities on the path have no room for delaying (Peng & Peng, 2022; Zhang & Han, 2023b). The buffer concept is introduced to reduce the impact of uncertain factors on the progress of project execution. By reducing the safety time and increasing the buffer, the critical chain method can shorten the project planning cycle, use the buffer to absorb the adverse impact of the delay on the overall schedule and improve the probability of completion on time (She et al., 2021; Zhang et al., 2018). A theoretical review of buffer management research shows two main directions, i.e., buffer settings and monitoring (Bredael & Vanhoucke, 2024; Zhang & Han, 2023a). The basic logic of using critical chain methods for multi-project management could also be summarised from previous literature (Zhang & Han, 2023b; Zhao et al., 2024).

The planning stage analyses the capacity constraint resource (CCR) of a multi-project to eliminate the resource conflict of the CCR and use heuristic algorithms (such as genetic algorithms etc.) to identify the system critical chain of the multi-project (Satic et al., 2024). At the same time, to avoid the delay of the drum activity of the former project affecting the start of the latter project, a capacity constraint buffer (CCB) is set for a multi-project system critical chain between the activity that uses the CCR last in the preceding project and the activity that uses the CCR first in the following project (Zhang & Han, 2023a). Here, a multi-project will be regarded as a single project with the critical activity chain as the critical chain (Bu et al., 2021). The execution phase analyses the buffer consumption based on the determined CCB to monitor whether multiple projects are delayed and corresponding measures are taken to ensure the timely completion of the projects (Trietsch, 2005).

Inspired by the literature and the concept of collaborative design (Dalouchei et al., 2022; Ghoddousi et al., 2017; Li et al., 2022; Luong Duc & Ohsato, 2008), this paper comprehensively considers factors such as project importance, resource scarcity, and risk resilience to set buffer size. Furthermore, real-time monitoring of the current active buffer consumption can be achieved by dynamically setting buffer monitoring points. Finally, combining case studies truly reflects the implementation of the proposed method, providing a basis for project schedule management decision-making and helping enterprises reduce the occurrence of erroneous warnings.

## 1. Literature review

Classic methods of buffer setting include the cut and paste method (C&PM) proposed by Goldratt and the root square error method (RSEM) proposed by Newbold in the 1990s (Bie et al., 2012). Although C&PM is simple to operate, the size of the buffer is directly related to the length of the critical chain. So, it is easy for C&PM to produce a buffer that is too large or too small; therefore, the method is rarely used in practice (Bakry et al., 2016). Contrastively, RSEM is based on the root of the sum of the variance of the safe time of all activities, which avoids arbitrarily reducing or increasing the estimated time of each activity and reduces subjective blindness (Hu et al., 2017).

On this basis, many scholars make improvements according to the project attributes (Goudarzi et al., 2024; Mozhdehi et al., 2024). For instance, Tukel et al. (2006) studied the impact of resource tension and network complexity on buffer determination and proposed an improved RSEM. Bie et al. (2012) analysed the influence of dependencies between activities on project duration and proposed a calculation method for a buffer size. Ghoddousi et al. (2017) studied indicators, such as project flexibility, complexity, robustness and dependency, and established a buffer size calculation method that considered internal and external risks. She et al. (2021) determined the buffer size based on the analysis of the structure of the project network. Some scholars also study project uncertainty. Luong Duc and Ohsato (2008) applied fuzzy theory to analyse the uncertainty of activity and project duration and converted it into a definite duration to ensure the reliability of the safe time estimation of each activity. Bakry et al. (2016) proposed an algorithm for optimal scheduling of repetitive construction projects under uncertain conditions and used fuzzy theory to model uncertainties related to various input parameters. Peng and Peng (2022) introduced risk integration shock rate to describe project process uncertainty and proposed a critical chain buffer setting method based on vulnerability theory. In addition, some scholars are concerned about the cost of project resources (Hu et al., 2017). Some scholars also used machine learning to study the method for setting a project buffer. Li et al. (2022) proposed a data-driven method for setting a buffer size, used full-factor experimental design and

Monte Carlo simulation to build data sets, trained the project buffer model through support vector regression, and constantly adjusted relevant parameters.

In terms of buffer monitoring, it usually divides buffers into three areas for monitoring (Kamandanipour et al., 2023; Zhao, Hu et al., 2024). When the buffer consumption is less than 1/3 of the total planned buffer, it indicates that the project activity is performing well (in the green area), and no action is required. When the buffer consumption is between 1/3 and 2/3 of the planned buffer total, it indicates possible problems in the project activities (in the yellow area), and it is necessary to identify the causes and take countermeasures. When the buffer consumption exceeds 2/3 of the planned buffer total (in the red area), it indicates that the project activity has a problem and immediate action is needed, such as rushing to work, adjusting the scope of the project and other measures to resolve the risk of delay, in which two trigger points (green zone and yellow zone, yellow zone and red zone demarcation points) are fixed (Zhang & Wan, 2019; Zhang & Wang, 2022). Though this static method (TBMM, traditional buffer management method) is easy to operate, it does not consider the possible increase of buffer consumption with the implementation of the project, which is a trigger to error warning (Zhang & Han, 2023a). Leach (2005) comprehensively considered the relationship between project implementation and buffer consumption and proposed the RBMM (relative buffer management method) but did not give the specific setting criteria for trigger points. Bei et al. (2010) proposed that the buffer required by the project should be adjusted in real-time according to the actual situation of the unfinished project. By dividing the remaining buffer at the current moment, dynamic monitoring of the buffer can be realised by setting dynamic monitoring points. On this basis, Bei et al. (2014) studied the impact of activities on buffering and further improved the dynamic buffering monitoring method based on activity-sensitive information. Zhang and Wan (2018) analysed the relationship between project uncertainty and activities and proposed that the buffer sizing should be allocated according to the risk exposure of activities, and the buffer monitoring margin of each monitoring point should be redistributed according to the risk weight factor of activities. In addition, there are also studies on building a dynamic buffer monitoring model based on project phase attributes (Zhang et al. 2018), activity reliability (Zhang et al. 2022), and other factors.

Summarising the above research results, some deficiencies were found in the existing methods. First, buffer management is a continuous work involving the planning stage to the execution stage; buffer setting and buffer monitoring correspond to the main tasks of these two stages, respectively. Most existing methods only focus on the research of one of the tasks, and there are few methods to design the coordination of the two from an overall perspective, which is not conducive to engineering practice applications. Secondly, the buffer setting of the existing methods takes economic indicators into less consideration, including project input cost, expected project return, etc. These factors imply the concept of more buffer protection for items with high economic indicators, and this concept is often given priority consideration in actual production. Based on this, this paper developed a new method for building a multi-factor critical chain buffer estimation model and designing a dynamic monitoring method based on project elements. This method improves the efficiency of buffer setting, ensures that high-yield projects obtain more buffer guarantees under the condition of limited resources, and reduces the occurrence probability of false warnings that often occur in traditional buffer monitoring. Specifically, it will include method analysis and case operation.

## 2. Research methods

### 2.1. Influential factors of a buffer

To ensure the stability of the multi-project schedule and improve the probability of completion on time, the critical chain method absorbs the uncertainties in the execution of activities through the use of buffering. When activities are affected by multiple uncertain factors. It is necessary to analyse the relationship between the multiple factors, establish a multi-factor critical chain buffer estimation model, and then obtain the project buffer sizing through the estimation of the activity buffer so as to calculate the overall buffer size of the multi-project. The project importance, resource tension, and on-time completion probability are discussed from the perspective of project management elements.

Project importance and its measurement coefficient

According to the definition of portfolio value management (Bu et al., 2021), all portfolios need to be managed to enhance and maintain organisational value. Value is a measure of the impact achieved by an entity or service.

To measure this value, refer to the definition of net income (profit) in economics.

First, consider the project in terms of net revenue, take the product of the project's expected return and project success probability minus the input cost as the net benefit. Then, divide the net benefit of a single project by the net benefit of the whole of multiple projects to obtain the final benefit, which is defined as the importance index *Qi* of project *i*. *Qi* reflects the importance of a single project *i* in a multi-project ensemble.

$$Q_i = \frac{E_i}{E} = \frac{V_i P_i - C_i}{VP - C}, i \in [1, N] \quad (1)$$

In Formula (1), a multi-project consists of *N* single projects. $E_i$、$V_i$、$P_i$ and $C_i$ represent the income, expected return, success probability, and input cost of project *i*, respectively. *E* represents the net benefit of all projects, *V* represents the expected return matrix, *P* represents the probability of success matrix, and *C* is the cost matrix (generally a constant matrix). There are correlations between projects in multiple projects. For example, $V_{12}$ represents the return from the joint implementation of Project *1* and Project *2*, $P_{12}$ represents the probability that Project *1* and Project *2* will succeed together depending on each other.

Second, the buffer sizing should meet the expectations of project value management. The measurement of returns is an influential factor in the buffer setting. For important projects with high returns, the appropriate buffer should increase to ensure its on-time completion principle and for general projects with lower returns, excess buffers should reduce. The importance measurement coefficient is defined as follows to reflect the importance difference between projects in the buffer setting.

$$\lambda_i = 1 + (Q_i - \overline{Q}) = 1 + (Q_i - \frac{\sum_{i=1}^{n} Q_i}{n}) \quad (2)$$

In Formula (2), $\overline{Q}$ means of the importance of items in multi-project. $\lambda_i$ as a product factor, when $Q_i$ exceeds $\overline{Q}$ and $\lambda_i$ is greater than 1, the buffer sizing is amplified. When $Q_i$ is less than $\overline{Q}$, the buffer sizing is compressed.

Resource tension coefficient

The intensity of active resource usage is another factor affecting buffer setting. On this basis, according to the characteristics of multi-project multi-resource constraints, the resource tension coefficient within the activity duration is defined as follows.

$$\alpha_{w_{(i,j)}} = Max \sum_{k=1}^{K} r_{w_{(i,j)},k} / R_{k,t}, \quad t \in d_{w_{(i,j)}} \quad (3)$$

In Formula (3), $w_{(i,j)}$ represents the activity *j* of project *i*, $r_{w_{(i,j)},k}$ represents the demand for resource *k* of activity *j* of project *i* within the duration $d_{w_{(i,j)}}$, $R_{k,t}$ represents the available use of resource *k* at time *t*, and *K* represents the set of all resources *k*.

In this step, the proportion of single resource demand of all activities in multiple projects in unit time is first judged, and then, the proportion of different resources is accumulated to find the sum of the proportion of demand of activities in unit time under multiple resources. Finally, the maximum value of the entire activity duration is taken and defined as the activity resource tension. The larger the $\alpha_{w_{(i,j)}}$, the more resources are required, and the smaller the value, the less resources are occupied. In addition, when designing multi-resource demand effects, the use of summation instead of product is to consider that the product may weaken the size of the value and does not reflect the characteristics that the buffer Settings should be increased when the tension of multiple resources is high.

On-time completion probability

The probability of completing each activity on time can be expressed through PERT's three-point estimation time, namely the most optimistic completion time, the most likely completion time and the most pessimistic completion time. The ratio of the difference between the most likely time and the most optimistic time to the difference between the most pessimistic time and the most optimistic time is defined as the elasticity coefficient as follows.

$$\beta_{w_{(i,j)}} = \frac{m_{w_{(i,j)}} - a_{w_{(i,j)}}}{b_{w_{(i,j)}} - a_{w_{(i,j)}}} \quad (4)$$

In Formula (4), $a_{w_{(i,j)}}$ is the most optimistic time of project *i* activity *j*, $m_{w_{(i,j)}}$ is the most likely time, and $b_{w_{(i,j)}}$ is the most pessimistic time. When $m_{w_{(i,j)}}$ approaches $a_{w_{(i,j)}}$, it means that the activity is more likely to be completed optimistically, the smaller $\beta_{w_{(i,j)}}$ is, the greater the probability that the activity will be completed on time. When $m_{w_{(i,j)}}$ approaches $b_{w_{(i,j)}}$, it means that the activity is more likely to be completed pessimistically, the greater the $\beta_{w_{(i,j)}}$ is, the smaller the probability that the activity will be completed on time.

### 2.2. MULTI-FACTOR BUFFER ESTIMATION MODEL

Based on the above analysis of income and the improvement of the root difference method, resource and on-time completion probability, a multi-factor capacity constraint buffer estimation model is proposed.

$$CCB = \left\{ \sum_{j \in CCB} \left[1 + \alpha_{w(i,j)}\right] \cdot \beta_{w(i,j)} \cdot \lambda_i \cdot (b_{w(i,j)} - m_{w(i,j)})^2 \right\}^{\frac{1}{2}} \quad (5)$$

$$CCB_i = \left[ \sum_{j \in CCB_i} m_{w(i,j)} / \sum_{i \in CCB} \sum_{j \in CCB_i} m_{w(i,j)} \right] \cdot CCB \quad (6)$$

In Formula (5) and Formula (6), $\alpha_{w(i,j)}$ and $\beta_{w(i,j)}$ are the resource tension and on-time completion of activity *j* of Project I, respectively, $\lambda_i$ is the measurement coefficient of importance of Project *i*, and activity *j* in all *i* is equal to this value. It can be seen that the improved *CCB* takes the activities involved on the critical chain of the system as the unit of calculation, and the final buffer size is obtained after summating and taking the root. *CCB* is the set of all activities on the critical chain of the system, and $CCB_i$ is the set of activities that use bottleneck resources from the first to the last of Project *i* on the critical chain of the system. Its size is assigned according to its proportion of the most likely completion time on the critical chain of the system.

### 2.3. BUFFER MONITORING POINT SETTING

After completing the buffer position setting, the monitoring study of the project execution began. First, the health of the overall progress at a given moment is determined by the consumption of the buffer at that moment.

$$BU_{j \in CC,t} = AT_{j,t} - PT_{j,t} = \sum_{k=1}^{j} (AT_{k,t} - PT_{k,t}) \quad (7)$$

In Formula (7), $AT_{j-1,t}$ is the actual cumulative construction period that has occurred before activity *j* (including activity *j*) on the critical chain of the system at time *t*, $PT_{j,t}$ is the planned cumulative construction period before activity *j* (including activity *j*) at time *t*, and the difference between the two is defined as the buffer consumption $BU_{j \in CC,t}$ before activity *j* (including activity *j*). If $BU_{j \in CC,t} > 0$, it indicates that the actual construction period is greater than the planned period, the progress is delayed, and the buffer is being consumed. If $BU_{j \in CC,t} \leq 0$, it indicates that the actual construction period is less than or equal to the planned construction period, the progress is ahead of schedule or normal, and the buffer is not consumed.

Second, compare the difference between the actual buffer consumption and the total buffer reserve, evaluate the buffer amount needed after the moment, and set the dynamic buffer monitoring point.

$$BR_{j \in CC,t} = PB_{j \in CC,t} - BU_{j \in CC,t} \quad (8)$$

In Formula (8), $BR_{j \in CC,t}$ is the remaining buffer after activity *j* at time *t*, and $PB_{j \in CC,t}$ is the total buffer reserve. When $BR_{j \in CC,t} \leq 0$, $BU_{j \in CC,t} \geq PB_{j \in CC,t}$, it means that buffer reserve has been exhausted and the project has been delayed. At this time, it is no longer meaningful to study the buffer monitoring settings, so this paper analyses the situation based on $BR_{j \in CC,t} > 0$, $BU_{j \in CC,t} < PB_{j \in CC,t}$. Refer to Goldratt's definition of buffer monitoring area division to divide the monitoring area. Since the value of buffer consumption $BU_{j \in CC,t}$ may be positive or negative, two cases are discussed. When $BU_{j \in CC,t} > 0$, $BR_{j \in CC,t} < PB_{j \in CC,t}$, the buffer is consumed, and the remaining buffer is the total buffer reserve minus the actual buffer consumption. Accordingly, dynamic monitoring point 1 and monitoring point 2 are set. When $BU_{j \in CC,t} \leq 0$, $BR_{j \in CC,t} \geq PB_{j \in CC,t}$, the buffer is not consumed, and the remaining buffer is the total buffer reserve. Dynamic monitoring point 1 and monitoring point 2 are set for this value. In addition, the absolute value of the actual buffer consumption, that is, the set of pre-completed quantities up to and including activity *j* at time *t*, can be regarded as an additional supplement to the total buffer reserve, the specific setting method of which is not discussed in this paper.

Based on the above analysis, the buffer monitoring point settings for activity *j+1* are defined as follows.

Monitoring point 1: $BM1_{j+1 \in CC,t} =$

$$= \begin{cases} BU_{j \in CC,t} + \frac{1}{3} BR_{j \in CC,t}, & BU_{j \in CC,t} \geq 0 \\ \frac{1}{3} PB_{j \in CC,t}, & BU_{j \in CC,t} < 0 \end{cases} \quad (9)$$

Monitoring point 2: $BM2_{j+1 \in CC,t} =$

$$= \begin{cases} BU_{j \in CC,t} + \frac{2}{3} BR_{j \in CC,t}, & BU_{j \in CC,t} \geq 0 \\ \frac{2}{3} PB_{j \in CC,t}, & BU_{j \in CC,t} < 0 \end{cases} \quad (10)$$

Third, according to the setting of dynamic monitoring points $BM1_{j+1 \in CC,t}$ and $BM2_{j+1 \in CC,t}$, classification processing is as follows. When $BU_{j+1 \in CC,t} < BM1_{j+1 \in CC,t}$, project execution is normal, and no action is required. When $BU_{j+1 \in CC,t} \geq BM1_{j+1 \in CC,t}$ and $BU_{j+1 \in CC,t} < BM2_{j+1 \in CC,t}$, there may be problems in project execution, and it is necessary to identify the causes and adopt countermeasures. When $BU_{j+1 \in CC,t} \geq BM2_{j+1 \in CC,t}$, project execution problems have occurred, and immediate action is needed to resolve the overall risk of delay by rushing work or adjusting the scope of the project.

Finally, based on the preceding monitoring point settings, define the dynamic monitoring procedure.

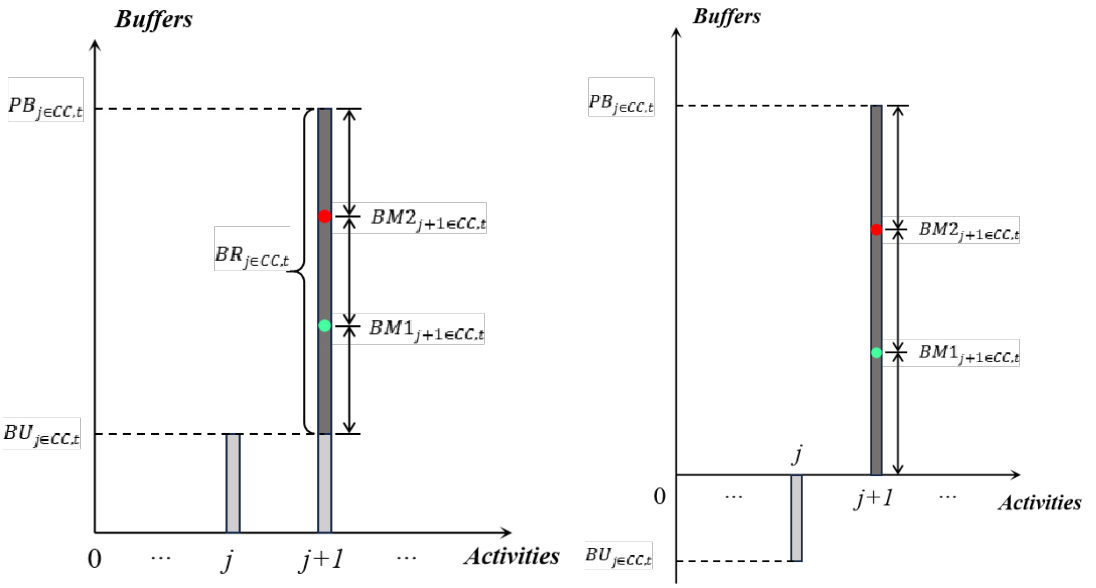



Fig. 1. Buffer monitoring point (left: when $BU_{j\in CC,t} > 0$; right: when $BU_{j\in CC,t} \leq 0$)

# 3. RESEARCH RESULTS AND DISCUSSION

## 3.1 BUFFER SIZING CALCULATION

A real case of a software development enterprise was used to verify the effectiveness of this paper. The company is a payment institution affiliated with a global top 500 financial group, focusing on internet payment, mobile phone payment, prepaid card issuance, and acceptance business in China. Its mobile payment transaction scale ranks among the top in the industry. In recent years, with the strong regularisation of China's financial industry and the intensification of competition within the industry, enterprises are facing a more severe market survival environment. Large internet software enterprises face challenges in internal project management. Among them, Internet IT projects have the characteristics of short duration, large quantity, and large scale (Bu et al., 2021). In the case of limited resources, they face the problem of resource conflicts among multiple projects. Moreover, project delays often occur due to an imperfect risk prevention mechanism and uncertain factors in the execution process.

In this context, three medium-sized projects carried out by the software development enterprise in the same period were selected as a verification case. The software development method adopts the "waterfall development" model. The cost input, expected return, and success probability of the project are shown in Table 1, and the working hours, activities relationships and resource requirements of each activity of the project are shown in Table 2. The resource constraints are mainly human constraints, with no more than eight development engineers and no more than four testing engineers. The critical chain of the system is 21-22-23-24-25-26-27-15-16-17-5-6-7-8-9.

According to Formulas (1) to (6), the capacity constraint buffer size $CCB_i$ of each project is obtained, as shown in Table 3.

By summating the $CCB_i$ values of each project, the capability constraint buffer value *CCB*=6.796 is obtained, and the total buffer reserve *PB*=6.796. Compared with the RSEM $CCB_{RSEM}$ =8.426, the buffer size is obviously reduced, and the overall schedule is shortened.

## 3.2. BUFFER POSITION SETTING

According to the multi-project critical chain buffer setting method, the capacity constraint buffer is set at the last bottleneck resource of Project C, Project B, and Project A, respectively. Considering that the actual software projects are set in a unit of person/day, this paper set $CCB_C$=3, $CCB_B$=2, $CCB_A$=2, and the total capacity buffer *CCB* = 7. Among them, TBMM monitoring point 1: 1/3·*PB* = 2.27, TBMM monitoring point 2: 2/3·*PB* = 4.53. The monitoring points of the RBMM are monitored by two linearly increasing trigger lines; refer to Bie's method (2010) for setting (Tab.4).

Tab. 1. Summary of case project cost returns

| Project | Cost $C_i$ | Return $V_i$ | Success PR $P_i$ | Effect $V, P$ | Income $E_i$ |
|---|---|---|---|---|---|
| A | 18 | 300 | 80% | $V_{A,B}$=200, $P_{A,B}$=0.9 | 222 |
| B | 32 | 800 | 70% | $V_{B,C}$=100, $P_{A,B}$=0.5 | 528 |
| C | 21 | 700 | 60% | | 399 |
| ALL | 71 | 1,800 | | | E=1,379 |

Notes: Total income $E$ consists of single item-income and inter-item related income. The units of Cost $C$, Return $V_i$ and Income $E_i$ are ten thousand yuan.

Tab. 2. Project activities hours and resources demand table (unit: person/day)

| Project | Activity | Immediate activity | Work hours | 50% hours | Optimistic hours | Coding engineer | Testing engineer |
|---|---|---|---|---|---|---|---|
| A | 0 requirement analysis | 1 | 3 | 2 | 1 | 2 | 0 |
| | 1 architecture design | 2 | 2 | 1 | 1 | 2 | 1 |
| | 2 code programming | 3 | 5 | 3 | 2 | 3 | 0 |
| | 3 self-testing | 4 | 3 | 2 | 1 | 2 | 0 |
| | 4 smoke-testing | 5 | 2 | 1 | 1 | 1 | 2 |
| | 5 system-testing | 6 | 8 | 5 | 3 | 1 | 3 |
| | 6 regression-testing | 7 | 9 | 5 | 4 | 1 | 2 |
| | 7 online release | 8 | 2 | 1 | 1 | 1 | 1 |
| | 8 whitelist verification | 9 | 3 | 2 | 1 | 0 | 1 |
| | 9 full open | 10 | 3 | 2 | 1 | 1 | 1 |
| B | 10 requirements analysis | 11 | 5 | 3 | 1 | 4 | 0 |
| | 11 architecture design | 12 | 2 | 1 | 1 | 2 | 0 |
| | 12 code programming | 13 | 4 | 2 | 1 | 4 | 0 |
| | 13 self-testing | 14 | 5 | 4 | 3 | 5 | 0 |
| | 14 smoke testing | 15 | 7 | 4 | 2 | 4 | 0 |
| | 15 system testing | 16 | 3 | 2 | 1 | 2 | 3 |
| | 16 regression testing | 17 | 10 | 7 | 5 | 2 | 4 |
| | 17 security scanning | 18 | 9 | 6 | 3 | 1 | 2 |
| | 18 online release | 19 | 4 | 3 | 2 | 2 | 1 |
| | 19 post-online verification | 20 | 3 | 2 | 1 | 2 | 1 |
| | 20 data statistics | 21 | 5 | 3 | 2 | 1 | 0 |
| C | 21 requirements analysis | 22 | 4 | 2 | 1 | 1 | 0 |
| | 22 architecture design | 23 | 2 | 1 | 1 | 2 | 1 |
| | 23 code programming | 24 | 3 | 2 | 1 | 4 | 0 |
| | 24 self-testing | 25 | 5 | 3 | 1 | 4 | 0 |
| | 25 system testing | 26 | 8 | 5 | 3 | 1 | 3 |
| | 26 regression testing | 27 | 6 | 4 | 2 | 1 | 2 |
| | 27 pressure testing | 28 | 4 | 3 | 2 | 2 | 2 |
| | 28 online release | 29 | 2 | 1 | 1 | 1 | 1 |
| | 29 technical support | / | 3 | 1 | 1 | 1 | 1 |

Notes: Considering the operability of 0.5 person/day in the project and the overall program programming, 0.5 person/day is calculated as 1 person/day, the size of the optimistic working hours comes from empirical assessment.

Tab. 3. Buffer estimation parameter calculation table (unit: person/day)

| *I* | *J* | | | | $R_1$ | $R_2$ | | | | *A* | *M* | *B* | | σ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0 | 0.161 | 0.277 | 0.884 | 2 | 0 | 0.250 | 0.000 | 0.250 | 1 | 2 | 3 | 0.500 | 1 | |
| | 1 | 0.161 | 0.277 | 0.884 | 2 | 1 | 0.250 | 0.250 | 0.500 | 1 | 1 | 2 | 0.000 | 1 | |
| | 2 | 0.161 | 0.277 | 0.884 | 3 | 0 | 0.375 | 0.000 | 0.375 | 2 | 3 | 5 | 0.333 | 2 | |
| | 3 | 0.161 | 0.277 | 0.884 | 2 | 0 | 0.250 | 0.000 | 0.250 | 1 | 2 | 3 | 0.500 | 1 | |
| | 4 | 0.161 | 0.277 | 0.884 | 1 | 2 | 0.125 | 0.500 | 0.625 | 1 | 1 | 2 | 0.000 | 1 | |
| | 5 | 0.161 | 0.277 | 0.884 | 1 | 3 | 0.125 | 0.750 | 0.875 | 3 | 5 | 8 | 0.400 | 3 | 2.039 ≈2.00 |
| | 6 | 0.161 | 0.277 | 0.884 | 1 | 2 | 0.125 | 0.500 | 0.625 | 4 | 5 | 9 | 0.200 | 4 | |
| | 7 | 0.161 | 0.277 | 0.884 | 1 | 1 | 0.125 | 0.250 | 0.375 | 1 | 1 | 2 | 0.000 | 1 | |
| | 8 | 0.161 | 0.277 | 0.884 | 0 | 1 | 0.000 | 0.250 | 0.250 | 1 | 2 | 3 | 0.500 | 1 | |
| | 9 | 0.161 | 0.277 | 0.884 | 1 | 1 | 0.125 | 0.250 | 0.375 | 1 | 2 | 3 | 0.500 | 1 | |
| B | 10 | 0.382 | 0.277 | 1.105 | 4 | 0 | 0.500 | 0.000 | 0.500 | 1 | 3 | 5 | 0.500 | 2 | |
| | 11 | 0.382 | 0.277 | 1.105 | 2 | 0 | 0.250 | 0.000 | 0.250 | 1 | 1 | 2 | 0.000 | 1 | |
| | 12 | 0.382 | 0.277 | 1.105 | 4 | 0 | 0.500 | 0.000 | 0.500 | 1 | 2 | 4 | 0.333 | 2 | |
| | 13 | 0.382 | 0.277 | 1.105 | 5 | 0 | 0.625 | 0.000 | 0.625 | 3 | 4 | 5 | 0.500 | 1 | |
| | 14 | 0.382 | 0.277 | 1.105 | 4 | 0 | 0.500 | 0.000 | 0.500 | 2 | 4 | 7 | 0.400 | 3 | |
| | 15 | 0.382 | 0.277 | 1.105 | 2 | 3 | 0.250 | 0.750 | 1.000 | 1 | 2 | 3 | 0.500 | 1 | 2.039 ≈2.00 |
| | 16 | 0.382 | 0.277 | 1.105 | 2 | 4 | 0.250 | 1.000 | 1.250 | 5 | 7 | 10 | 0.400 | 3 | |
| | 17 | 0.382 | 0.277 | 1.105 | 1 | 2 | 0.125 | 0.500 | 0.625 | 3 | 6 | 9 | 0.500 | 3 | |
| | 18 | 0.382 | 0.277 | 1.105 | 2 | 1 | 0.250 | 0.250 | 0.500 | 2 | 3 | 4 | 0.500 | 1 | |
| | 19 | 0.382 | 0.277 | 1.105 | 2 | 1 | 0.250 | 0.250 | 0.500 | 1 | 2 | 3 | 0.500 | 1 | |
| | 20 | 0.382 | 0.277 | 1.105 | 1 | 0 | 0.125 | 0.000 | 0.125 | 2 | 3 | 5 | 0.333 | 2 | |
| C | 21 | 0.289 | 0.277 | 1.012 | 1 | 0 | 0.125 | 0.000 | 0.125 | 1 | 2 | 4 | 0.333 | 2 | 2.718 ≈3.00 |
| | 22 | 0.289 | 0.277 | 1.012 | 2 | 1 | 0.250 | 0.250 | 0.500 | 1 | 1 | 2 | 0.000 | 1 | |
| | 23 | 0.289 | 0.277 | 1.012 | 4 | 0 | 0.500 | 0.000 | 0.500 | 1 | 2 | 3 | 0.500 | 1 | |
| | 24 | 0.289 | 0.277 | 1.012 | 4 | 0 | 0.500 | 0.000 | 0.500 | 1 | 3 | 5 | 0.500 | 2 | |
| | 25 | 0.289 | 0.277 | 1.012 | 1 | 3 | 0.125 | 0.750 | 0.875 | 3 | 5 | 8 | 0.400 | 3 | |
| | 26 | 0.289 | 0.277 | 1.012 | 1 | 2 | 0.125 | 0.500 | 0.625 | 2 | 4 | 6 | 0.500 | 2 | |
| | 27 | 0.289 | 0.277 | 1.012 | 2 | 2 | 0.250 | 0.500 | 0.750 | 2 | 3 | 4 | 0.500 | 1 | |
| | 28 | 0.289 | 0.277 | 1.012 | 1 | 1 | 0.125 | 0.250 | 0.375 | 1 | 1 | 2 | 0.000 | 1 | |
| | 29 | 0.289 | 0.277 | 1.012 | 1 | 1 | 0.125 | 0.250 | 0.375 | 1 | 1 | 3 | 0.000 | 2 | |

Notes: The columns in Table 3, from left to right: project number (*i*), activity number (*j*), Project average importance degree (), importance measurement coefficient ($Q_i$), number of coding engineer ($r_1$), number of testing engineer (r2), coding engineer tension (α1), testing engineer tension (α2), resource constraints coefficient (α), the most optimistic time (a), the most probable time (m), the most pessimistic time (b), on-time completion probability (β), *b* and *m* deviation (σ).

Tab. 4. RBMM buffer settings

| Completion ratio of system critical chain (%) | Proportion of buffer consumption | |
|---|---|---|
| | monitoring point 1 (%) | monitoring point 2 (%) |
| 0 | 15 | 30 |
| 100% | 75 | 90 |

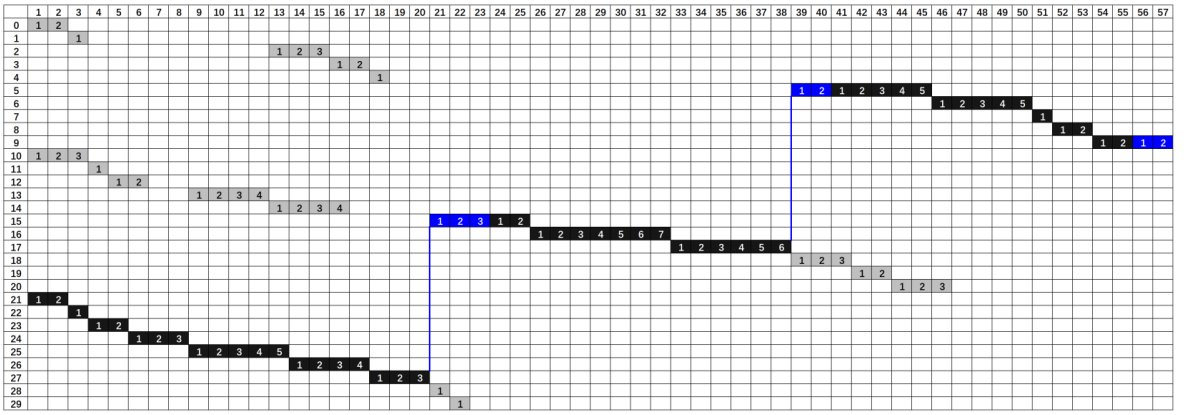


Fig. 2. Multi-project scheduling after adding capacity constraint buffer

### 3.3. Dynamic buffer monitoring

The simulated working hours of activities are randomly generated from the lognormal distribution (Tukel et al., 2006); the actual construction period can be simulated based on the "lognrnd()" function in MATLAB. To facilitate observation, the critical chain activities of the system were renumbered according to serial number 1 to serial number 15. Through computer programming simulation, TBMM, RBMM, and the method in this paper were respectively used for buffer monitoring. The following results were obtained in a single run.

In Figure 3, Figure 4, and Figure 5, the horizontal axis represents the activities on the critical chain of the multi-item system, and the vertical axis represents the capacity constraint buffer consumption. The blue dot represents the buffer consumption to simulate the actual activity. The green line represents the warning monitoring line 1 linked by the buffer monitoring point 1, beyond which the activity may be problematic. The red line represents the warning monitoring line 2 linked by buffer monitoring point 2, beyond which the activity has encountered problems.

By comparison, it can be found that, with the completion of activities, activity 3 to activity 15 in TBMM exceeded the warning monitoring line 1, with a proportion of more than 86 %, activity 2 to activity 7 in RBMM exceeded the warning monitoring line 1, with the proportion reaching 40 %. However, no situation beyond the monitoring line occurred in the method presented in this paper.

To verify the reliability of the above results further, the Monte Carlo technique was used to perform 1000 simulations, and TBMM, RBMM and the proposed method in this paper were analysed and compared.

Table 5, Table 6 and Table 7 show the numbers of the active buffer consumption values of the three methods that fall into the three monitoring areas. In Figure 6, Figure 7 and Figure 8, the buffer monitoring results of three methods are visually displayed in three different colours in terms of activity, buffer consumption, and execution times.

An analysis shows that the buffer consumption of TBMM does not fall into the red area in the early stage of the project, and the frequency of falling into the yellow area is also small, but the frequency of falling into the yellow area and the red area is constantly increasing in the later stage, and the frequency of falling into the yellow area and the red area reaches the highest in the last activity. This is inconsistent with the situation when the actual project goes to the later stage, and the uncertainty is smaller, as well as the warning frequency. It shows that the TBMM method does not consider the actual situation of project execution.

Compared with TBMM, the buffer warning frequency of RBMM is relatively small, but due to the linear characteristics of the relative buffer monitoring points, the lower monitoring points in the early stage lead to more warning frequency, showing a phenomenon of more before and less after. In actual work, there is generally no direct delay at the beginning of the project, and a large number of buffer warnings may cause wrong judgment of monitoring, which also indicates that the setting of RBMM monitoring points is too subjective.

The method proposed in this paper avoids TBMM and RBMM problems better, and the buffer warning frequency is the least. The buffer consumption of activities on the critical chain of the whole system is basically in the green area, and some activi-

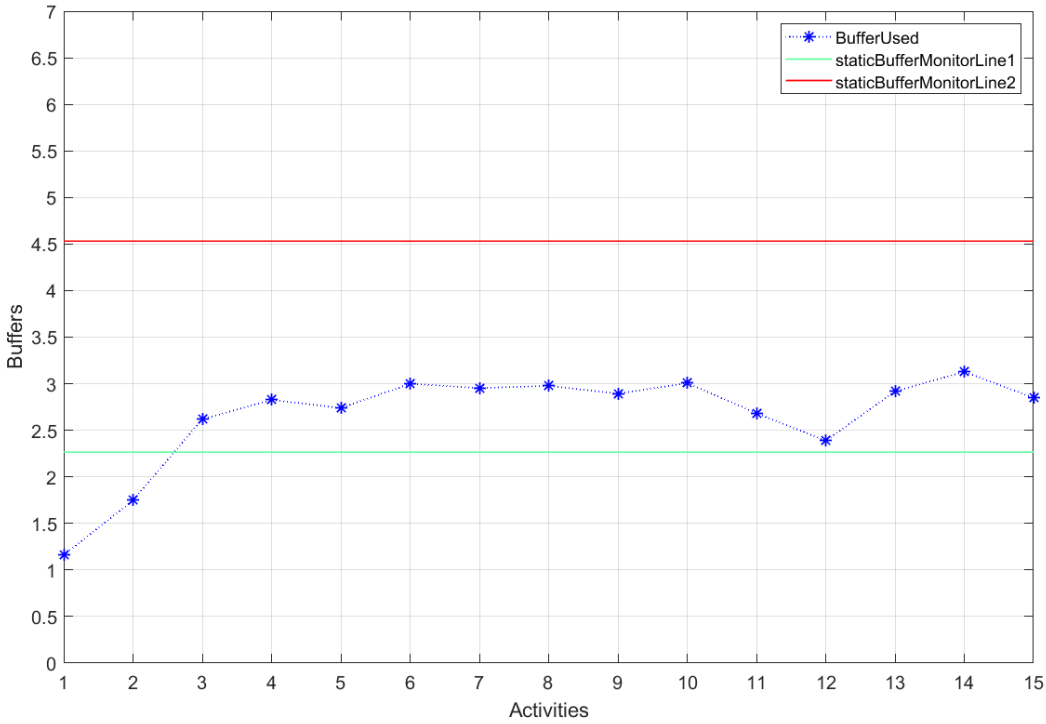



Fig. 3. TBMM method's buffer monitoring results (single time)

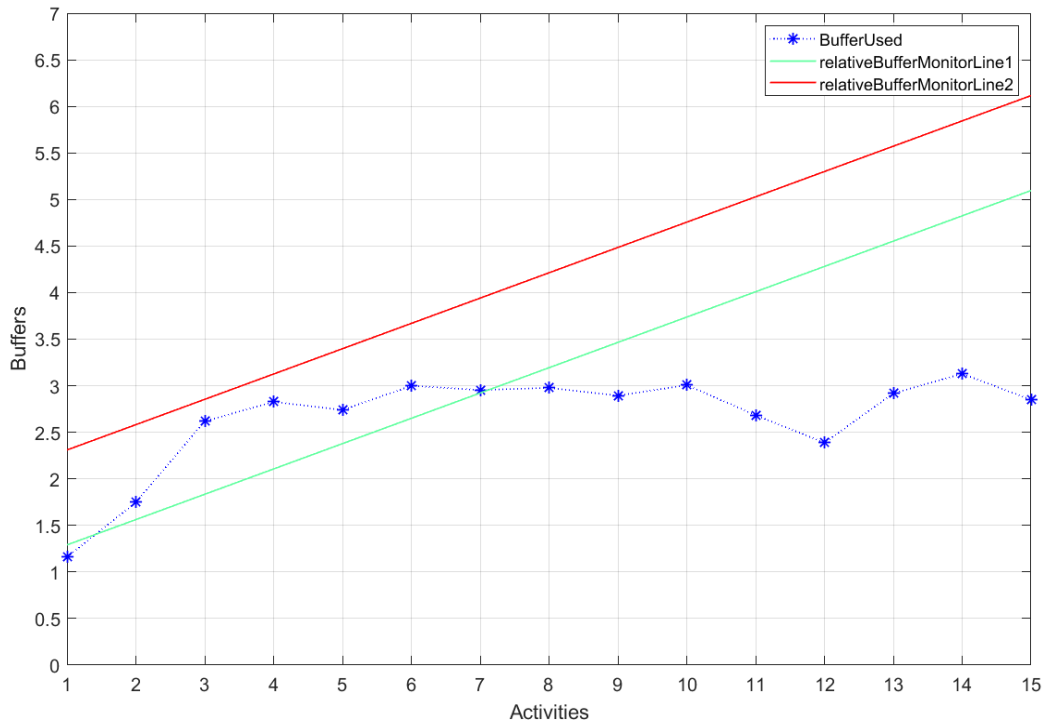



Fig. 4. RBMM method's buffer monitoring results (single time)

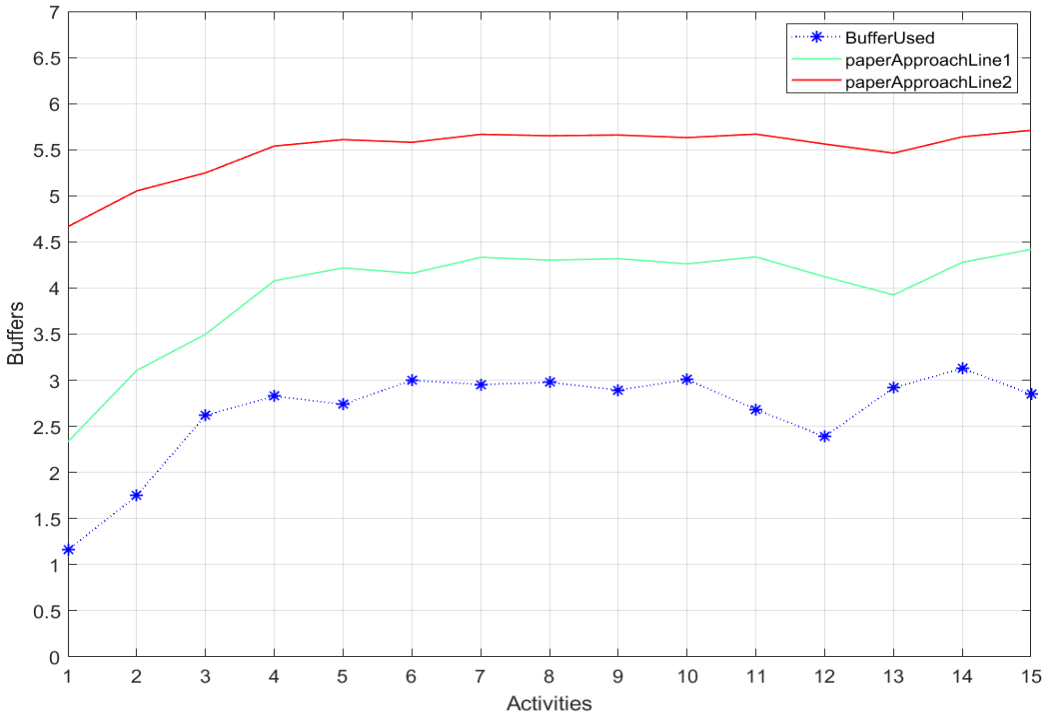



Fig. 5. Buffer monitoring results of the method described in this paper (single time)

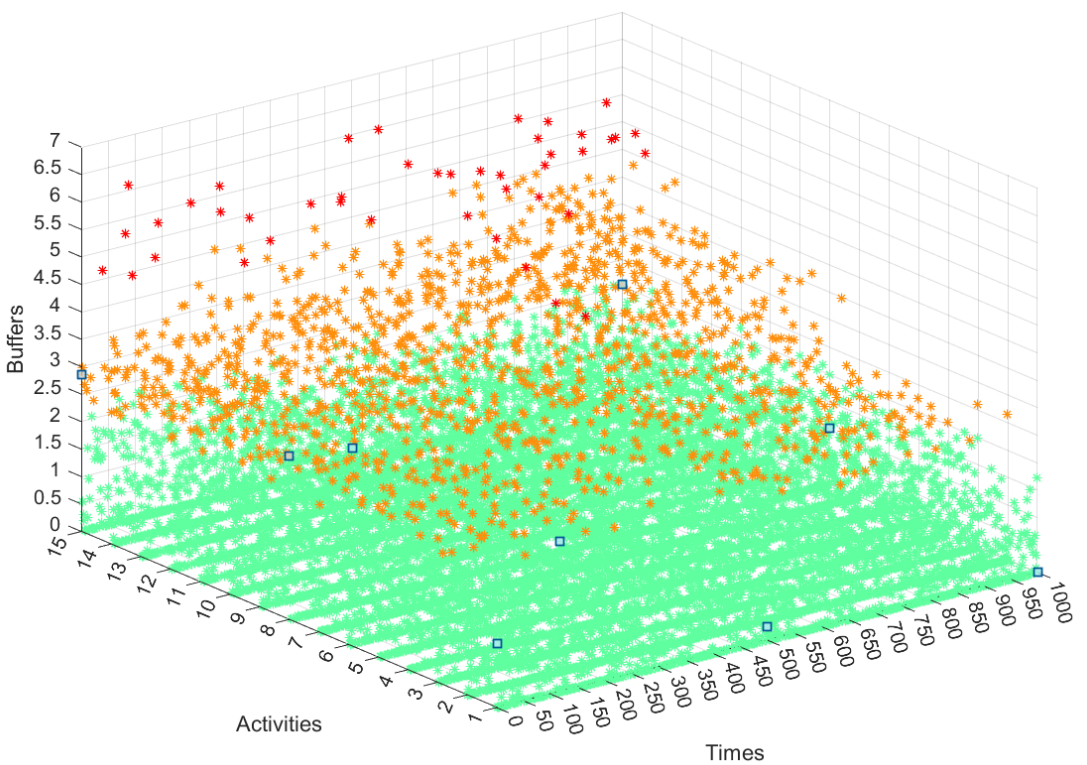



Fig. 6. TBMM method's buffer monitoring results (1000 times)

Tab. 5. TBMM method's regional distribution of the monitoring results (1000 times)

| Activities / Area | 1 (21) | 2 (22) | 3 (23) | 4 (24) | 5 (25) | 6 (26) | 7 (27) | 8 (15) | 9 (16) | 10 (17) | 11 (5) | 12 (6) | 13 (7) | 14 (8) | 15 (9) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Less than the monitoring line 1 (green area) | 997 | 972 | 951 | 944 | 948 | 947 | 915 | 922 | 916 | 909 | 896 | 893 | 885 | 830 | 820 |
| Greater than monitoring line 1 and smaller than monitoring line 2 (yellow area) | 3 | 27 | 49 | 56 | 52 | 53 | 84 | 77 | 83 | 90 | 103 | 105 | 113 | 154 | 162 |
| Greater than the monitoring line 2 (red area) | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 16 | 18 |

Tab. 6. RBMM method's regional distribution of the monitoring results (1000 times)

| Activities / Area | 1 (21) | 2 (22) | 3 (23) | 4 (24) | 5 (25) | 6 (26) | 7 (27) | 8 (15) | 9 (16) | 10 (17) | 11 (5) | 12 (6) | 13 (7) | 14 (8) | 15 (9) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Less than the monitoring line 1 (green area) | 921 | 911 | 915 | 932 | 957 | 976 | 964 | 974 | 980 | 987 | 987 | 995 | 989 | 990 | 993 |
| Greater than monitoring line 1 and smaller than monitoring line 2 (yellow area) | 76 | 75 | 75 | 61 | 41 | 23 | 32 | 22 | 19 | 12 | 12 | 4 | 10 | 9 | 7 |
| Greater than the monitoring line 2 (red area) | 3 | 14 | 10 | 0 | 2 | 1 | 4 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |

Tab. 7. Regional distribution of the monitoring results of the method described in this paper (1000 times)

| Activities / Area | 1 (21) | 2 (22) | 3 (23) | 4 (24) | 5 (25) | 6 (26) | 7 (27) | 8 (15) | 9 (16) | 10 (17) | 11 (5) | 12 (6) | 13 (7) | 14 (8) | 15 (9) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Less than the monitoring line 1 (green area) | 997 | 999 | 1000 | 1000 | 1000 | 1000 | 997 | 1000 | 1000 | 1000 | 998 | 999 | 1000 | 980 | 999 |
| Greater than monitoring line 1 and smaller than monitoring line 2 (yellow area) | 3 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 2 | 1 | 0 | 19 | 1 |
| Greater than the monitoring line 2 (red area) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

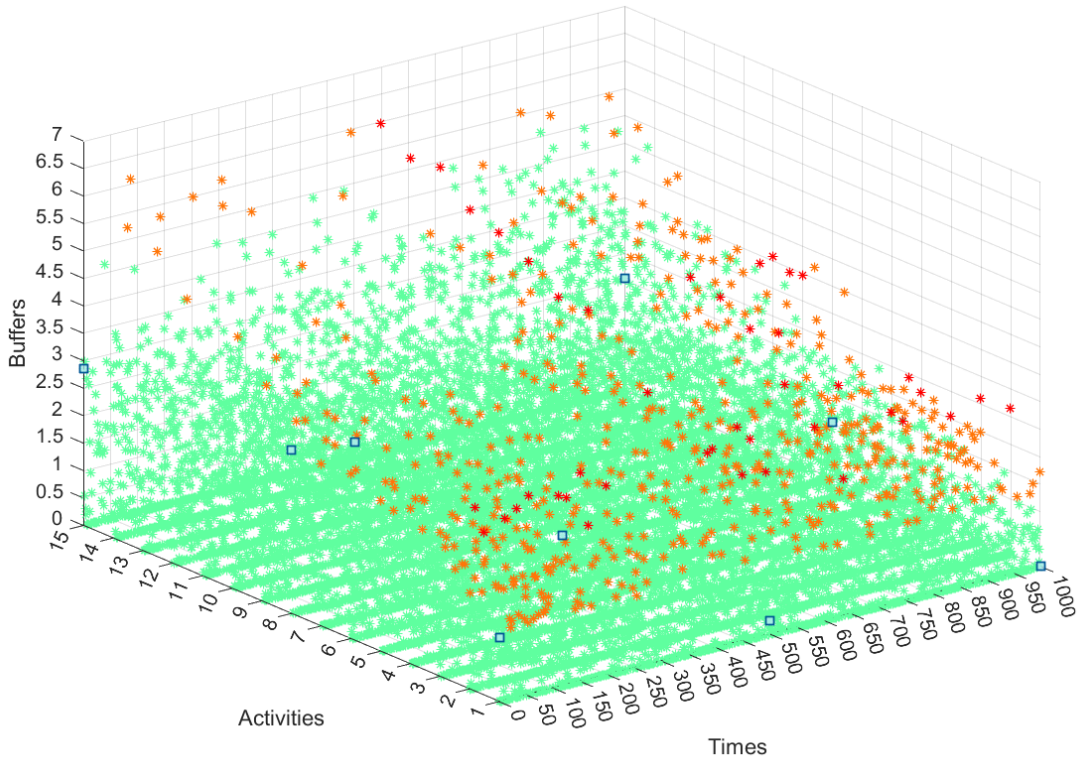



Fig. 7. RBMM method's buffer monitoring results (1000 times)

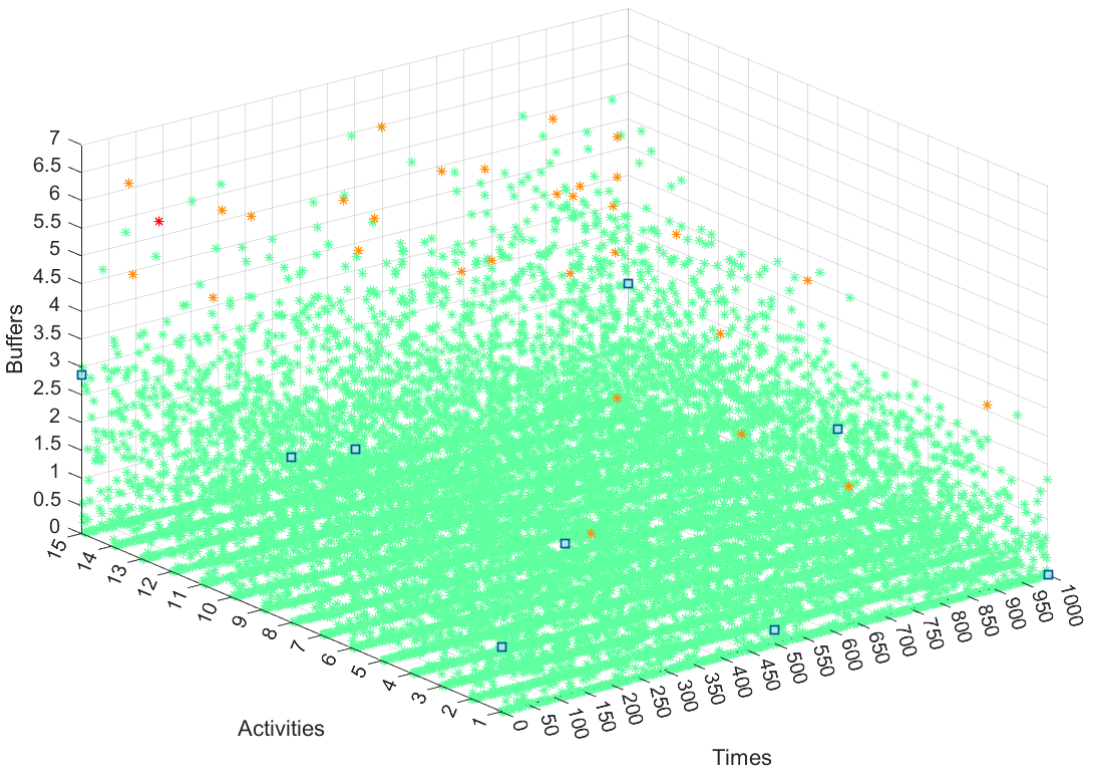



Fig. 8. Buffer monitoring results of the method described in this paper (1000 times)

ties in the yellow area are scattered, especially in the tail, due to the delay caused by uncertain factors, the buffer consumption of the last activity is processed and solved. The overall buffer monitoring cost is the best among the three methods, which obviously reduces the additional resource expenditure caused by the error buffer warning and ensures the timely completion of the multi-project execution.

## Conclusions

This paper proposes a critical chain buffer setting and dynamic monitoring method based on project elements. In the planning stage, the calculation model of the multi-project capacity constraint buffer is designed by analysing three factors: project income, resources, and on-time completion probability. In the execution stage, according to the dynamic change of buffer consumption in the process of multi-project implementation, dynamic buffer monitoring points are set to monitor the buffer surplus in real-time. This method improves the efficiency of buffer setting, ensures that high-yield projects can obtain more buffer protection under the condition of limited resources, and reduces the probability of false warnings, which often occur in traditional buffer monitoring.

Managerial implications: By studying the project schedule management problem of software development enterprises under resource constraints, this paper proposes a critical chain buffer estimation model considering the influence of multiple factors from the project elements and designs a dynamic buffer monitoring model. A financial payment enterprise is taken as a case to verify the effectiveness of this method. First, in the planning stage, project importance was defined according to the input and income of multiple projects, resource tension was designed based on the activity's bottleneck resource occupation, and on-time completion probability was designed based on the PERT estimation of the activity. The capacity constraint buffer estimation model of multiple projects was established by integrating three kinds of factors, and the buffer setting size and location were determined. Second, the buffer consumption is determined according to the current activity buffer consumption, the remaining buffer is calculated through the buffer reserve, and the buffer monitoring of the next activity is dynamically set by setting dynamic monitoring points on the remaining buffer so as to achieve rolling dynamic monitoring and warning. Finally, the proposed method in this paper is verified by computer simulation with a real case, and the results are compared with TBMM and RBMM to demonstrate the effectiveness and advantages of the proposed method.

The empirical results show that compared with the traditional critical chain management method, this paper considers the characteristics of IT multi-projects under resource constraints. In terms of buffer setting, it reflects the design intention of giving priority to high-yield and high-risk projects to obtain more buffer protection and effectively shorten the overall schedule and construction period. In terms of buffer monitoring, dynamic analysis of the current buffer consumption is used to set the next activity monitoring point, and real-time prediction and judgment of buffer monitoring are carried out, which effectively reduces the occurrence frequency of false buffer warnings, makes the division of buffer monitoring area more reasonable, and significantly improves the efficiency of buffer monitoring compared with traditional methods. This paper provides a reference for the research and practice of progress assurance and monitoring in multi-project parallel software enterprises.

Limitations and further research: Limitations of this paper are reflected in two aspects: case study and buffer monitoring point settings. In terms of the case research, three software projects carried out by the enterprise within the same time period were selected. The scale of the project portfolio was relatively small and could not fully reflect the current situation of the enterprise's project portfolio. Therefore, it is necessary to expand the scope of project portfolio management and increase the scale of the project portfolio in the next step to further improve the buffer settings and monitoring methods proposed in this paper. In terms of the buffer monitoring point settings, there is no further discussion of situations where buffer consumption may be negative (such as activities being completed ahead of schedule), which can be considered as an additional supplement to buffer reserves. The buffer can be called a "pseudo-buffer"; it can appropriately reduce the dependency of the delay on the buffer amount and should be considered in dynamic buffer monitoring. Therefore, the next step is to study the influence of "pseudo-buffer" on dynamic buffer monitoring on the existing basis.

## Literature

An, J., Li, S. C., & Wu, X. P. (2024). Multi-objective planning for time-cost trade-offs in multi-project parallel environment. *Engineering Construction and Architectural Management*, *32*(6), 4051-4073. doi: 10.1108/ecam-08-2023-0867

Bakry, I., Moselhi, O., & Zayed, T. (2016). Optimized scheduling and buffering of repetitive construction projects under uncertainty. *Engineering Construction and Architectural Management*, *23*(6), 782-800. doi: 10.1108/ecam-05-2014-0069

Bie, L., & Cui, N.. (2010). Research on dynamic buffer monitoring in critical chain project management. *Chinese Journal of Management Science*, *18*(6), 97-103. doi: 10.16381/j.cnki.issn1003-207x.2010.06.003

Bie, L., Cui, N. F., & Zhang, X. M. (2012). Buffer sizing approach with dependence assumption between activities in critical chain scheduling. *International Journal of Production Research*, *50*(24), 7343-7356. doi: 10.1080/00207543.2011.649096

Bie, L., Cui, N., & Tian, W. (2014). Research on activity sensitivity index based dynamic buffer monitoring method. *Chinese Journal of Management Science, 22*(10), 113-121. doi: 10.16381/j.cnki.issn1003-207x.2014.10.005

Bredael, D., & Vanhoucke, M. (2024). A genetic algorithm with resource buffers for the resource-constrained multi-project scheduling problem. *European Journal of Operational Research*, *315*(2), 19-34. doi: 10.1016/j.ejor.2023.11.009

Bu, Z., Liang, X., & Meng, H. (2021). Improved Method and Application of Critical Chain Identification Based on Multi-Project Important Degree. *Industrial Engineering and Management*, *26*(5), 187-194. doi: 10.19495/j.cnki.1007-5429.2021.05.023

Dalouchei, F., Mousavi, S. M., Antucheviciene, J., & Minaei, A. (2022). A Bi-Objective Model for Scheduling Construction Projects Using Critical Chain Method and Interval-Valued Fuzzy Sets. *Buildings*, *12*(7), Article 904. doi: 10.3390/buildings12070904

Ghoddousi, P., Ansari, R., & Makui, A. (2017). A Risk-Oriented Buffer Allocation Model Based on Critical Chain Project Management. *Ksce Journal of Civil Engineering*, *21*(5), 1536-1548. doi: 10.1007/s12205-016-0039-y

Goudarzi, E., Esmaeeli, H., Parsa, K., & Asadzadeh, S. (2024). Energy-efficient resource-constrained multi-project scheduling problem with generalized precedence relations and multi-skilled resources. *Journal of Supercomputing*, *80*(10), 13837-13872. doi: 10.1007/s11227-024-05933-0

Hu, X. J., Demeulemeester, E., Cui, N. F., Wang, J. J., & Tian, W. D. (2017). Improved critical chain buffer management framework considering resource costs and schedule stability. *Flexible Services and Manufacturing Journal*, *29*(2), 159-183. doi: 10.1007/s10696-016-9241-y

Kamandanipour, K., Tavakkoli-Moghaddam, R., & Yakhchali, S. H. (2023). A discrete time/resource trade-off problem with a critical chain method under uncertainty: a hybrid meta-heuristic algorithm. *Soft Computing*, *27*(23), 17867-17885. doi: 10.1007/s00500-023-09065-0

Leach, L. P. (2005). *Critical Chain Project Management*. London, England: Artech House.

Li, H. B., Cao, Y. W., Lin, Q., & Zhu, H. Y. (2022). Data-driven project buffer sizing in critical chains. *Automation in Construction*, *135*, 104134. doi: 10.1016/j.autcon.2022.104134

Luong Duc, L., & Ohsato, A. (2008). Fuzzy critical chain method for project scheduling under resource constraints and uncertainty. *International Journal of Project Management*, *26*(6), 688-698. doi: 10.1016/j.ijproman.2007.09.012

Mozhdehi, S., Baradaran, V., & Hosseinian, A. H. (2024). Multi-skilled resource-constrained multi-project scheduling problem with dexterity improvement of workforce. *Automation in Construction*, *162*, 105360. doi: 10.1016/j.autcon.2024.105360

Peng, J. L., & Peng, C. (2022). Buffer Sizing in Critical Chain Project Management by Brittle Risk Entropy. *Buildings*, *12*(9), 1390. doi: 10.3390/buildings12091390

Satic, U., Jacko, P., & Kirkbride, C. (2024). A simulation-based approximate dynamic programming approach to dynamic and stochastic resource-constrained multi-project scheduling problem. *European Journal of Operational Research*, *315*(2), 454-469. doi: 10.1016/j.ejor.2023.10.046

She, B. L., Chen, B., & Hall, N. G. (2021). Buffer sizing in critical chain project management by network decomposition *. *Omega-International Journal of Management Science*, *102*, 102382. doi: 10.1016/j.omega.2020.102382

Trietsch, D. (2005). The effect of systemic errors on optimal project buffers. *International Journal of Project Management*, *23*(4), 267-274. doi: 10.1016/j.ijproman.2004.12.004

Tukel, O. I., Rom, W. O., & Eksioglu, S. D. (2006). An investigation of buffer sizing techniques in critical chain scheduling. *European Journal of Operational Research*, *172*(2), 401-416. doi: 10.1016/j.ejor.2004.10.019

Zhang, J. G., & Han, Q. (2023a). Adaptive capacity constraint buffer monitoring of the multi-project system based on the relevance of drum activities. *Kybernetes*, *52*(10), 4668-4685. doi: 10.1108/k-11-2021-1147

Zhang, J. G., & Han, Q. (2023b). Buffer Monitoring of Critical Chain Projects Based on Support Vector Machine Prediction. *Ksce Journal of Civil Engineering*, *27*(7), 2745-2755. doi: 10.1007/s12205-023-0033-0

Zhang, J. G., Jia, S. K., & Diaz, E. (2018). Dynamic monitoring and control of a critical chain project based on phase buffer allocation. *Journal of the Operational Research Society*, *69*(12), 1966-1977. doi: 10.1080/01605682.2017.1415641

Zhang, J. G., Wan, D. (2018). Dynamic real-time monitoring and control of the critical chain project. *Chinese Journal of Management Science, 26*(4), 171-178. doi: 10.16381/j.cnki.issn1003-207x.2018.04.019

Zhang, J. G., & Wan, D. (2019). Integrated buffer monitoring and control based on grey neural network. *Journal of the Operational Research Society*, *70*(3), 516-529. doi: 10.1080/01605682.2018.1447251

Zhang, J. G., & Wang, M. H. (2022). Differential buffer monitoring for critical chain projects based on comprehensive activity reliability. *Journal of the Operational Research Society*, *74*(9), 2064-2079. doi: 10.1080/01605682.2022.2125845

Zhao, Y., Hu, X. J., Wang, J. J., & Cui, N. F. (2024). A robust multi-project scheduling problem under a resource dedication-transfer policy. *Annals of Operations Research*, *337*(1), 425-457. doi: 10.1007/s10479-024-05854-4

Zhao, Y., Jing, H. T., Yu, M., & Li, J. M. (2024). Efficient project management of historical building groups: a BIM-based approach with critical chain method and 5D simulation. *International Journal of Construction Management*, *25*(7), 826-839. doi: 10.1080/15623599.2024.2366711

Volume 17 • Issue 2 • 2025

sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 1 October 2024
accepted: 1 May 2025

pages: 106-116



# Expert-innovator behaviour questionnaire as a new tool for selecting potential experts

Anna M. Deptuła
Czesław S. Nosal

ABSTRACT

The effects of global change on development and the sustainability of the economy are visible in implementing innovations. Appropriate selection of experts, considering various knowledge areas, should use selected tools. The article introduces a research tool for selecting experts for innovation risk assessment. In particular, it aims to present individual measurement scales and their reliability assessment (Cronbach's alpha). The article presents the factor structure of a potential expert's competencies measured using a questionnaire. The questionnaire was developed for the appropriate selection of specialists from various industries, including production, mechanics, and management. The questionnaire constitutes a tool applicable for assessing and selecting people involved in the implementation and risk assessment of innovations. It is based on the following four scales (factors): open mind, closed mind, cognitive motivation, and response to uncertainty. The questionnaire's effectiveness was studied using two research samples, n= 224 and n= 349, comprised of entrepreneurs and individuals professionally related to business and development and implementation of innovations. The interviewees were selected randomly. Factor analysis was used to reduce the test items in the questionnaire. The internal reliability of items was analysed using Cronbach's alpha. The proposed questionnaire forms a new tool that can be used in selecting experts who deal with the risk assessment of innovations and in the broadly understood process of recruiting staff with appropriate competencies in terms of mindset characteristics. The article presents an analysis related to the conduct of typical research in Management and Quality Sciences, as well as practical principles guiding the use of the questionnaire, which may have wider application in the practice of risk management. The article presents the measurement tool with the answer key, which is a valuable guide for interpreting the results. The questionnaire facilitates the selection of individuals focused on independent and courageous problem-solving and the statement of evaluations. At the same time, an adequate level of caution should be respected, characterising people with risk aversion. Furthermore, creativity and openness are coupled with a considerable ability to develop new solutions and rationally respond to difficult and unpredictable situations.

KEY WORDS

**expert, innovation, creativity, risk, questionnaire, assessment, open mind, closed mind, cognitive motivation, response to uncertainty**

**10.2478/emj-2025-0015**

**Anna M. Deptuła**
Opole University of Technology,
Ozimska 75, 45-370, Opole, Poland
ORCID 0000-0002-2073-4971

Corresponding author:
e-mail: an.deptula@po.edu.pl

**Czesław S. Nosal**
SWPS University in Wroclaw,
Ostrowskiego 30b, 53-238, Wrocław, Poland
ORCID 0000-0002-1693-2692
e-mail: cnosal@swps.edu.pl

## Introduction

The need to develop a research tool in the form of a questionnaire was mainly due to the research gap that was perceived in this area and the experience of practitioners involved in project implementation in the enterprise. The key objective of this activity was to increase the effectiveness of the work carried out by specialists dealing with innovation, particularly those related to risk assessment. The guidelines related to the assessment procedure and, more broadly, to risk management became the determinant for the study of

Deptuła, A. M., & Nosal, C. S. (2025). Expert-innovator behaviour questionnaire as a new tool for selecting potential experts. *Engineering Management in Production and Services, 17*(2), 106-116. doi: 10.2478/emj-2025-0015

the new research tool. The right selection of experts is the key to effective risk analysis, devoid of overestimated and underestimated assessments (Deptuła, in preparation).

The expert-innovator behaviour questionnaire (EIBQ) forms a new tool that offers support in selecting potential experts for tasks related to creating and implementing innovations. Its structure is a response to the research needs related to the correct process of assessing the risk of innovation on a narrow scale of an enterprise and a wide scale of society. In the absence of available tools for diagnosing of experts' competencies (Deptuła & Nosal, 2021), corresponding to, among others, the purposeful selection of experts for the team that implements innovations, the authors developed a new measurement tool. The questionnaire facilitates the selection of individuals who fulfil the following criteria:

1. They are focused on independent and courageous problem-solving and the statement of evaluations. At the same time, they respect an adequate level of caution characterising people with risk aversion.
2. They are creative and open, coupled with a considerable ability to develop new solutions.
3. They can respond rationally to difficult and unpredictable situations.

The article aims to present a new tool for employee evaluation, EIBQ. The article proposes new scales, which, according to the authors, best describe selected characteristics of the expert mindset. The previous article (Deptuła & Nosal, 2021) presented a draft version of the questionnaire based on the three measured characteristics of the mindset and personality of the expert. In this article, the discussion focuses on factor analysis results based on which new measurement scales (sets of items) were selected. An answer key is also presented that can be applied to selecting potential experts. The paper presents the individual test items for measurement: open mind, closed mind, cognitive motivation, and response to uncertainty.

It also provides the proposed answer key for diagnosing the expert's state of mind.

It should be noted that EIBQ does not measure direct expert behaviour as the questionnaire is based on self-assessment.

## 1. Literature review

Choosing the right expert is a difficult and complicated task. Creating and implementing technological and organisational innovations is a task that requires the innovator to be open-minded and adapt quickly to changes that will undoubtedly take place in the future. Another important expert attribute is a balanced reaction to risk and estimation of uncertainty. In creating the new method for diagnosing expert attributes, the authors attempted to balance the psychological attributes with the diversity of tasks.

The expert-innovator behaviour questionnaire (EIBQ) proposed in this article is a new method attempting to consider the characteristics of the expert's mind. No such diagnostic tool was found in the literature. As a rule, the literature is dominated by efforts focused on the content of projected changes or narrow technological solutions. There is a distinct lack of methods for diagnosing the mind and personality traits of experts. Also, as a rule, these analyses used the Delphi method and did not consider the mind or personality traits of experts. Therefore, EIBQ fills an important gap in diagnosing the expert mind and personality.

The interdisciplinary nature of innovation projects means that the expert is required to have knowledge of, among other things, economics, management, design, marketing, environmental protection or legal regulations. Increasing the effectiveness of the work carried out by specialists in innovation, particularly related to risk assessment, may prove to be a recipe for many of the shortcomings and deficiencies faced by the company in adapting innovation to the requirements of society. The right selection of experts will play a key role here, and the effects will be seen in sustainable, socio-economic, scientific and integrated approaches to sustainable development. However, many efforts in the literature indicate the importance of this task (Ayag & Özdemir, 2007; Dey & Ogunlana, 2004; Dotgson, 2008; Kahraman et al., 2007; Lee et al., 2006; Manuel, 2007; Mobey & Parker 2002; Mullins & Sutherland, 1998; Nadkarni & Shenoy, 2001; Robin & Riedel, 1997; The Institute of Operational, 2012; Simon, 1992; Tversky & Kahneman, 1983; Ericsson, 1981; Ericsson, 1982; Larkin et al., 1980; Bouman, 1980; Reber et al., 2007; Ungson & Braunstein, 1984; Busenitz, 1999; Austin et al., 2007; Neubauer & Freudenthaler, 2005). An extensive group of methods is represented by behavioural models that use the knowledge and skills of the expert (Mobey & Parker, 2002; Mullins & Sutherland, 1998; Slovic, 2000).

Most often, these relate to describing the effectiveness of an expert's predictive judgments of a patient's condition (Sarbin, 1944; Dörfler & Ackermann, 2012). It turns out that learning about the cognitive qualities of expert minds lays the groundwork

for more effective and efficient decision-making (Dörfler & Ackermann, 2012; Petrides et al., 2022; Garb, 1989; Gustafson, 1963; Goldberg, 1970; Goldberg, 1992; Alegre, Pérez-Escoda, López-Cassá, 2019). However, these studies are rather selective and directed at a general overview of the expert's role in decision-making.

Contemporary support in this area is personality questionnaires. They are designed to assess a person's selected behaviour. Currently, tools, such as self-assessment questionnaires, are used where the individual answers questions about their behaviour, feelings, or thoughts (Credé et al., 2012). Most commonly, questionnaires address self-reported neuroticism, extraversion, cognitive motivation, novelty tolerance, and positive valuing of life (The Big Five questionnaires) and have dimensions to support psychological issues or the HR process more broadly (Gustafson, 1963; Goldberg, 1970; Goldberg, 1992; Alegre et al., 2019; Andrei, 2016). Available methods use various combinations of measurement of selected traits of both personality and human temperament (Catell, 1943; Catell et al., 1993; Cloninger et al., 1993; Costa & McCrae, 1992). However, the available questionnaires are not geared towards the purposeful selection of an expert with an emphasis on elements of a managerial and organisational nature. In practice, companies are looking for the right specialists to direct to specific tasks due to selected personality traits. Particularly important in this regard is selecting experts for innovation risk assessment.

In the literature, most of the methods used for risk assessment are expert-based. In particular methods of project management are currently used in the innovation risk assessment (Ayag & Özdemir, 2007; Dey & Ogunlana, 2004; Dotgson, 2008; Kahraman et al., 2007; Lee et al., 2006; Manuel, 2007; Mobey & Parker 2002; Mullins & Sutherland, 1998; Nadkarni & Shenoy, 2001; Robin & Riedel, 1997; The Institute of Operational, 2012). An extensive group of methods is represented by behavioural models that use the knowledge and skills of the expert (Mobey & Parker, 2002; Mullins & Sutherland, 1998; Slovic, 2000).

The authors also frequently use fuzzy logic for the risk assessment process (Deptula & Rudnik, 2017; Deptula & Rudnik, 2018; Choi & Ahn, 2010; Kahraman et al., 2007).

Differences between expert opinions significantly affect the outcome of the assessment; therefore, there is a need to continuously improve and objectify such methods (Choi & Ahn, 2010; Chin et al., 2009). Thus, the method of selecting experts gains importance. The selection of an expert should be based on certain rules. Personality questionnaires can be helpful in this regard. An example of a measurement tool characterising selected personality traits of an innovator is the expert-innovator behaviour questionnaire (EIBQ) presented in this article.

The starting point for developing the EIBQ questionnaire was the analysis of 54 statements, presented in detail by Deptuła and Nosal (2021). The statements describing the practically observable behaviours of innovators concerned three personality characteristics related to locus of control, motivation in activities, and the need for stimulation (Deptuła & Nosal, 2021; Deptuła & Nosal, 2021; Knosala & Deptuła, 2018; Larsen et al., 2013; Nosal, 1990, 1992, 2001; Rokeach, 2015; Rotter, 1966; Singer, 1975; Strelau, 2014; Zimmerman, 2000; Zuckerman, 2007).

## 2. Research methods

In the development of the version of the questionnaire presented in the article, the following research procedure was carried out several times:

1. Correlation coefficients within each variant (different variants were created by changing the number of statements in the measurement scales) of the questionnaire were analysed.

2. Factor analysis was conducted to extract measurement scales for assessing a person's predisposition to be an expert.

3. Basic statistics of the model with the extracted factors (measurement scales) were developed. Each of the extracted scales was developed according to the following assumptions:

- The measurement strongly impacts the decision-maker's preference for risk perception.
- The items address behaviours from everyday life and work.
- The questions provide a broad spectrum of research on the decision-maker's personality and cognitive abilities.

4. The reliability and relevance of the questionnaire were determined based on Cronbach's alpha.

Two research samples were used in the development of the questionnaire. The first was used to develop IBQ. The sample selection was random; however, the survey was conducted among entrepreneurs implementing innovations. The survey was conducted using paper questionnaires. The sample size, in this case, was 224. The survey was conducted in early 2020 in Poland. The first version of the ques-

tionnaire was used to conduct research to develop EIBQ. The research sample was purposive. The research was directed at entrepreneurs, innovators, designers, researchers, and people implementing innovations that appeared in publicly available reports and registers on innovation development in Poland. The research was funded by the research grant MINIATURA 4, 2020/04/X/HS4/00632. The survey was conducted in 2021 in Poland. As the study coincided with the period of the COVID-19 pandemic, in the implementation of the survey, an electronic version of the questionnaire was used.

The study of the EIBQ was carried out among 224 entrepreneurs and individuals professionally related to business, development, and implementation of innovations. The interviewees were selected randomly. Their number was determined by the minimum threshold resulting from the methodological rules for developing a personality questionnaire (Cohen et al., 1999; Crocker & Algina, 1986; Cronbach, 1990; Hornowska, 2007; Nunnally, 1978; Zakrzewska, 1994).

Factor analysis is a multivariate method used for data reduction purposes. It was originally developed by Spearman (1927) to analyse scores on mental tests and extract the general factor (factor g) as a structural basis of human intelligence. Based on a long history, the factor analysis methods are useful in a much wider range of situations. The basic idea is to represent a set of variables (matrix of correlations) with fewer new variables as vectors. In this solution, the vectors are called factors. These factors can be thought of as underlying correlations. Factor analysis is designed for interval data, although it can also be used for ordinal data (e.g., scores assigned to Likert scales) (Anderson, 2003; Manly & Navarro, 2016; Rencher, 2002).

In the studies, the following steps were taken to determine the final structure of the factors:

Tab. 1. Factor structure of items

| Factory structure of test items after rotation - main components | | | | | |
|---|---|---|---|---|---|
| **Lp** | Statement | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
| **1** | When buying new products, I choose well-known brands (companies) | | 0.492 | | |
| **2** | I like to accomplish new tasks as a means to learn something | | | 0.244 | |
| **3** | Too many changes make me feel uncomfortable | | | | 0.414 |
| **4** | In my life, I am guided by the advice/suggestions provided by my parents | | 0.426 | | |
| **5** | Sometimes I do not know how to act in many situations | | | | 0.477 |
| **6** | I like to deal with difficult issues | 0.563 | | | |
| **7** | Most success stories of companies are related to pure coincidence | | | -0.406 | |
| **8** | At school, I sought honours and good grades | | | 0.520 | |
| **9** | Rapid changes make me feel confused | | | | 0.539 |
| **10** | I feel anxious when a new employee appears in the company | | 0.406 | | |
| **11** | Existing work methods and techniques should not be changed | | 0.522 | | |
| **12** | Companies should offer products similar to their competitors | | 0.490 | | |
| **13** | Most of the company's successes are the result of its employees' activities | | | 0.114 | |
| **14** | In general, I am looking for tasks that I have never performed | 0.515 | | | |
| **15** | When I feel unsure of something, I do not undertake any action | | | | 0.558 |
| **16** | I like to follow/walk/run along unknown paths | 0.516 | | | |
| **17** | When I carry out tasks in a company, I like to focus on new solutions | 0.581 | | | |
| **18** | I cope well with uncertain situations | | | | -0.116 |
| **19** | I am happy to provide new solutions and ideas | 0.697 | | | |
| **20** | I like complicated tasks | 0.676 | | | |

1. The correlation coefficients between the 54 variables (IBQ items) were determined and analysed.
2. An exploratory factor analysis was performed for the matrix of all variables using the principal components method.
3. Values of factorial loadings for four factors were analysed. Values greater than 0.4 were qualified for the content analysis. Factor loadings prior to and following rotations were considered.
4. Names were specified to four main factors, which will further be referred to as scales for measuring an expert's mind competencies: open mind, closed mind, cognitive motivation, and reaction to uncertainty.

The coefficient of factor loads (for the four identified scales) is provided in Appendix 1. Table 1 shows the factor structure of the 20 items that make up the final draft of the questionnaire. The final version also had four control statements from the social desirability scale.

## 3. Research results

As a result of factor analysis, four factors characterising the competencies of the potential expert were identified. In total, these four factors all account for 13.26% of the identified variance. This is not a large range of explanations but indicates that the identified factors are specific in their nature. Further studies have been planned to extend (add to) this range with new factors.

The first factor, open mind (OM), forms the major factor, as it means directing the mind to new information and preference to coping with difficult tasks. However, other studies demonstrate that open-mindedness is positively correlated with intuition, imagination, and curiosity (Nosal, 2021). It is these qualities of mind that will be the focus of the projected research and factor analysis.

The second factor, defined as the closed mind (CLM), can be understood in a nearly literal sense. It means a mind that limits its field of exploration and new information. The problem of characteristics of such a mind has a long history dating back to Rokeach's pioneering research on dogmatism (1960). The dogmatic mind limits the scope of information search because it is thinking about formulated judgments based on narrow conceptual categories.

Factor three (cognitive motivation, CM) and factor four (response to uncertainty, RU) involve the attributes that characterise the stability of this type of motivation and behaviour under uncertainty. A favourable organisation of the mind should condition not only a constant tendency to seek information but also an appropriate attitude toward an emergence of uncertainty. This attitude is a kind of substitute for the ability to anticipate and not block the mind in the face of emerging changes and various forms of uncertainty in the environment. As emphasised by Prigogine (1996), we live in the conditions marked by uncertainty.

All the presented factors of the expert's mind form a rather clear structure. The justification for its feasibility is provided by the correlation matrix between the factors (Table 2).

Tab. 2. Matrix of correlation between factors

| | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| **Factor 1** | 1.00 | 0.04 | -0.42 | 0.02 |
| **Factor 2** | 0.04 | 1.00 | -0.24 | 0.14 |
| **Factor 3** | -0.42 | -0.24 | 1.00 | 0.14 |
| **Factor 4** | 0.02 | 0.14 | 0.14 | 1.00 |

Table 2 shows that the open mind is independent of the closed mind. These two types of minds and their associated cognitive orientation clearly oppose each other. In turn, the third factor (cognitive motivation) correlates negatively with the second factor (closed mind). Summarising the results of the factor analysis, it can be argued that competencies related to the open mind may constitute adequate criteria for the selection of potential experts. In turn, the indicator of mental closure can be used as a control since it is not beneficial for an expert to have this kind of mental orientation. The questionnaire also proposed a general simplified version of the control scale (SD).

The resulting factor structure was developed for a research sample equal to n=224. It concerned the analyses of the first version of IBQ presented in the article (Deptula & Nosal, 2021). The following section presents a comparison of the factor analysis results for the new research sample n=349 (EIBQ 2).

## 4. Discussion of the results

Item internal reliability EIBQ was analysed using Cronbach's alpha. Cronbach's alpha is certainly the most used statistic in the area of social sciences. It is a measure of reliability, specifically internal consistency reliability or item interrelatedness, of a scale or test (e.g., questionnaire). Internal consistency refers to

Tab. 3. Analyse Cronbach's alpha for two questionnaire versions

| IBQ* | Cronbach's alpha N=224 | EIBQ** | Cronbach's alpha n= 349 |
|---|---|---|---|
| UO<br>6 items | 0.7 | UO<br>4 items | 0.71 |
| UZ<br>5 items | 0.38 | UZ<br>6 items | 0.57 |
| MP<br>4 items | -0.06 | MP<br>4 items | 0.68 |
| RNNS<br>5 items | -.04 | RNNS<br>4 items | 0.77 |
| Number of items tested | 54* | | 41** |
| Number of items of the Q | 24 | | 22 |
| *The factor structure in Appendix 1 **The factor structure in Appendix 2 | | | |

the extent to which all items on a scale or test contribute positively towards measuring the same construct.

It should be noted that the test items presented in the article are still an object of research and will be subject to changes in subsequent versions of the questionnaire. Currently, two estimations of the reliability of the developed scales have been carried out for a sample equal to n=224 and n=49. The results show an upward trend for the index Alpha.

A summary of the results is presented in Table 3.

A further stage of the study will also be to conduct a CFA to determine the validity of the factorial structure of the questionnaire. However, this requires the collection of new data, which is currently underway. Nevertheless, the visible progress already suggests that the effectiveness of the developed scales is relatively high.

### 4.1. Brief characteristic of measurement scales in EIBQ — originality

Statistical analysis forms an important element of the assessment of the effectiveness and accuracy of each questionnaire; however, the knowledge of the scope of the presented topic and experience in conducting research is an equally important and significant element that determines the final draft of the research tool. Considering these two aspects means that the EIBQ presented in the work may be recommended for use as an assessment tool for people involved in the risk assessment of innovation and support in the recruitment process to the team of experts.

The statements in the first factor, open mind, characterise the desired expert mind orientation, so the more points a person scores on this scale, the higher their competencies as an expert.

The statements in the second factor, closed mind, indicate that if a person agrees with them, such individuals are challenging themselves with difficult, complicated and paradoxical experiences. Hence, the higher the score on this scale, the lower their competencies as experts. In such conditions, the selection of such a person should not be considered for tasks related to risk assessment and innovation implementation.

The statements in the third factor, cognitive motivation, are intended to determine the level of a person's motivation to create new solutions and ideas. People who need to create new solutions are very desirable for tasks where the scope of uncertainty and risk is high. The more points a person scores on this scale, the higher their competence as an expert.

The statements included in the reaction to uncertainty scale refer to the analysis of the familiar and commonly accepted attitudes towards risk: risk aversion, neutrality, and preference. Determining how a person reacts to uncertain situations may significantly limit the irrational decisions that bypass the substantive analysis of facts and data. The higher the result obtained by an employee on this scale, the more it proves the lack of resistance to action in unexpected

and uncertain situations, i.e., their poor adaptation to work in stressful conditions and a low level of competence as an expert.

In the final draft, the new questionnaire (EIBQ) contains 18 basic statements and four control statements from the social desirability scale. Apart from the many advantages of the questionnaire survey, there are some difficulties associated with the use of this survey form. The biggest problem is related to the degree of honesty (truthfulness) and reliability of the data derived from the questionnaire. Some respondents may consciously or unconsciously evaluate individual statements using conformist and schematic behaviour patterns and actions undertaken under the effect of generally accepted social norms and principles. Therefore, the use of the social desirability scale is widely known and accepted and recommended for this type of research tool.

It should also be emphasised that the independence of the created measurement scales allows them to be used as control scales. For example, the closed mind scale can be applied to successfully verify the reliability of the respondent's answers for the open mind scale. Moreover, the reaction to the uncertainty scale in relation to the cognitive motivation scale may have a similar application. The independence of the created scales can also be a good starting point for measuring the selected competencies of people applying for various positions in the company. Detailed principles for using the developed scales as control scales and the possibility of verifying the dependencies that can be established between them will be found in the reported research (Deptula, in preparation).

The factor analysis and the authors' experience in conducting research allow for proposing a presented version of EIBQ (Appendix 3).

### 4.2. Implications — the interpretation of questionnaire results

Completing the questionnaire, the respondents reply by expressing their opinions in support or rejection of a statement. A 5-point rating scale is employed to evaluate the statements: "I completely disagree", "I disagree", "It is difficult to say", "I agree", and "I completely agree", which correspond to a point scale from 1 to 5, respectively.

If an overall result equals 0, the respondent's competency as an expert can be assumed to be considerable. In the interpretation of the questionnaire results, it is assumed that the lower the result, the higher the competency. The overall result of the questionnaire or partial results for individual measurement scales can be interpreted (Table 4). The overall score for EIBQ is calculated according to the formula:

$$EIBQ = |-31+OM-CLM+CM-RU| \quad (1)$$

where:
OM — point value obtained on the open mind scale,
CLM — point value obtained on the closed mind scale,
CM — point value obtained on the cognitive motivation scale,
RU — point value obtained on the response to uncertainty scale.

The most suitable candidates for experts include candidates with a score in the range of 0 to 21. These people are creative and open-minded, and, at the same time, they can offer stability to the implementation of the innovation process while maintaining an appropriate level of caution. They are not excessively focused on generating risky activities, and, at the same time, they cope well with the fast pace of activity and accept the need for frequent divergence from a plan.

The individuals scoring between 22 and 41 points will be slightly worse at adjusting to frequent changes. Such individuals can display many contradictory features and strongly depend on specific events and situations.

The least desirable candidates include persons with scores above 42 points. In this situation, an individual will very often make decisions in a fearful or hasty manner and, thus, very often ill-considered. It is recommended that people with a score above 42 points should not be involved in innovative tasks.

The interpretation of the results of the questionnaire based on the given scale is only indicative. It was developed to evaluate expert applicants quickly. Detailed and unambiguous assessment requires experience and knowledge of the constructs described in the questionnaire. The authors encourage readers to carry out the questionnaire themselves and then interpret its result, and in case of doubts, to contact the authors to clarify any ambiguities. One should be aware that the interpretation of the questionnaire given in the article is simplified and does not consider the feedback between the personality traits tested in the questionnaire. Detailed rules for the interpretation of the results can be found in (Deptuła, in prepa-

ration). Nevertheless, the simplified interpretation of the EIBQ offers a tool for an accurate and independent characterisation of the candidate for an expert.

## Conclusions

The article proposes an expert-innovator behaviour questionnaire (EIBQ) to facilitate the effective process of selecting experts for projects related to implementing innovations.

Individual statements were selected so that it was possible to indicate individuals who are potentially the most suitable candidates for work on innovative projects. The practical use of the questionnaire will make it possible to eliminate people who may overestimate or underestimate the innovation risk assessment. In this respect, the results from the literature and this research concerning the features of the mind of an innovator's expert were extensively considered. On this basis, the following factors were considered to be the determinants of selecting an expert: a mental quality defined as an open mind, cognitive motivation, and a stable reaction to uncertain situations. For this purpose, potential innovators were tested, and the described version of the EIBQ was prepared using statistical and content analysis.

The questionnaire presented in the paper is a new tool that can be used to characterise experts-innovators, but also for the broadly understood process of recruiting employees. A broader profile of experts will be possible thanks to the IMQ Innovator Mind Questionnaire, which is in the development of o measuring the expert's intuition (IMQ characteristics will soon appear in subsequent papers).

The results presented in the article may constitute the basis for the development of research in the discipline of Management and Quality Sciences, as well as have an extensive practical application due to the advantages of the presented research tool.

The EIBQ questionnaire is considered a preliminary solution, although it meets basic psychometric standards such as consistency coefficients (Cronbach's alpha) and factor structure indices. The EIBQ requires further research and confirmation in different criterion groups. The authors also plan to include attrib-

Tab. 4. Interpretation of the EIBQ results

| Competence level | LOW | MEDIUM | HIGH |
|---|---|---|---|
| **OM**<br><br>**Items:**<br><br>8, 17, 19, 20 | **Score from 4 to 10 points**<br><br>Individuals scoring in this range are characterised by an open mind to a limited degree. When such a person is selected for the expert team, it is recommended to undertake decisions on the basis of the overall result of the questionnaire. | **Score from 11 to 15 points**<br><br>Individuals scoring in this range tend to vary in their beliefs, may be hesitant, and defining their mindsets as open may be ambiguous. In the case a high result is obtained on the control scale, the participation of such a person in the team implementing innovations should be considered. | **Score from 16 to 20 points**<br><br>Scores in this range demonstrate strong open-mindedness. In combination with a moderate result from the control scale, such candidates can be successfully selected for tasks related to the implementation of innovations.<br><br>An employee scoring in this range will be creative, curious and tolerant of discrepancies. |
| **CLM**<br><br>**Items:**<br><br>2, 5, 6, 13, 14, 15 | **Score from 30 to 20**<br><br>**points**<br><br>Individuals who score in this range are characterised by a closed mind. They do not tolerate unclear and uncertain situations and, at the same time, may demonstrate the characteristics of overly self-confident people, which, in the case of innovation, could limit its implementation. When selecting such a person for the expert team, it is recommended to make decisions based on the overall result of the questionnaire. | **Score from 13 to 19**<br><br>**points**<br><br>Individuals who score in this range are characterised by changing mindsets, may be indecisive, and the ability to define their mind as closed may be ambiguous. | **Score from 6 to 12**<br><br>**points**<br><br>People with scores in this range demonstrate little characteristics of a closed mind. In combination with a moderate result in the control scale, they can be successfully selected for tasks related to the implementation of innovations. |

Tab. 5. Interpretation of the EIBQ results - continued

| Competence level | LOW | MEDIUM | HIGH |
|---|---|---|---|
| **CM**<br>**Items:**<br>3, 10, 21, 22 | **Score from 4 to 10**<br>**points**<br>Individuals scoring the results in this range have a low need to learn new things. They prefer well-proven solutions and working methods and, therefore, may be reluctant to constant changes and improvements related to the development of innovation. | **Score from 11 to 15**<br>**points**<br>Individuals who score in this range may be hesitant or reluctant to provide the actual reasons for their behaviour. Their participation in the team of experts should be monitored and confronted with the needs of the project on an ongoing basis. In the case of obtaining a high result on the control scale, the participation of such a person in the team implementing innovations should be considered. | **Score from 16 to 20**<br>Individuals who achieve results in this range demonstrate a willingness to create new solutions. In combination with a moderate result in the control scale, they can be successfully selected for tasks related to the implementation of innovations. In conjunction with the low RNNS score, their participation in the team of experts is desirable. |
| **RU**<br>**Items:**<br>4, 9, 11, 18 | **Score from 16 to 20**<br>**points**<br>Individuals who score high on the scale of responding to uncertain situations could be susceptible to suspending its implementation; therefore, their participation in the expert team should be limited or strongly controlled. | **Score from 15 to 9**<br>**points**<br>Individuals who score in this range may be hesitant or reluctant to reveal the true reasons for their behaviour. Their involvement in the team of experts should be monitored and confronted with the current needs of the project. In the event of a high score on the control scale, the participation of such a person in the team should be considered. | **Score from 8 to 4**<br>**points**<br>Individuals who achieve scores in this range demonstrate resistance to new and difficult situations. They do not immerse themselves in an unforeseen situation and try to deal with it so they can be successfully selected for tasks related to the implementation of innovations. |
| **SD**<br>**Items:**<br>**1, 7, 12, 16** | **Score from 4 to 8**<br>**points**<br>A result in this range may mean that the employee will be pushing through their ideas with determination and will not be afraid to express their personal opinions. | **Score from 9 to 15**<br>**points**<br>The result in this range is difficult to interpret. It can indicate individuals who do not intend to stand out too much and choose safe answers, and thus, at work, will act in accordance with the "instructions" guidelines. | **Score from 16 to 20**<br>A score in this range means that the employee tried to fit into the generally accepted patterns and norms of a creative employee. Their answers might not be entirely dictated by the inner conviction of the will to get a good result. |
| **Overall test result** | **Score from 42 to 64**<br>Scorers in this range are individuals who like to be supervised. They work better when they know exactly what to do and their participation in the project implementation is limited to following orders. Their participation in the team of experts implementing innovations is not excluded; however, it needs to be adequately planned. | **Score from 22 to 41 points**<br>Individuals who score in this range may be creative and inventive; however, they also show many contradictory features and are strongly dependent on specific circumstances and situations. They may also tend to fit in with generally accepted standards and recommendations. Their participation in the team of experts, however, will not impede the potential innovation development opportunities. | **Score from 0 to 21 points**<br>The individuals who achieve a result in this range are most likely creative and outgoing individuals. Their participation in a team of experts is very welcome. |

utes of the mind such as intuition and critical thinking. It is understood that a method of this kind requires constant changes in language and terminology. Regardless of the practical (diagnostic) purpose of EIBQ, we view the method as a step on the road to better understanding the characteristics of the mind of experts.

## Literature

Alegre, A., Pérez-Escoda, N., & López-Cassá, E. (2019). The relationship between trait emotional intelligence and personality. Is trait EI really anchored within the Big Five, big two and big one frameworks? *Frontiers in Psychology*, *10*, 866. doi: 10.3389/fpsyg.2019.00866

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis. Third Edition*. New York, USA: A John Wiley & Sons.

Andrei, F., Siegling, A. B., Aloe, A. M., Baldaro, B., & Petrides, K. V. (2016). The Incremental validity of the Trait Emotional Intelligence Questionnaire (TEIQue): a systematic review and meta-Analysis. *Journal of personality assessment*, *98*, 261-276. doi: 10.1080/00223891.2015.1084630

Austin, E. J., Farrelly, D., Black, C., & Moore, H. (2007). Emotional intelligence, Machiavellianism and emotional manipulation: Does EI have a dark side? *Personality and Individual Differences*, *43*, 179-189. doi: 10.1016/j.paid.2006.11.019

Ayağ, Z., & Özdemir, R. G. (2007). An Analytic Network Process-Based Approach to Concept Evaluation in a New Product Development Environment. *Journal of Engineering Design*, *18*, 209-226. doi: 10.1080/09544820600752740

Bouman, M. J. (1980). Application of information-processing and decision-making research. In G. R. Ungson & D. N. Braunstein (eds.), *Decision making: An interdisciplinary inquiry* (pp. 129-167). Boston, USA: Kent Publishing.

Busenitz, L. W. (1999). Entrepreneurial Risk and Strategic Decision Making. *The Journal of Applied Behavioral Science*, *35*, 325-340.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The journal of abnormal and social psychology*, *38*, 476-506.

Cattell, R. B., Cattell, A. K., & Cattell, H. E. P. (1993). The sixteen personality factor fifth edition questionnaire. *The SAGE Handbook of Personality Theory and Assessment: Personality Measurement and Testing*, *2*(2), 135.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81.

Chin, K. S., Tang, D. W., Yang, J. B., Wong, S. Y., & Wang, H. (2009). Assessing New Product Development Project Risk by Bayesian Network with a Systematic Probability Generation Methodology. *Expert Systems with Applications*, *36*(6), 9879-9890. doi: 10.1016/j.eswa.2009.02.019

Choi, H. G., & Ahn, J. (2010). Risk Analysis Models and Risk Degree Determination in New Product Development: A Case Study. *Journal of Engineering and Technology Management*, *27*(1–2), 110–124. doi: 10.1016/j.jengtecman.2010.03.006

Cloninger, R. C., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of general psychiatry*, *50*(12), 975-990.

Cohen, R. J., Swerdlik, W. J., & Tobin, R. M. (2021). *Psychological testing and assessment: An introduction to tests and measurement (10 th ed.)*. New York, USA: McGraw-Hill Education.

Costa, P. T., & McCrae, R. R. (1992). The NEO Personality Inventory Manual Revised. Odessa, USA: Psychological Assessment Resources.

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology*, *102*(4), 874-88. doi: 10.1037/a0027403

Crocker, L., & Algina J. (1986). *Introduction to classical and modern test theory*. New York, USA: Harcourt Brace Jovanovich.

Cronbach, L. J. (1990). *Essentials of Psychological Testing (5th ed.)*. New York, USA: Harber Collins Publisher.

Deptuła, A. M., & Nosal C. S. (2021). Innovator Behavior Questionnaire as an Expert Selecting for Technical Innovation Risk Assessment. *European Research Studies Journal*, *14*(1), 107-119. doi:10.35808/ersj/2032

Deptuła, A. M., & Knosala R. (2015). Risk Assessment of the Innovative Projects Implementation. *Management and Production Engineering Review*, *6*, 15-25. doi: 10.1515/mper-2015-0032

Deptuła, A. M. (In preparation). *Criteria of expert selection for the risk management process in an enterprise. Supporting tools - Method of selection - Research results*. Oficyna Wydawnicza Politechniki Opolskiej.

Deptuła, A. M., & Rudnik, K. (2017). Risk Assessment of Innovative Project Using Fuzzy Inference System with Expert's Psychological Conditions. In: *Proceedings of the 29th International Business Information Management Association (IBIMA)*, p. 3093-3105. Vienna, Austria.

Deptuła, A. M., & Rudnik, K. (2018). Fuzzy Approach Using Experts' Psychological Conditions to Estimate the Criteria Importance for the Assessment of Innovative Projects Risk. *Management and Production Engineering Review*, *9*(1), 13-23. doi: 10.24425/119396

Dey, P. K., & Ogunlana, S. O. (2004). Selection and Application of Risk Management Tools and Techniques for Build-Operate-Transfer Projects. *Industrial Management and Data Systems*, *104*(3), 334-346. doi: 10.1108/02635570410530748

Dörfler, V., & Ackermann, F. (2012). Understanding intuition: The case for two forms of intuition. *Management Learning*, *43*(5), 545-564. doi: 10.1177/1350507611434686

Dotgson, M., Gann, D. M., & Salter, A. (2008). *The Management of Technological Innovation*. Oxford, Great Britain: Oxford University Press.

Ericsson, K. A., & Chase, W. G. (1981). Exceptional memory. *American Scientist*, *70*(6), 607-615.

Ericsson, K. A., & Polson, P. G. (1988). An experimental analysis of the mechanisms of a memory skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 305-316.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin*, *105*, 387-396.

Goldberg, L. R. (1970). Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin*, *73*, 422-432.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26-42. doi: 10.1037/1040-3590.4.1.26

Gustafson, J. E. (1963). The computer for use in private practice. In: *Proceedings of Fifth IBM Medical Symposium*, p. 101-111. White Plains, USA: mM Technical Publication Division.

Hornowska, E. (2007). *Testy psychologiczne. Teoria i praktyka* [*Psychological tests. Theory and practice*]. Warsaw, Poland: SCHOLAR.

Kahraman, C., Büyüközkan G., & Ateş N. Y. A. (2007). Two Phase Multi-Attribute Decision-Making Approach for New Product Introduction. *Information Sciences*, *177*, 1567-1582. doi: 10.1016/j.ins.2006.09.008

Knosala, R., & Deptuła, A. M. (2018). *Ocena ryzyka wdrażania innowacji* [*Risk assessment of implementing innovations*]. Warsaw, Poland: Polskie Wydawnictwo Ekonomiczne.

Larkin, J., McDennott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335-1342.

Larsen, R., & Buss, D. (2013). *Personality Psychology: Domains of Knowledge About Human Nature*. London, England: McGraw –Hill, Higer Education.

Lee, S. F., Tsai Y. C., & Jih W. J. (2006). An Empirical Examination of Customer Perceptions of Mobile Advertising. *Information Resources Management Journal*, *19*, 39-55. doi: 10.4018/irmj.2006100103

Manly, B. F. J., Jorge, A., & Navarro, A. (2016). *Multivariate statistical methods: a primer*. Boca Raton, USA: Chapman and Hall/CRC.

Manuel, E. (2007). Innovation and Risk Management. *Munich Personal RePEe Archive*, *2277*. doi: 10.2139/ssrn.969028

Mobey, A., & Parker, D. (2002). Risk Evaluation and Its Importance to Project Implementation. *Work Study*, *51*, 202-208. doi: 10.1108/00438020210430760

Mullins, J., & Sutherland, D. J. (1998). New Product Development in Rapidly Changing Markets: An Exploratory Study. *Journal of Product Innovation Management*, *15*(3), 224-236. doi: 10.1016/S0737-6782(97)00081-7

Nadkarni, S., & Shenoy, P. P. (2001). Bayesian Network Approach to Making Inferences in Causal Maps. *European Journal of Operational Research*, *128*, 479-498. doi: 10.1016/S0377-2217(99)00368-9

Neubauer, A. C., & Freudenthaler, H. H. (2005). Models of emotional intelligence. In Schultz R and Roberts RD (eds.), *Emotional intelligence: An international handbook*, 31-50.

Nosal, C. S. (1990). *Psychologiczne modele umysłu* [*Psychological models of mind*]. Warszawa, Polska: Państwowe Wydawnictwo Naukowe.

Nosal, C. S. (2001). *Psychologia Myślenia i Działania Menedżera: Rozwiązywanie Problemów, Podejmowanie Decyzji, Kreowanie Strategii* [*Psychology of Managerial Thinking and Acting: Problem Solving, Decision Making, and Strategy Formation*]. Wrocław, Poland: Akade.

Nosal, C. S. (1992). *Diagnoza typów umysłu* [*Diagnosis of mind type*]. Warszawa, Poland: Państwowe Wydawnictwo Naukowe.

Nosal, C. S. (2021). On the relationship between intuition, consciousness and cognition: in search of a unified concept of mind. *Annals of Psychology: Perception and Cognition*, *24*(3-4), 345-360.

Nunnally, J. C. (1978). *Psychometric theory*. New York, USA: McGraw-Hill.

Petrides, K. V., Perazzo, M. F., Pérez-Díaz, P. A., Jeffrey, S., Richardson, H. C., Sevdalis, N., & Ahmad, N. (2022). Trait Emotional Intelligence in Surgeons. *Frontiers in Psychology*, *13*, doi: 10.3389/fpsyg.2022.829084

Prigogine, I. (1996). *La fin des certitudes*. Truffaut, Paris: Editions Odile Jacob.

Reber, R., Ruch-Monachon, M. A., & Perrig, W. J. (2007). Decomposing intuitive components in a conceptual problem solving task. *Consciousness and Cognition*, *16*(2), 294-309. doi: 10.1016/j.concog.2006.05.004

Rencher, A. C. (2002). *Methods of Multivariate Analysis. Second edition*. New York, USA: Wiley-Interscience.

The Institute of Operational. (2012). *Sound Practice Guidance.* Retrieved from https://www.ior-institute.org/sound-practice-guidance, access 8.04/2022

Robin, R., & Riedel, J. C. K. H. (1997). Design and Innovation in Successful Product Competition. *Technovation*, *17*(10), 537-548. doi: 10.1016/s0166-4972(97)00050-3

Rokeach, M. (2015). The Open and Closed Mind. Investigations into the Nature of Belief Systems and Personality Systems. Connecticut, USA: Martino Fine Books.

Rotter, J. B. (1966). Generalized Expectancies for Internal versus External Control of Reinforcement. *Psychological Monographs*, *80*, 1-28. doi: 10.1037/h0092976

Sarbin, T. R. (1944). The logic of prediction in psychology. *Psychological Review*, *51*, 210-228.

Simon, H. A. (1992). What is an 'Explanation' of behavior? *Psychological Science*, *3*(3), 150-161.

Singer, R. (1975). *Motor Learning and Human Performance*. New York, USA: Macmillan and Company.

Slovic, P. (2000). The Perception of Risk. In: P. Slovic (Ed.), *The Perception of Risk*. London, England: Earthscan Publications.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. London, England: Macmillan.

Strelau, J. (2014). *Różnice Indywidualne. Historia – Determinanty – Zastosowania* [*Individual Differences: History – Determinants – Applications*]. Warsaw, Poland: Wydawnictwo Naukowe Scholar.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293-315.

Ungson, G. R, & Braunstein, D. N. (1984). Decision Making: An Interdisciplinary Inquiry. *Administrative Science Quarterly*, 29, 479.

Zakrzewska, M. (1994). *Analiza czynnikowa - zastosowania naukowe. Psychologia – metodologia* [*Factor analysis - scientific applications. Psychology – methodology*]. Poznań, Poland: Wydawnictwo Naukowe UAM.

Zimmerman, B. J. (2000). Attaining Self-Regulation. *Handbook of Self-Regulation*, 13-39. doi: 10.1016/b978-012109890-2/50031-7

Zuckerman, M. (2007). Sensation Seeking and Risky Behavior. Sensation Seeking and Risky Behavior. *American Psychological Association*; New York, doi: 10.1037/11555-000

**Erratum**

**Eratum to: Katarzyna Halicka and Arkadiusz Gola**

**Gerontechnology ranking using the TOPSIS methods**

**Engineering Management in Production and Services. Volume 16, Issue 1, pages 93-103.**
**(DOI: 10.2478/emj-2024-0007)**

In the Acknowledgements (page 101) the funding source of the research was erroneously stated. The correct Acknowledgements section appears below:

The research leading to these results has received funding from the commissioned task entitled "VIA CARPATIA Universities of Technology Network named after the President of the Republic of Poland Lech Kaczyński", under the special purpose grant from the Minister of Science, contract no. MEiN/2022/DPI/2577 action entitled "In the neighborhood - inter-university research internships and study visits.

https://doi.org/10.2478/emj-2025-0016