



A Bibliometric Review of Portfolio Optimization Models: An Operations Research Perspective

Dr. Richa Kalpesh Saxena

Richa.Saxena@nmims.edu

NMIMS, Anil Surendra Modi School
of Commerce, Mumbai Campus,
India

Devansh Bhatthena

devanshbhatthena@gmail.com

NMIMS, Anil Surendra Modi School
of Commerce, Mumbai Campus,
India

Aryan More

more.aryan1307@gmail.com

NMIMS, Anil Surendra Modi School
of Commerce, Mumbai Campus,
India

Dhairya Patel

dhairyapatel795@gmail.com

NMIMS, Anil Surendra Modi School
of Commerce, Mumbai Campus,
India

Dev Darira

devdarira2007@gmail.com

NMIMS, Anil Surendra Modi
School of Commerce, Mumbai
Campus, India

Disha Tiwari

tiwaridisha09@gmail.com

NMIMS, Anil Surendra Modi
School of Commerce, Mumbai
Campus, India

ABSTRACT

This study presents a bibliometric review of portfolio optimization research from an Operations Research perspective, covering literature published between 2015 and 2026. A final corpus of 708 documents was identified through a PRISMA-based selection process using Scopus. Biblioshiny and Microsoft Excel were used for data extraction, analysis, and visualization. The study examines publication and citation trends, source performance, author productivity, institutional and country-level output, international collaboration, keyword patterns, conceptual clusters, and thematic evolution. Four research questions address publication and citation activity; prominent journals, authors, institutions, and countries; the field's main conceptual and methodological approaches; and changes in themes between earlier and later periods. The results show that portfolio optimization and portfolio selection remain central to the literature. Robust optimization, robust portfolio optimization, genetic algorithms, asset allocation, machine learning, portfolio management, multi-objective optimization, and Conditional Value-at-Risk also appear among frequently observed concepts. Publication output is concentrated in a relatively small group of sources, while researchers and countries contribute across a broad international base. China, the United States, and India are major contributors, and the collaboration analysis shows links among leading research-producing regions. The thematic analysis indicates continuity in the field's core portfolio optimization theme alongside increasing specialization in uncertainty modelling, advanced risk measures, metaheuristics, machine-learning applications, and dynamic stochastic models. The review identifies research opportunities at the intersection of robust or distributionally robust optimization and machine-learning-integrated decision models. It highlights the need to compare these techniques with classical mean-variance models using consistent datasets, realistic constraints, and common evaluation standards. Future research directions include out-of-sample testing, method benchmarking, transparent data definitions, ESG integration, emerging computational approaches, and broader geographic collaboration. Findings remain bounded by the Scopus-based search strategy and the 2015–2026 observation window; citation-age effects and the partial nature of the final year should also be considered.

Keywords: Operations Research, Portfolio Optimization, Risk Management, Bibliometric Analysis, Machine Learning

1. INTRODUCTION

Portfolio construction has long been a proving ground for operations research methods in finance. Ever since the problem of balancing between expected return and acceptable variance was first framed as a quadratic programme over the early 1950s, the challenge of allocating capital across assets in a mathematically defensible manner has drawn contributions from mathematical programming, stochastic programming, heuristic search and, more recently, machine learning. Each of these methodological traditions has come to occupy a substantial literature of its own, and together have formed a corpus of work large enough that any

individual narrative review cannot hope to encapsulate in its entirety. It is this scale, and the associated difficulty to track where the field has been and where it appears to be headed, that motivates a bibliometric approach to studying portfolio optimization.

Bibliometric analysis offers a perspective that allows stepping back from individual papers to examine the structure of a research area as a whole: how much is being published and when, which journals, authors, institutions and countries are driving that output, which concepts recur most often, and how the relationships between those concepts change over time. Rather than summarizing a curated sample of studies, as a conventional narrative review would, a bibliometric study works from the citation record of a field, using tools such as Bradford's Law, Lotka's Law, factorial analysis and thematic mapping to describe patterns that would be difficult to see from reading papers one at a time. Applied to portfolio optimization, this kind of analysis can show whether the field is still expanding or beginning to plateau, whether its research effort is concentrated in a small number of outlets and institutions or spread broadly, and whether newer computational approaches such as robust optimization and machine learning are displacing classical mean-variance methods or simply being layered on top of them.

This study applies exactly that lens to the literature on portfolio optimization published between 2015 and 2026. Using Scopus as the source database and the Biblioshiny interface to the Bibliometrix R package for the bulk of the analysis, together with Microsoft Excel for supporting calculations, the study draws on a final corpus of 708 documents assembled through a PRISMA-based selection of identification, screening, eligibility assessment and inclusion. The dataset spans several document types, predominantly journal articles, alongside a smaller number of conference papers, reviews, books and book chapters, and its scale is sufficient to support the kind of aggregate, pattern-level analysis that a bibliometric study depends on.

The analysis is organized around four research questions. The first asks how publication output and citation activity in portfolio optimization research have progressed across the 2015 to 2026 period, which requires separating the trajectory of how much is published from the trajectory of how much that work is subsequently cited, since the two need not move together. The second asks which journals, authors, academic institutions and countries have played the most prominent roles in shaping the field, an inquiry that combines source-level indicators such as Bradford's Law and the local H-index with author-level measures such as Lotka's Law and country-level measures of both output and collaboration. The third asks what conceptual frameworks and methodological approaches characterize the literature, addressed through keyword frequency analysis and factorial mapping of how concepts cluster together. The fourth asks how the field's thematic structure has evolved between an earlier period and a later one, using thematic mapping to identify which topics function as well-established, central themes, which remain specialized niches, which serve as foundational but still-developing building blocks, and which are emerging or declining at the margins of the field.

Beyond simply describing these patterns, the study is also concerned with what they imply for where portfolio optimization research might productively go next. A recurring observation that motivates this concern is that much of the field's apparent innovation takes the form of established techniques being added onto classical portfolio formulations rather than being tested rigorously against them. Robustness features, machine-learning-based forecasts, metaheuristic solvers and other extensions frequently appear as enhancements to mean-variance optimization without a parallel body of work that directly compares their real-world, out-of-sample performance to the classical approaches they are meant to improve upon. Recommendations that research should simply incorporate "more" artificial intelligence or "more" sustainability considerations, without specifying what comparative standard that research should be judged against, offer comparatively little guidance for where effort is best directed. By combining a structural account of the literature's development with an assessment of where its methodological gaps lie, this study aims to provide both a map of portfolio optimization research as it currently stands and a more specific agenda for the comparative, benchmark-driven work that appears still to be missing.

The remainder of the paper is organized as follows. Section 2 presents the research methodology, including the search strategy, screening process, and analytical tools. Section 3 reviews the relevant literature, and Section 4 states the research objectives and questions. Section 5 reports the bibliometric results on publication and citation trends, research sources, contributors, geographic distribution, conceptual structure, and thematic evolution. Section 6 discusses the findings, research gaps, and future research agenda. Section 7 concludes the paper and outlines its limitations.

2. RESEARCH METHODOLOGY

2.1 Article Selection, Inclusion, and Exclusion Criteria

This study adopted a bibliometric research design to examine the development and research trends in portfolio optimization within the field of Operations Research. An article selection process using a PRISMA-based analysis that included identification, screening, eligibility, and inclusion phases was conducted.

The literature search was carried out using Scopus. A predetermined search strategy was implemented to retrieve the main lines of research related to portfolio optimization; where five groups of search query were considered: portfolio optimization, portfolio selection, asset allocation, robust and stochastic portfolio optimization, and portfolio optimization in combination with heuristic and meta-heuristic methods (e.g. Genetic algorithm, particle swarm optimization, and simulated annealing). The search was made within titles, abstracts and keywords, getting 960 documents.

Of the resulting articles, the records retrieved were further screened and curated. 113 duplicate records were removed, resulting in 847 articles after duplicates screening. The remaining records were then scrutinized for relevance to the scope of the study, leading to 131 articles to be discarded, leaving 716 potentially relevant articles.

The remaining records were subsequently re-screened for relevance and underwent a metadata quality assessment. After this process, eight studies were removed due to concerns regarding their bibliographic metadata and the final corpus consisted of 708 published studies. The entire process of finding published articles can be seen in the PRISMA flow diagram in Figure 1.

2.2 Data Extraction and Analysis

Citation information of the 708 final studies was obtained from Scopus, consisting of authorship details, publication sources, keywords, citations, and abstracts. The data covered 623 articles, 43 conference papers, 29 reviews, 8 books, and 4 book chapters covering the years 2015-2026.

The collected records were carefully checked and recoded, such as authors' names normalization before conducting the bibliometric analysis. A web interface tool for Bibliometrix R package, also called Biblioshiny, was used to conduct the bibliometric analysis, with Microsoft Excel used for data searching and cross-comparing.

Within the framework of the study, five major perspectives of the portfolio optimization literature were examined: development, source dynamics, author productivity, production by country, and conceptual and theoretical structure of the field. Development of the research was investigated by yearly publication and citation figures. Performance of sources was explored with the application of Bradford's Law and local H-indexes. Productivity of the researcher was studied with the application of the author's publication and Lotka's Law, production by country with countries' productivity and international collaboration, and conceptual and theoretical structures of the field with relevant concept analysis, factorial analysis, thematic mapping and thematic evolution.

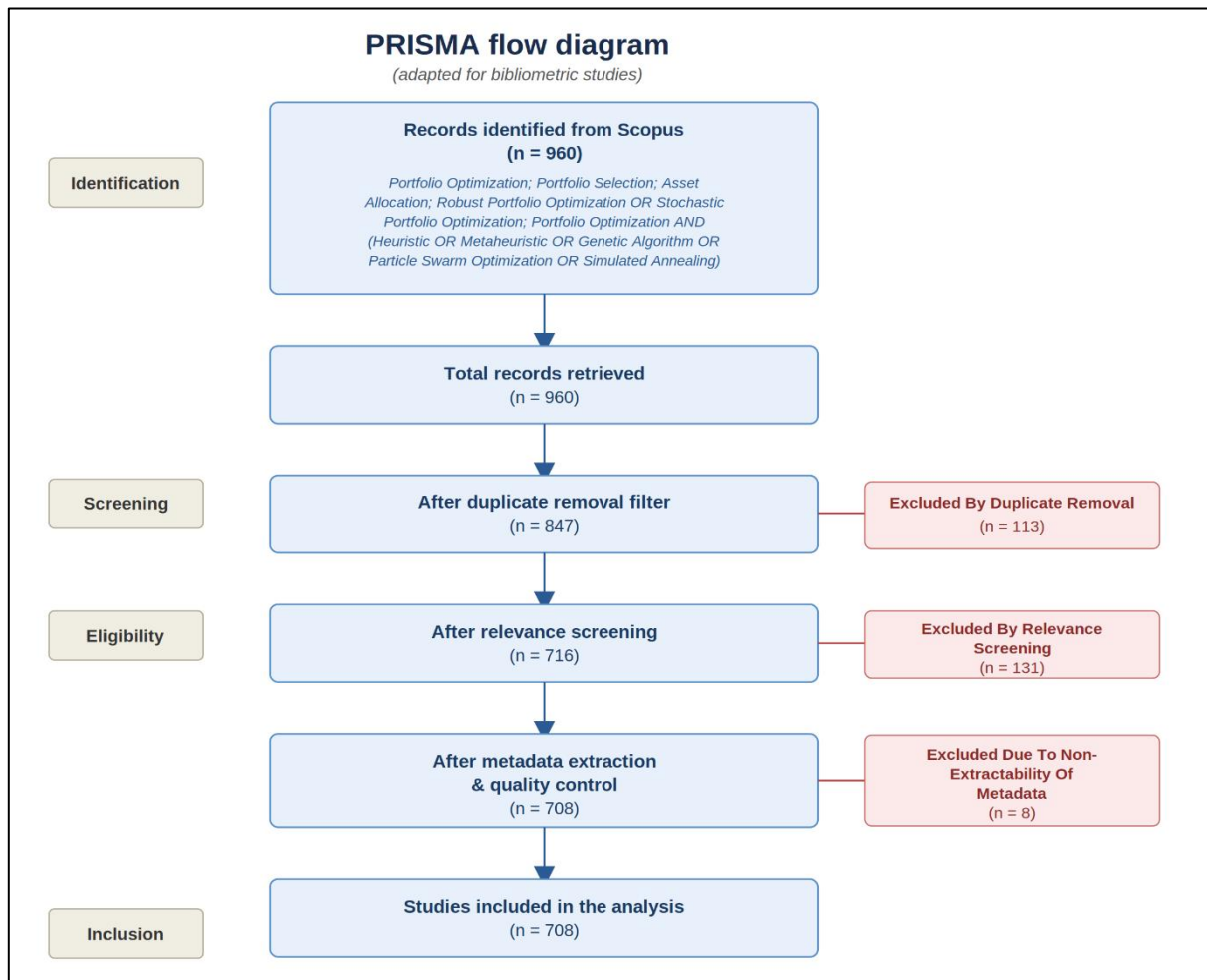


Figure -1: PRISMA-Flow Diagram (extracted from Biblioshiny – author's work)

3. LITERATURE REVIEW

3.1 Introduction

[1] framed portfolio selection around the relationship between expected return and variance, establishing a foundational mean–variance approach to portfolio construction. Subsequent work has extended this problem through mathematical and stochastic programming, heuristic search, robust optimization, and machine-learning methods.

The bibliometric analysis covers 708 documents from 2015 to 2026. This corpus represents approximately 72.8% of the estimated total of 973 documents implied by the fitted growth curve. The life-cycle analysis suggests that the field is comparatively mature overall, although particular methodological intersections - especially robust optimization combined with machine-learning-integrated optimization - remain underexplored.

3.2 Classic and Risk-Measure Extensions Mean-Variance (MV)

Optimization is the reference case for many OR applications in finance; the assumption it penalizes downside and upside deviation equally means however that other risk measures should be preferred when using optimisation, preferably such that this is at scale.

Conditional Value-at-Risk (CVaR) optimisation is one of the most prevalent OR-applied measures as it can be expressed as a linear program and is coherent (while standard measures such as variance or Value-at-Risk do not adhere to these consistency principles).

Conditional Value-at-Risk (CVaR) optimization represents a prominent Operations Research application in the portfolio-optimization literature. In the keyword analysis, CVaR appears among the ten most frequent author keywords, with 16 occurrences, and is connected to “portfolio optimization” and “robust optimization” in the keyword-cluster analysis. “Robust optimization” and “robust portfolio optimization” appear 37 and 31 times, respectively, indicating the importance of uncertainty-aware methods in the reviewed corpus.

3.3 Metaheuristic and Evolutionary Optimization Approaches

A key application of Operations Research in portfolio optimization is to account for factors beyond return and risk by incorporating realistic constraints, including integer constraints such as minimum trade size and limits on the number of assets in a portfolio.

Because realistic constraints such as order lot sizes or restrictions on short selling can produce mixed-integer formulations that standard methods may not solve within reasonable time limits, the literature has extensively applied metaheuristics. These methods include evolutionary computation, such as genetic algorithms, and nature-inspired optimization. Genetic algorithms are the fifth most frequent keyword, with 30 occurrences, and appear in a cluster distinct from classical mean–variance and Sharpe-ratio problems. The trend-map analysis indicates that genetic algorithms were established before the more recent wave of machine-learning methodologies discussed in the next section.

Survey literature such as [2] for the application of swarm intelligence metaheuristics to portfolio selection shows the scale of research this sub-topic has garnered.

3.4 Robust and Distributionally Robust Optimization

Another strand, which is, surprisingly often, cited by a very significant amount, refers to the problem of optimizing a portfolio subject to estimated and uncertain mean-return and covariance parameters. These models assume an uncertainty set for the unknown parameters, and optimize the portfolio against the worst-case realization inside this set. A version of this concept, which tries to control for worse than possible worst-case realization is known as distributionally robust optimization (DRO) - it controls for the worst-case probability distribution over the set of possible distributions.

An influential contribution to distributionally robust optimization is the paper by [3], which develops tractable reformulations using the Wasserstein metric and has 1,603 citations in the analysed collection. This is approximately 2.58 times the 622 citations reported for the next-ranked paper, showing the prominence of this contribution within the corpus. Robust optimization remains a frequent keyword, with 37 occurrences.

More broadly, the citation profile varies substantially across papers and publication years. Older studies have had more time to accumulate citations than recently published studies, so citation counts should be interpreted alongside publication age. Robust optimization, including variants that address uncertainty in model parameters, remains an important methodological theme in the collection.

3.5 Machine Learning in Operations Research

Most recent machine-learning applications for portfolio selection generally take one of two approaches: they either use a machine-learning algorithm to predict financial returns, which are then fed into classical mean-variance or robust optimizer, or machine learning is integrated with other forms of O.R.

Approaches that combine prediction and optimization either use a separate forecasting model whose outputs feed into a classical mean–variance or robust optimizer, or integrate prediction with downstream decision-making. [4] present the predict-then-optimize framework and are the second-ranked paper by citations in the analysed collection, with 622 citations. [5] provide an example of portfolio optimization using return prediction with machine-learning and deep-learning methods. Although these approaches are receiving increased attention, systematic comparisons between separate forecast-then-optimize pipelines and integrated approaches remain limited.

3.6 Field-level Patterns

The field-level analysis confirms a trend across the papers: we see that majority of our selected papers come from OR and computer science themed publications.

Among the top 10 of sources, one can see journal such as ‘Expert Systems with Applications’, ‘European Journal of Operational Research’, ‘Annals of Operations Research’, ‘IEEE Access’ or ‘Applied Soft Computing’.

The authorship analysis shows that approximately 87.1% of contributors in the analysed set published only one paper. The results also show that many countries contribute to this field, with China, the United States, and India occupying the top three positions in this collection. China has the highest publication output and the strongest collaborative involvement among these countries. Given the field’s relatively mature state, contributions appear more often to extend established methodological foundations than to introduce entirely new approaches.

4. RESEARCH OBJECTIVE

Implementing a systematically reviewed bibliometric approach, this study seeks to unearth primary trends and prevailing themes within the published literature. The extent of the investigation is described by the following research questions:

RQ1: How publication outputs and citation flows have progressed within portfolio optimization research between 2015 and 2026?

RQ2: Which key journals, authors, academic institutions and geographic regions have played important roles in the area?

RQ3: What primary conceptual frameworks and methodological approaches characterize this body of literature?

RQ4: How have the key thematic areas evolved from the earlier period to the later period?

5. DATA ANALYSIS AND RESULTS

5.1 Publication and Citation Trends

An analysis of the publication trends for portfolio optimization models from 2015 to 2026, summarized in Figure 2, shows that there was a steady pace of research published over time before a sudden drop off. The total of papers published per year began at 63 in 2015 and 62 in 2016, followed by a peak of 89 in 2017. The results remained fairly high for the following four years with 71, 74, 71 and 80 published in 2018, 2019, 2020 and 2021 respectively. Since this time however, there has been a steady downtrend annually, with the totals dropping to 65, 59, 33, 25 and 15 in 2022, 2023, 2024, 2025 and 2026 respectively.

The publication series shows an initial increase from 2015 to the 2017 peak, followed by relatively high output through 2021 and a decline thereafter. The largest year-on-year reduction in the reported series occurred between 2023 and 2024, when the number of publications fell from 59 to 33. The overall decline from 80 publications in 2021 to 15 in 2026 spans several years and should not be described as a single-year fluctuation.

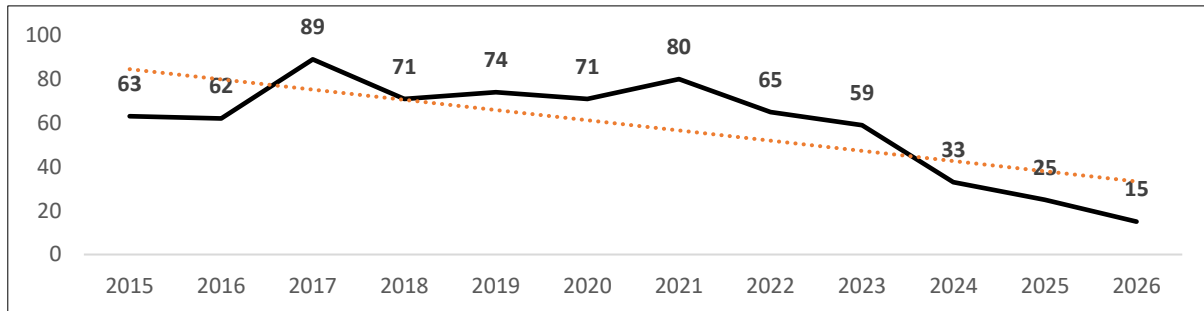


Figure -2: Yearly publication volume in Portfolio Optimization Research (2015-2026) (extracted from Excel – author’s work)

Citation performance shows a different pattern over time compared with overall publication output. As shown in Figure 3, average total citations per year grew from 6.75 in 2015 to a peak of 12.70 in 2023 before dropping in the following years. Papers published in 2018 achieved an average of 82.94 citations per article, whereas papers published in 2026 averaged 5.60 citations per article.

However, these citation differences need to be understood in terms of paper age. Older articles have had much more time to be cited by successive studies, while newer papers have had very little time to gather citations [6]. Therefore, lower citation numbers for recent years should not be viewed as affirmation of lower research quality or less impact. Instead, they mostly reflect the difference in time available for citations to build up [7].

As a result, citation trends do not track publication trends exactly. Although annual publication numbers dropped observably after 2021, the average citations per year continued to rise until 2023 before turning downward. This highlights that publishing activity and citation growth follow different timelines across the study period.

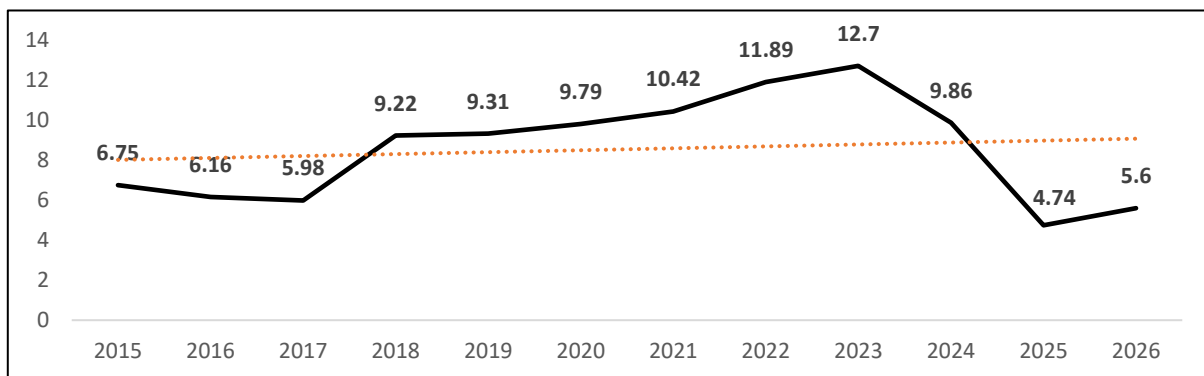


Figure -3: Annual trend of Average Total Citations (2015-2026) (extracted from Excel – author’s work)

5.2 Research Sources, Contributors and Geographic Distribution

5.2.1 Key Journals and Sources

The published literature on portfolio optimization models is centred around a relatively small group of key journals. As illustrated in Figure 4, Expert Systems with Applications leads the field with 52 published articles. It is followed by European Journal of Operational Research (32 articles), Annals of Operations Research (17 articles), IEEE Access (16 articles), and Applied Soft Computing (13 articles). Other leading journals each contributed between 7 and 12 articles.

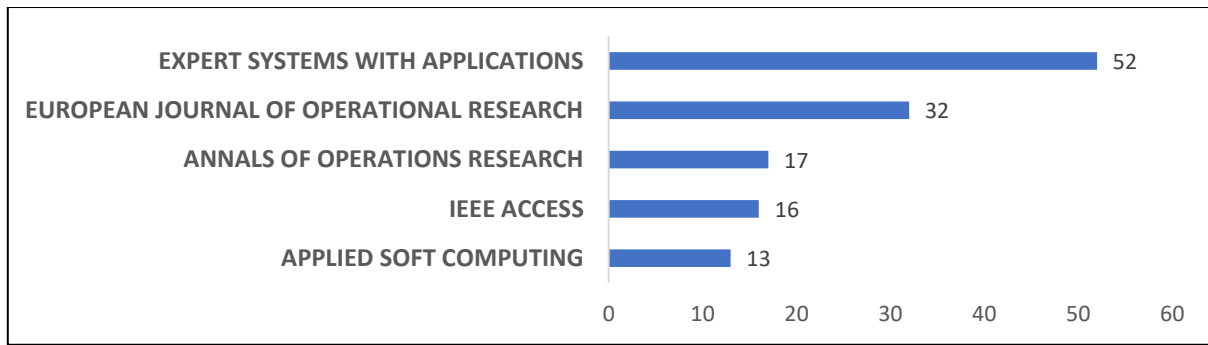


Figure -4: Most Relevant Sources in Portfolio Optimization Research (extracted from Excel – author’s work)

Bradford’s Law [8], illustrated in Figure 5, shows that portfolio-optimization research is concentrated in a relatively small core of journals. In this analysis, 17 core sources account for 238 papers, 76 sources form the second zone, and 213 sources form the third zone. This pattern indicates that a relatively small group of journals publishes a substantial share of the literature, while the remaining papers are distributed across a much larger set of outlets.

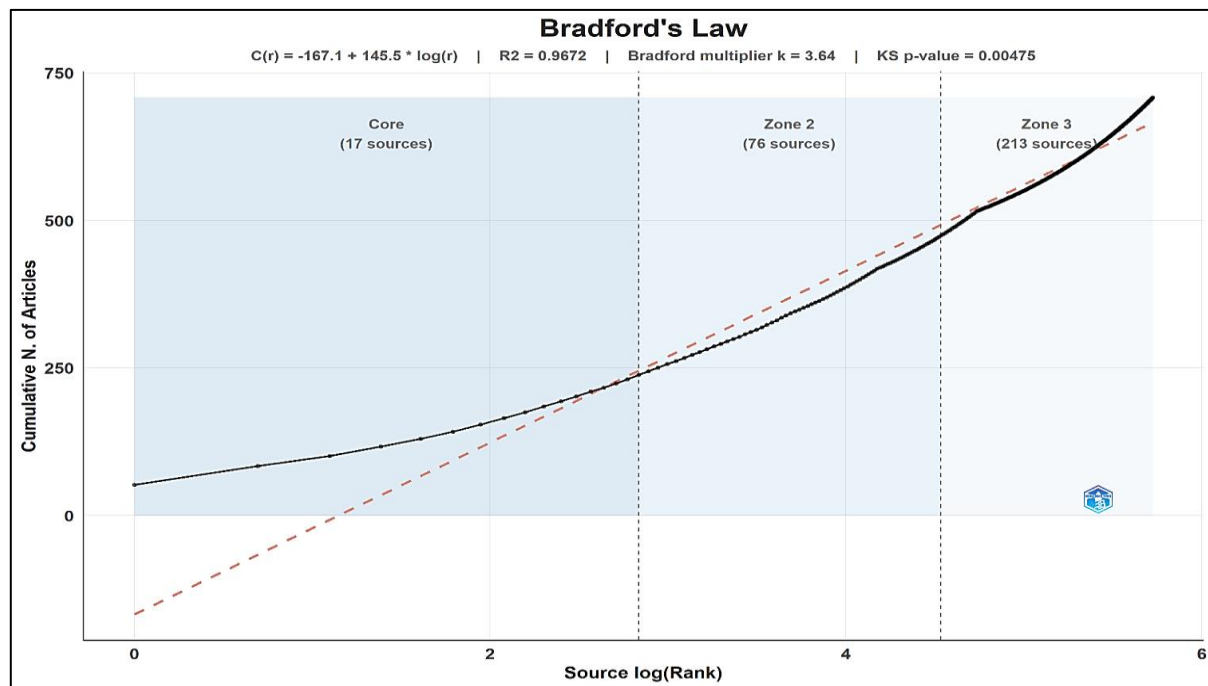


Figure -5: Bradford’s Law Distribution of Sources (extracted from Biblioshiny – author’s work)

Alongside total publication counts, the local H-index values in Figure 6 help measure overall journal impact [9], [10]. Expert Systems with Applications leads with the highest local H-index of 40, succeeded by the European Journal of Operational Research at 25 and IEEE Access at 16. Annals of Operations Research and Applied Soft Computing achieve local H-indices of 14 and 13 respectively. Overall, the journals that publish the highest number of papers also rank among the top in terms of local H-index impact and the overall journal-level analysis indicates that a small group of outlets occupies prominent positions in both publication output and local citation impact.

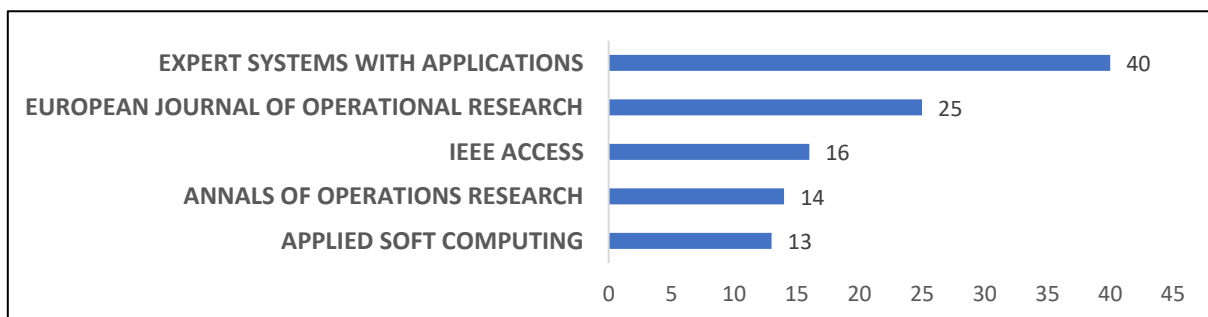


Figure -6: Local H-index of Leading Sources (extracted from Excel – author’s work)

5.2.2 Key Authors

An analysis of individual authors shows that research in this field is spread across a wide range of scholars, with even the most active researchers publishing a modest number of papers. As displayed in Figure 7, Chen W leads in total publication output with 7 articles, followed by Chen C-H with 6 articles. Meanwhile, Chou Y-H and Zhang Y authored 5 articles each. Another group of researchers, including Cui T, Jalota H, Li B, Ma Y, Mehlawat MK and Platanakis E contributed 4 articles each.

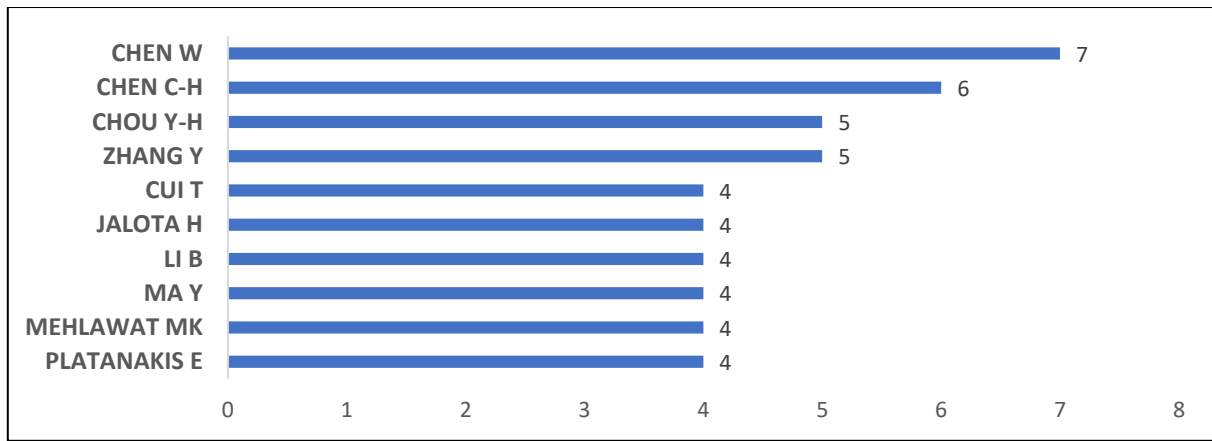


Figure -7: Most Relevant Authors in Portfolio Optimization Research (extracted from Excel – author’s work)

In an attempt to better understand author output, Figure 8 uses Lotka’s Law [11] to analyse researcher productivity. The results indicate that 512 authors, reported as approximately 87.1% of contributors in this analysis, wrote only one paper, whereas 51 authors published two papers. This pattern indicates that most contributors in the analysed set published once, with a smaller group contributing repeatedly.

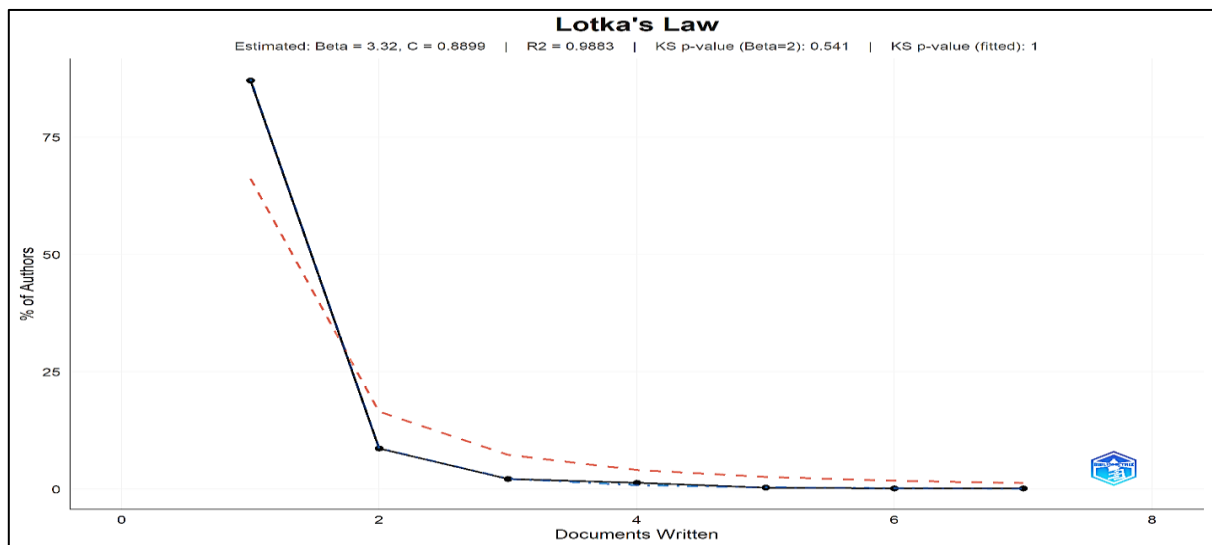


Figure -8: Lotka’s Law of Scientific Productivity (extracted from Biblioshiny – author’s work)

5.2.3 Leading Academic Institutions

Looking at academic institutions, published research is being done by many universities, but only a handful of institutions account for the highest number of papers. As illustrated in Figure 9, Islamic Azad University ranks first with 20 articles, followed by the University of Delhi with 12 and Beihang University with 10. The Hong Kong Polytechnic University added 9 articles, while Capital University of Economics and Business, the Chinese University of Hong Kong, IIT Delhi, IIT Mandi, South China University of Technology and Tamkang University all reporting 8 papers.

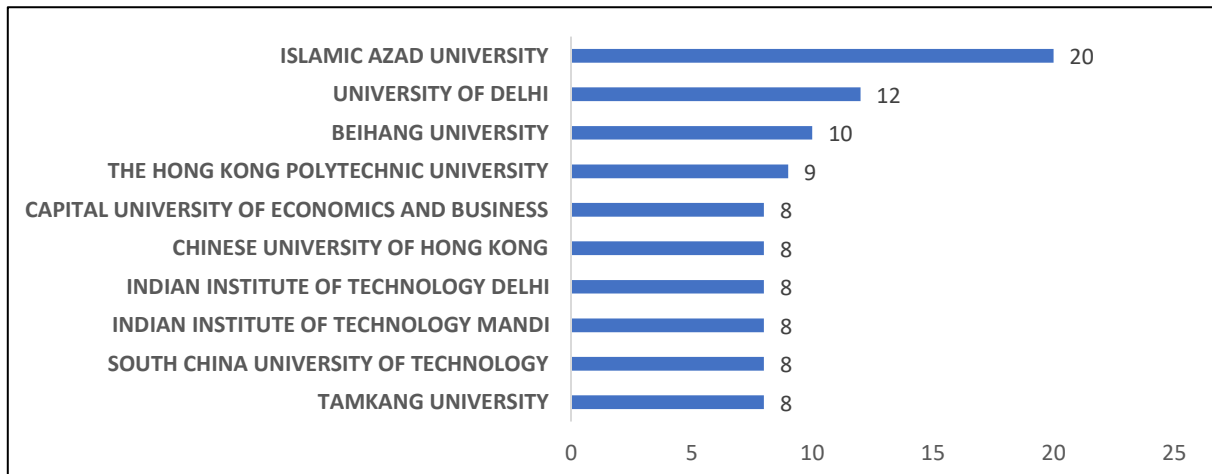


Figure -9: Most Relevant Academic Affiliations in Portfolio Optimization Research (extracted from Excel – author’s work)

The institutional analysis indicates that although the study involved many different schools and centres, only a few institutions account for the majority of total output. Also, the institutional results should be read in terms of overall publishing, not academic impact, since the institutional analysis presented here tracks only output volume.

5.2.4 Geographic Distribution and Collaboration

The country-level analysis indicates that portfolio optimization research maintains a wide international base, although most of the published articles are more centred in some major countries. Illustrated in Figure 10, China leads the field with 363 publications, followed by the United States with 180 and India with 109. The United Kingdom (79) and Iran (73) hold the next highest positions, while Italy, Germany, France, Spain and Turkey complete the top ten contributors. The noticeable distance between China's total output and other top-ranking countries highlight a clear collection of research in leading nations.

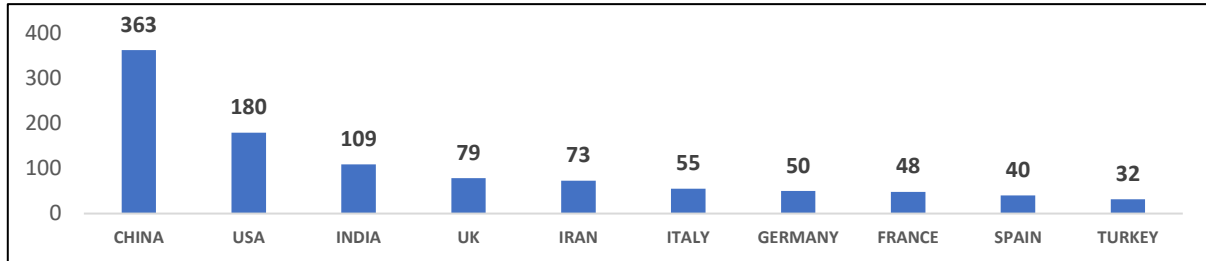


Figure -10: Most Productive Countries in Portfolio Optimization Research (extracted from Excel – author's work)

Figure 11 presents country productivity across the 2015–2026 observation window. By the end of the study period, the cumulative totals were 363 publications for China, 180 for the United States, and 109 for India. The country series indicates different rates of accumulation over the period; these end-of-period totals should not be interpreted as publications produced during 2026 alone.

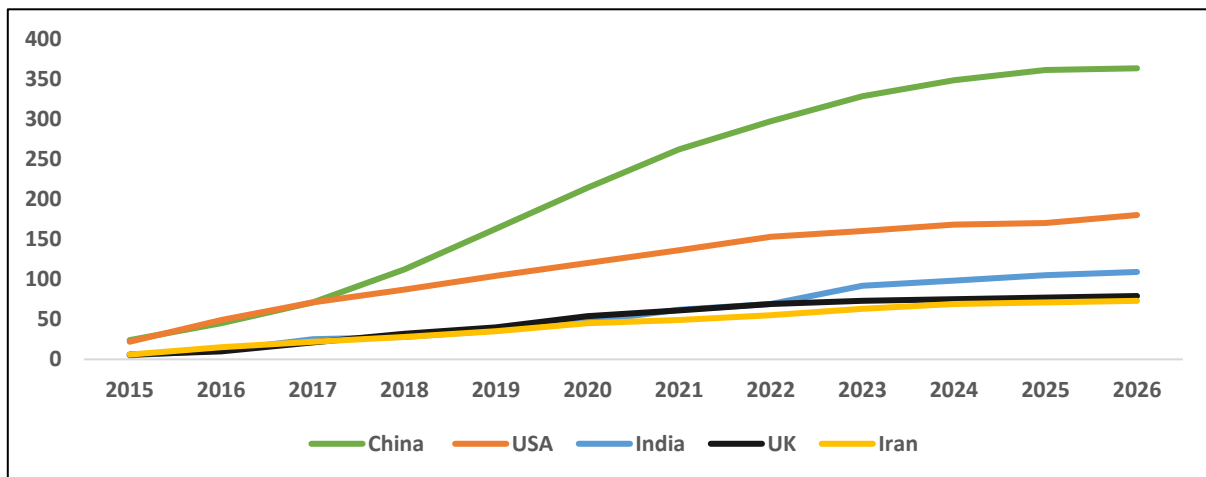


Figure -11: Country Productivity Over Time (2015-2026) (extracted from Excel – author's work)

The wide geographic spread of the literature is supported by international research collaboration. Figure 12 highlights China as a central node, with collaborative links to the United States (18 links), Hong Kong (16), Singapore (11), India (10), and the United Kingdom (10). These links indicate that leading research-producing countries collaborate across borders; the concentration of publication output in a few countries does not imply geographic isolation.

Overall, the geographic analysis shows that portfolio optimization research has an international publication base alongside clear differences in publication output across countries.

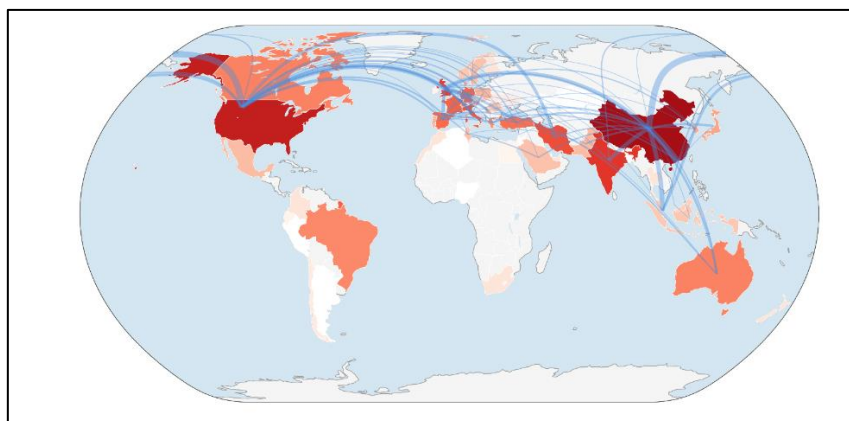


Figure -12: International Collaboration Network in Portfolio Optimization Research (extracted from Biblioshiny – author's work)

5.3 Conceptual and Methodological Structure

5.3.1 Dominant Conceptual Frameworks

The overall body of literature revolves around core ideas in investment and portfolio decision-making. As seen in Figure 13, Portfolio Optimization stands out as the most frequent concept, recording 179 occurrences, while Portfolio Selection ranks second with 80. Other noteworthy concepts featured in the data are Robust Optimization (37), Robust Portfolio Optimization (31) and Genetic Algorithm (30), Asset Allocation (23), Machine Learning (21), Portfolio Management (19), Multi-Objective Optimization (17) and Conditional Value-at-Risk (16).

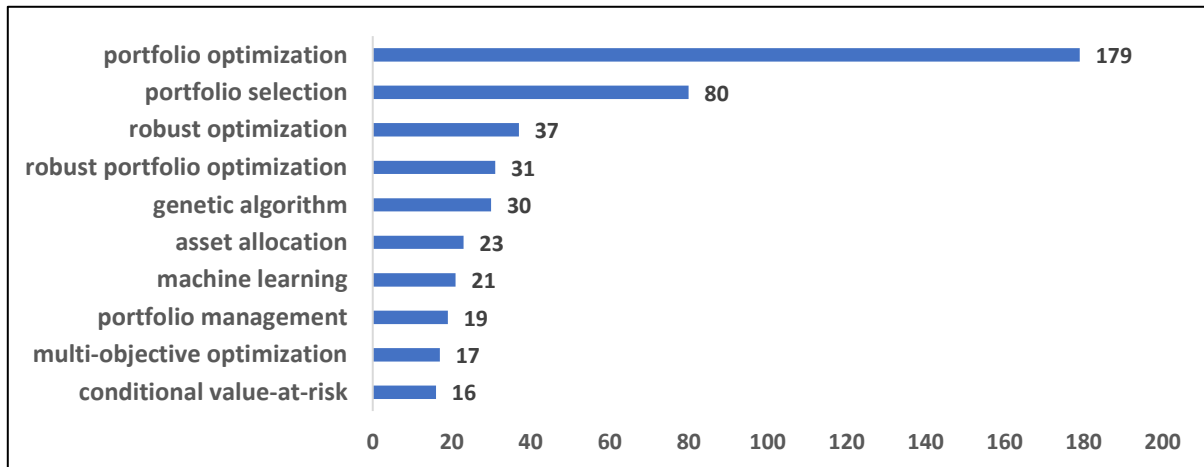


Figure -13: Most Relevant Concepts in Portfolio Optimization Research (extracted from Excel – author’s work)

The evidence at hand indicates that although the traditional Portfolio Optimization makes the base of the body of work, newer research may in the future include advanced risk management solutions, for instance Conditional Value-at-Risk [12], [13]. Additionally, the commonness of Machine Learning, Genetic Algorithms and Robust Optimization highlights an evolution towards multi-objective computational modelling. Overall, these concepts reflect a field driven by algorithmic and data-informed decision making [14], [15].

5.3.2 Methodological Approaches and Conceptual Clusters

Factorial Analysis further explains how major themes are structured across the literature [16]. As illustrated in Figure 14, keyword patterns reveal clear clusters corresponding to core portfolio decision-making, metaheuristics, machine learning, classical finance and uncertainty modelling. The main cluster comprises *Portfolio Selection*, *Portfolio Optimization*, *Risk Management*, *Portfolio Management*, *Asset Allocation*, *Robust Optimization* and *Multi-Objective Optimization*. Secondary clusters capture algorithmic methods such as *Swarm Intelligence* and *Genetic Algorithms*, data science tools like *Machine Learning* and *Deep Learning* and traditional financial concepts including *Mean-Variance* and the *Sharpe Ratio*.

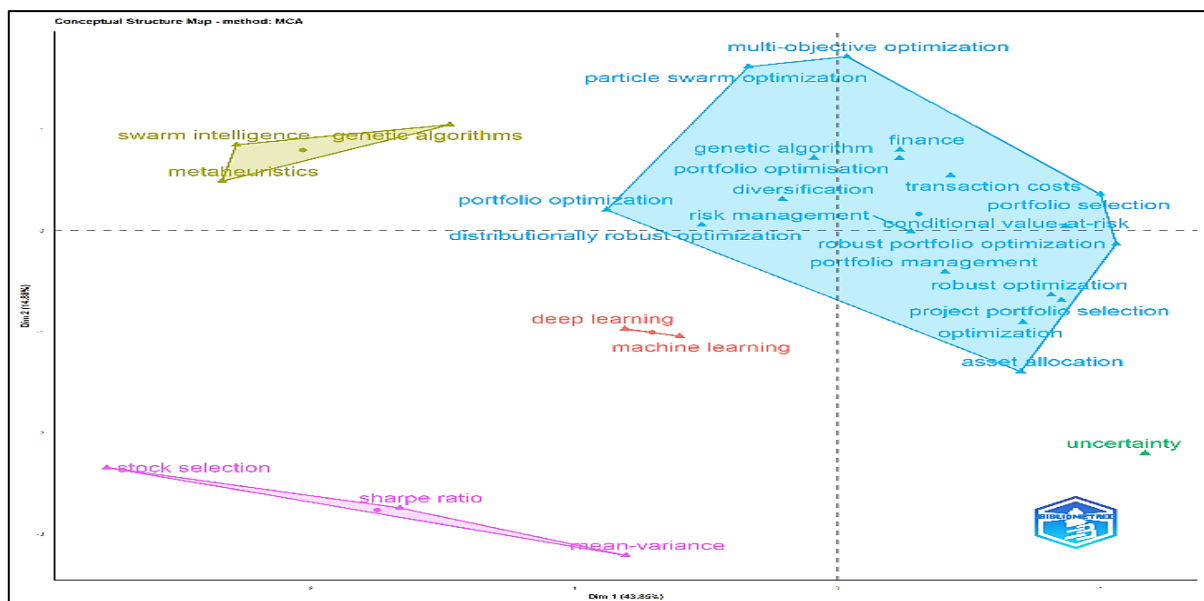


Figure -14: Factorial Map of Keyword Clusters in Portfolio Optimization Research (extracted from Biblioshiny – author’s work)

Altogether, the grouping of these themes shows how the field integrates time-tested financial frameworks with modern computational tools [14], [17].

5.4 Thematic Evolution Across Study Periods

5.4.1 Thematic Structure

This thematic map is a representation of the co-occurrences of themes: it reflects a complete research scheme, which covers basic portfolio theories, optimization algorithms, risk-related subjects, and particular financial applications. Within the motor-theme (central and well-developed) region, topics such as *Portfolio Selection*, *Genetic Algorithm*, *Multi-Objective Optimization*, *Optimization*, *Finance*, *Stock Selection*, *Particle Swarm Optimization* and the *Sharpe Ratio* display high internal development and strong external ties, indicating that these themes occupy well-developed and highly connected positions within the thematic structure.

By comparison, the niche-theme (outermost and developed) region contains specialized concepts including *Copula*, *Stock Markets*, *COVID-19*, *Gold* and *Dependence Structure*. These themes show high internal maturity but lower integration with the broader network, pointing to specialized research pockets.

In the basic-theme (central and undeveloped) region, topics like *Robust Optimization*, *Robust Portfolio Optimization*, *Asset Allocation*, *Risk Management* and *Uncertainty* dominate. Although these themes show relatively low density, their high centrality highlights their role as fundamental building blocks across the discipline.

The emerging and declining (outermost and undeveloped) region captures several developing or specialized topics, including *Value-at-Risk*, *Conditional Value-at-risk*, *Machine Learning*, *Portfolio Management*, *Deep Learning*, *Cryptocurrency*, *Reinforcement Learning* and *Multi-Objective Portfolio Selection*. Their outermost location indicates that while actively researched, they occupy emerging or context-specific positions rather than central hubs [18].

In summary, the results are consistent with a history of progressive interaction between core portfolio-selection theories and the growth of computing and empirical methods. This progression provides a basis for considering how the field may develop across different time horizons. Figure 15 presents the thematic map that groups these topics into motor, niche, basic, and emerging or declining themes.

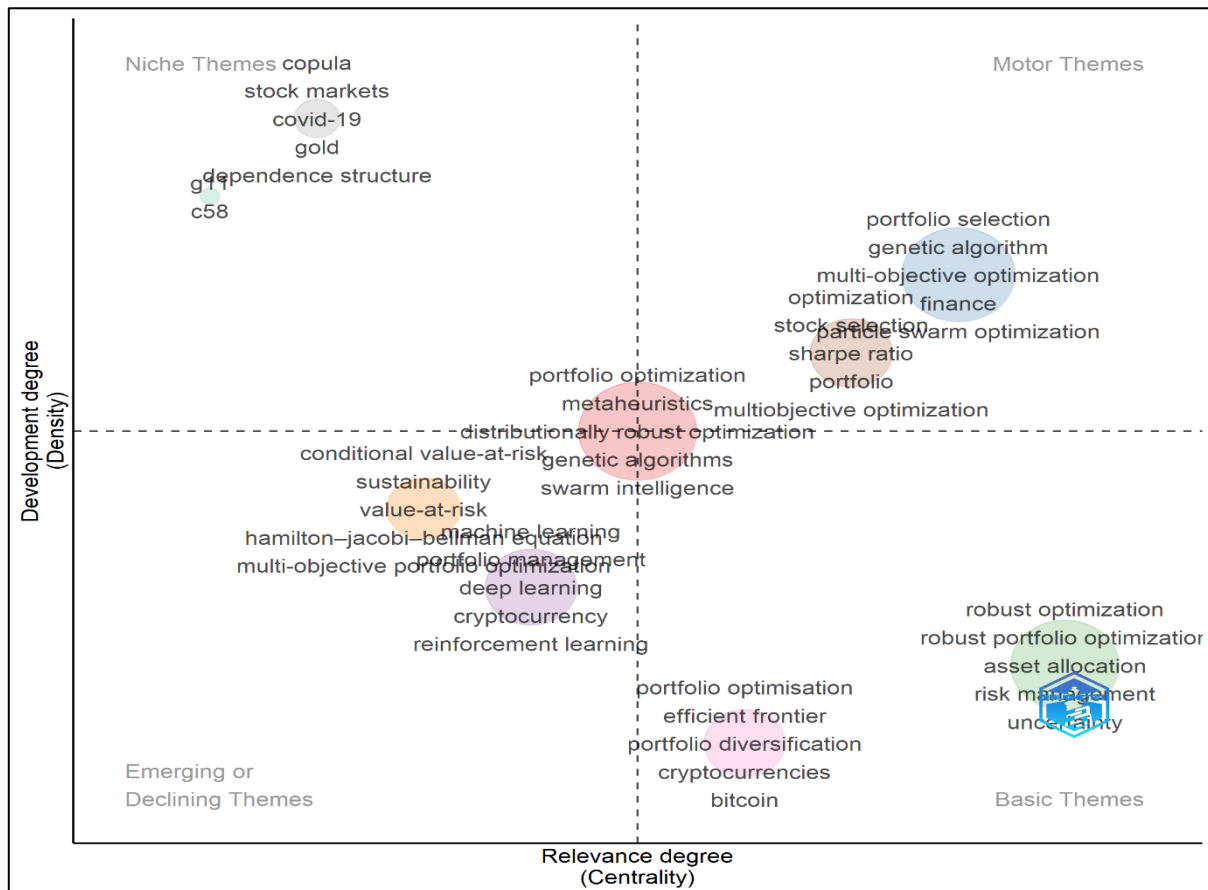


Figure -15: Thematic Map of Portfolio Optimization Research (extracted from Biblioshiny – author’s work)

5.4.2 Thematic Evolution Between the Two Periods

Analysing thematic evolution from 2015–2019 to 2020–2026 highlights the evolution of main research concepts over time. As visualized in Figure 16, the strongest thematic sequence occurs between Portfolio Optimization in the first phase and Portfolio Optimization in the second, with 82 co-occurrences and a weighted inclusion index of 0.592, indicating that the thematic core of the literature remained stable.

At the same time, the map reveals how earlier topics evolved into more refined and focused sub-domains in the latter period. For instance, Robust Optimization contributed directly to the emerging Robust Portfolio Optimization structure while maintaining links to broader portfolio studies [15]. Portfolio Selection carried forward into 2020–2026, expanding its connections to Conditional Value-at-Risk, Diversification, Metaheuristics, Portfolio Management and Portfolio Optimization [2], [12]. Additionally, Robust

Portfolio Optimization established links to Dynamic Stochastic Portfolio Optimization [19], while Risk Management aligned closely with Portfolio Management and Portfolio Optimization.

These cases of thematic change demonstrate a combination of a core of continuity and an extension of breadth: the core topics remained stable, while new extensions broadened the field's scope. Consequently, while the central core of portfolio optimization remained constant, related sub-fields became increasingly sophisticated through interactions with robust modelling, specialized risk tools, stochastic models, metaheuristic algorithms and practical management applications.

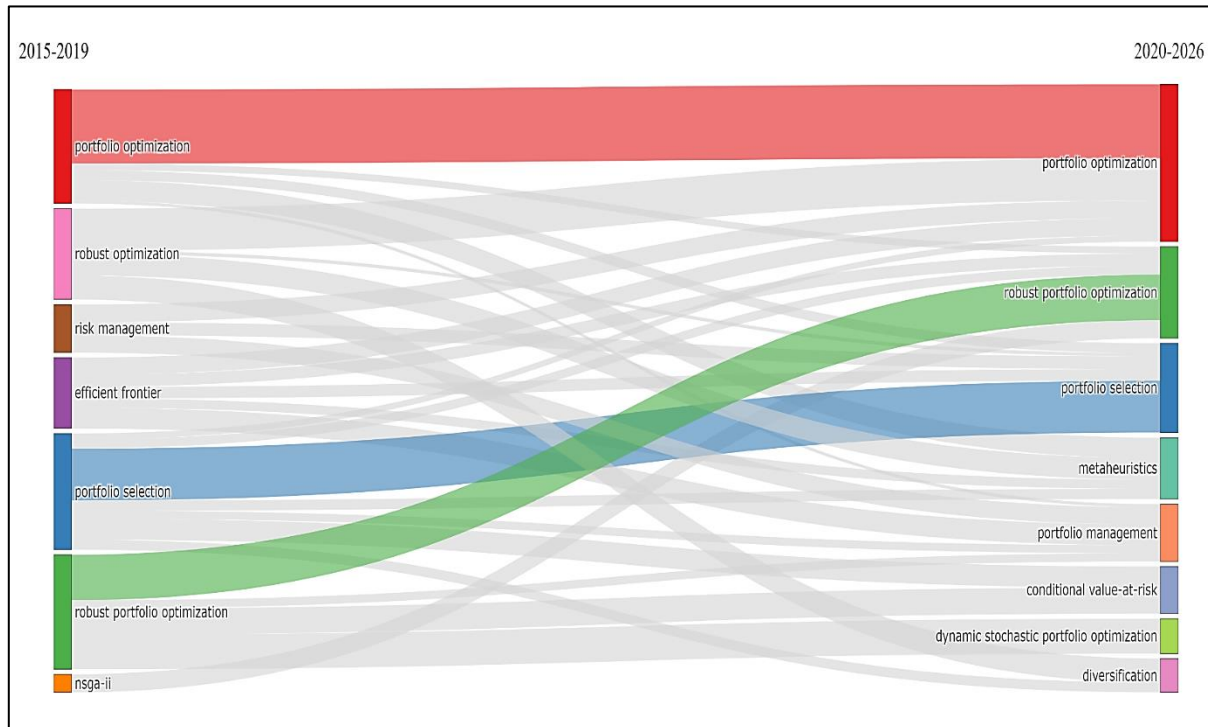


Figure -16: Thematic Evolution of Portfolio Optimization Research (2015-2026) (extracted from Biblioshiny – author's work)

6. DISCUSSION AND FUTURE RESEARCH AGENDA

6.1 Findings

All in all, the results of this survey show that the field of portfolio optimization is a well-established and ever-growing scientific field. It is linked through a common thread while the scope of methodologies is ever-expanding. Core portfolio selection and optimization remain the frame for reference, while robust optimization, metaheuristics, machine learning, deep learning and sophisticated risk measures are current components of the modern research landscape.

In addition, publication activity reflects a distinct core-periphery pattern. A small, concentrated cluster of journals, universities and nations drive a significant portion of overall output, while author participation is widely distributed. Yet, this institutional and geographic concentration is mitigated by robust cross-border collaboration, ensuring that research moves forward through an interconnected global network rather than isolated clusters.

Most significantly, thematic tracking reveals a process of evolutionary maturation rather than wholesale replacement. The underlying portfolio optimization theme maintains its dominance throughout the study period, even as earlier concepts branch into more specialized domain areas. The primary thematic courses directly through core portfolio optimization across both timeframes, while surrounding themes extend into CvaR, diversification, metaheuristics, dynamic stochastic models and portfolio management.

In essence, modern literature has not discarded classic portfolio theory; instead, it has progressively enhanced how those core decision problems are mathematically modelled, empirically evaluated and applied in real-world contexts [15], [17].

6.2 Research Gaps

Synthesizing the field and methodology viewpoints, it appears there are very few papers at the intersection of robust/distributionally robust optimization and machine-learning integrated optimization techniques, although both methods individually have blossomed into established sub-topics. A review of the map of thematic development shows that most of the investment is focused toward topics with links to the development of robust portfolios instead of to those linked with machine learning clusters. However, machine learning/deep learning, remain closer to the thematic niche areas of learning-based diversification, reinforcement learning than the centre theme of robust optimization. Based on this observation, we find that further quantitative comparative studies to establish the relative performance among metaheuristic, robust, and learning-based portfolio formulation frameworks under uniform, realistically grounded assumptions of constraints: budget constraints, cardinality bounds, transaction costs and minimum lot sizes, within a single framework for evaluation, represent a still-open contribution, even though recent trends show diminishing publication growth and the studied family of methodology is, often in separation from other methods, to a greater extent isolated in existing works.

6.3 Future Research Agenda

6.3.1 Basis and Approach

This future research agenda is based on the bibliometric analysis of 708 portfolio-optimization documents from 2015 to 2026. The fitted growth curve projects approximately 973 documents in total; the 708-document corpus therefore represents about 72.8% of that estimate. The reported Bradford zones comprise 17 core sources accounting for 238 papers, 76 sources in the second zone, and 213 sources in the third zone - a total of 306 sources across the three zones. Lotka's Law indicates that approximately 87.1% of contributors in the analysis published one paper. Together, these patterns suggest a comparatively mature field with concentrated publication sources and a need for rigorous, transparent comparisons of new methods against classical optimization approaches.

6.3.2 Robustness, Uncertainty, and Distributionally Robust Optimization

[3] examine distributionally robust optimization using the Wasserstein metric and report performance guarantees and tractable reformulations; this paper has 1,603 citations in the analysed collection, approximately 2.58 times the 622 citations reported for the next-ranked paper. [20] provide a broad treatment of robust portfolio optimization and management. The thematic map places "robust optimization" among the basic themes in 2015–2019 rather than among the driving themes in 2020–2026, suggesting that robustness is often added to mean–variance models. Future research should test Wasserstein-based, moment-based, or box-uncertainty approaches against non-robust baselines using out-of-sample performance and realistic constraints, including transaction costs, cardinality limits, and estimation error.

6.3.3 Machine Learning Plus Prescriptive Optimization

Research on machine learning (21 keyword occurrences) and deep learning has expanded since approximately 2018. [5] provide an example of portfolio optimization using return prediction with machine-learning and deep-learning methods. [21] propose a deep-learning approach to hedging, while [4] develop the predict-then-optimize framework. Much of the literature still feeds predictions into a downstream optimizer; systematic comparisons between these two-step pipelines and integrated, end-to-end predict-and-optimize approaches remain limited. Future work should compare decision quality, interpretability, and out-of-sample performance across forecast horizons and asset classes.

6.3.4 Metaheuristics versus Exact and Robust Methods

A second cluster centres on genetic algorithms, recorded 30 times as a keyword, alongside population-based methods, metaheuristics, swarm methods, and particle swarm intelligence [2]. These methods may be useful when asset-count or transaction-cost constraints make exact algorithms difficult to apply. However, few studies compare these methods with exact or robust solvers on the same instances. Comparative studies should use common test instances and similar computational limits, evaluating solution quality, convergence, and scalability.

6.3.5 New Model of Computing: Hybrid and Quantum Techniques

Quantum computing and distributionally robust optimization are the newest topics, with work from ca. 2021, which still falls in the niche-theme area of the thematic map. The earliest work concerned [22] on quantum annealing for trading portfolios and [23] on quantum portfolio optimisation-proof-of-concept studies as opposed to generic alternative to state-of-the-art solvers. The quantum and hybrid quantum-classical algorithms should be evaluated on larger portfolios once hardware improves, in comparison to exact and metaheuristic solvers.

6.3.6 Sustainability and ESG-Integrated Optimization

Few words about sustainability are found in the dominant portfolio- and robust-optimization themes of the thematic map, even though ESG considerations are increasingly relevant for investors. An extension of the research is to integrate ESG scores, carbon exposure, or regulations directly into the objective function (rather than screening after the fact) and test for risk-adjusted return versus sustainability trade-offs in different economic conditions.

6.3.7 Broadening Geographical, Institutional, and Author Scope

China (363 papers), the US (180), and India (109) are the main prolific countries, but China is more central in collaboration networks with the US, Hong Kong, Singapore, India, and the UK. Because most authors only publish once [11], we are not sure how representative these few countries and authors are. Future work should extend the methodology to less-studied countries and focus on continuous collaborative research rather than one-time publications.

6.3.8 Outline for the Future Research Agenda

Robustness & uncertainty: test Wasserstein-based and distributionally robust models with classical mean-variance under estimation error and ambiguity set dimensions, based on out-of-sample performance and tractability, to identify these designs have real benefit. Predictive AI + prescriptive OR: benchmark ML/deep-learning predictions using integrated predict-then-optimize + hedging systems on standard cases, rated by accuracy + explanation ability to identify the place where end-to-end approaches outperform what is currently a decoupled pipeline.

Metaheuristics vs. exact/robust: compare genetic algorithms and swarm intelligence against exact and robust baselines on the same instances, in terms of solution quality, convergence, and scalability.

Benchmarking & validation: create comparable, well-understood datasets and transparent, document constraints tested across markets and timeframes and on criteria of comparability and reproducibility.

Emerging computation: benchmark quantum annealing, gate-based, and hybrid algorithms against classical solvers on QUBO problems, evaluated for solution quality and scalability. Sustainability/ESG integration: incorporate ESG scores and carbon exposure into multi-objective and robust optimization techniques, based on risk-return and ESG performance criteria, taking sustainability from the keyword to the modelled optimization target. Geographic & institutional diversification – scale up

comparative collaborative research to underserved markets on external-validity benchmarks to evaluate the generalisability of China/USA/India led research elsewhere.

6.3.9 Future Direction

This bibliometric study recommends new research directions in this field and calls for systematic comparisons of the existing methods instead of developing metaheuristics, robust optimizers or AI methods in every new paper. All these approaches such as robust optimization, metaheuristics, ML, and quantum computing have been extensively studied; however, rigorous comparisons on realistic, common problems are still missing. Future studies will figure out when these approaches provide true advantage over classical approaches and under fair benchmarks while sharing data and collaboration to form cumulative knowledge across time and space.

7. CONCLUSION

This study was done to map the structure and trajectory of portfolio optimization research published between 2015 and 2026 through a bibliometric analysis of 708 documents, and to identify where the field's future research effort might be most productively directed. Taken as a whole, the findings point to a field that has reached a stage of maturity: publication output rose through an initial growth phase from 2015 to around 2017, remained comparatively high and stable for several years afterward, and has since entered a marked decline, with output in the final years of the study period falling well below its earlier peak. Because the estimated total output the field is converging toward is only modestly larger than what has already been published, this pattern of publication activity is consistent with a research area whose expansion is slowing rather than one still in an early, fast-growing stage. Citation activity followed a related but distinct pattern. It continued to grow for some years after publication volume had already begun to fall. This divergence most likely reflects the additional time older papers have had to accumulate citations rather than any decline in the quality or relevance of more recent work.

The distribution of this output across journals, authors, institutions and countries reveals a consistent core-periphery structure. A relatively small number of journals, led by outlets such as Expert Systems with Applications and the European Journal of Operational Research, account for a disproportionate share of published articles, a concentration confirmed by Bradford's Law analysis showing that a core group of journals produced roughly a third of all papers in the corpus while the remaining output was spread thinly across a much larger set of sources. Author-level analysis shows an even sharper concentration at the level of individual contribution. The overwhelming majority of authors in the field appear only once in the record, consistent with Lotka's Law, even though a small set of researchers publish repeatedly. A similar pattern holds geographically, where China, the United States and India account for the largest shares of published work, with China in particular standing out both for its overall volume and for its centrality in international co-authorship networks.

At the conceptual level, the analysis confirms that classical portfolio optimization and portfolio selection remain the dominant organizing ideas of the field, with related concepts being the most prominent extensions built around that core. Thematic mapping across the earlier and later halves of the study period reinforces this picture of continuity rather than replacement. Portfolio optimization itself shows the strongest link between the two periods, while related themes such as robust portfolio optimization, dynamic stochastic modelling and various risk measures have branched outward from that stable centre into more specialized sub-domains. Machine learning and deep learning, despite their visibility in recent discussions of the field, remain positioned at the margins of this thematic structure rather than at its core. They appear among the emerging and still-developing themes rather than among the well-established motor themes that anchor the literature.

Read together, these patterns support a central conclusion that runs through the study's proposed future research agenda: much of what appears as innovation in portfolio optimization research consists of established techniques, whether robust optimization, metaheuristic search or machine-learning-based prediction, is being layered onto classical models as incremental additions, rather than being tested systematically against those classical models or against each other under shared, realistic conditions.

The study's future research agenda sets out seven directions along which more comparison-oriented work could be pursued. These include testing Wasserstein-distance-based and other distributionally robust formulations against classical mean-variance optimization under realistic estimation error, ambiguity and constraint conditions; benchmarking end-to-end predict-and-optimize machine learning pipelines against conventional two-step forecasting-then-optimization approaches across different forecast horizons and asset classes; comparing metaheuristic solvers directly against exact and robust methods on common problem instances under matched computational limits; developing standardized, shared datasets and constraint definitions to make results across studies comparable; extending early, proof-of-concept work on quantum and hybrid quantum-classical optimization to larger, more realistic portfolio problems as the underlying hardware matures; embedding environmental, social and governance considerations directly into the optimization objective rather than applying them as a post-hoc screen; and broadening the geographic and institutional base of the field's collaborative research beyond the small number of countries that currently dominate its output.

As with any bibliometric study, these conclusions are shaped by the boundaries of the method and the corpus on which it rests.

The analysis is limited to documents indexed in Scopus between 2015 and 2026 and retrieved through a specific, keyword-based search strategy; work published outside this window, indexed in other databases, or using terminology not captured by the search queries would not be reflected in the resulting patterns. Citation-based indicators such as the local H-index and average citation counts are also inherently sensitive to the age of a publication, meaning that the apparent lower impact of very recent work should be interpreted as a reflection of limited accumulation time rather than of diminished quality. Similarly, the institutional and country-level findings describe patterns of publication output rather than direct measures of scientific influence or research quality, and should be read accordingly. Within these boundaries, however, the study offers a structured account of how portfolio optimization research has developed over the past decade and where its comparative, benchmark-driven gaps remain. What emerges most clearly is that the field does not appear to be short of new methods; robust optimization, metaheuristic search, machine learning and even early-stage quantum computing techniques have each been explored in depth. What the field appears lacks is a sustained programme

of fair, shared-data comparison capable of establishing, under realistic and consistent conditions, when each of these approaches genuinely improves on the classical foundations from which it departs. Addressing that gap, more than the introduction of newer individual techniques, is what this study identifies to be the most likely direction for future research in portfolio optimization.

REFERENCES

- ADDIN Mendeley Bibliography CSL_BIBLIOGRAPHY [1] F. J. Fabozzi, H. M. Markowitz, P. N. Kolm, and F. Gupta, "Portfolio Selection," in *The Theory and Practice of Investment Management*, no. 60, Wiley, 2011, pp. 45–78. doi: 10.1002/9781118267028.ch3.
- [2] O. Ertenlice and C. B. Kalayci, "A survey of swarm intelligence for portfolio optimization: Algorithms and applications," *Swarm Evol. Comput.*, vol. 39, pp. 36–52, 2018, doi: 10.1016/j.swevo.2018.01.009.
- [3] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, no. 1–2, pp. 115–166, 2018, doi: 10.1007/s10107-017-1172-1.
- [4] A. N. Elmachtoub and P. Grigas, "Smart 'Predict, then Optimize,'" *Manage. Sci.*, vol. 68, no. 1, pp. 9–26, 2022, doi: 10.1287/mnsc.2020.3922.
- [5] Y. Ma, R. Han, and W. Wang, "Portfolio optimization with return prediction using deep learning and machine learning," *Expert Syst. Appl.*, vol. 165, 2021, doi: 10.1016/j.eswa.2020.113973.
- [6] B. Hanafi and M. Ali, "Research evolution in security involved for quantum computing and machine learning through scientometric review analysis," *Discov. Networks*, vol. 2, no. 1, 2026, doi: 10.1007/s44354-026-00029-y.
- [7] C. W. Belter, "Bibliometric indicators: Opportunities and limits," *J. Med. Libr. Assoc.*, vol. 103, no. 4, pp. 219–221, 2015, doi: 10.3163/1536-5050.103.4.014.
- [8] B. C. Brookes, "'Sources of information on specific subjects' by S.C. Bradford," *J. Inf. Sci.*, vol. 10, no. 4, pp. 173–175, Apr. 1985, doi: 10.1177/016555158501000406.
- [9] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *J. Informetr.*, vol. 11, no. 4, pp. 959–975, 2017, doi: 10.1016/j.joi.2017.08.007.
- [10] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 46, pp. 16569–16572, 2005, doi: 10.1073/pnas.0507655102.
- [11] A. Lotka, "The frequency distribution of scientific productivity Author (s): Alfred J . Lotka Source : Journal of the Washington Academy of Sciences , Vol . 16 , No . 12 (June 19 , 1926), Published by : Washington Academy of Sciences Stable URL : <https://www.jst.org>," *J. Washingt. Acad. Sci.*, vol. 16, no. 12, pp. 317–323, 1926.
- [12] M. Norton, V. Khokhlov, and S. Uryasev, "Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation," *Ann. Oper. Res.*, vol. 299, no. 1–2, pp. 1281–1315, 2021, doi: 10.1007/s10479-019-03373-1.
- [13] Y. Zhang, X. Li, and S. Guo, "Portfolio selection problems with Markowitz's mean–variance framework: a review of literature," *Fuzzy Optim. Decis. Mak.*, vol. 17, no. 2, pp. 125–158, 2018, doi: 10.1007/s10700-017-9266-z.
- [14] Z. X. Loke, S. L. Goh, G. Kendall, S. Abdullah, and N. R. Sabar, "Portfolio Optimization Problem: A Taxonomic Review of Solution Methodologies," *IEEE Access*, vol. 11, pp. 33100–33120, 2023, doi: 10.1109/ACCESS.2023.3263198.
- [15] P. Xidonas, R. Steuer, and C. Hassapis, "Robust portfolio optimization: a categorized bibliographic review," *Ann. Oper. Res.*, vol. 292, no. 1, pp. 533–552, 2020, doi: 10.1007/s10479-020-03630-8.
- [16] B. Chen, J. Zhong, and Y. Chen, "A hybrid approach for portfolio selection with higher-order moments: Empirical evidence from Shanghai Stock Exchange," *Expert Syst. Appl.*, vol. 145, 2020, doi: 10.1016/j.eswa.2019.113104.
- [17] A. Gunjan and S. Bhattacharyya, "A brief review of portfolio optimization techniques," *Artif. Intell. Rev.*, vol. 56, no. 5, pp. 3847–3886, 2023, doi: 10.1007/s10462-022-10273-7.
- [18] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field," *J. Informetr.*, vol. 5, no. 1, pp. 146–166, Jan. 2011, doi: 10.1016/J.JOI.2010.10.002.
- [19] S. Kilianová and D. Ševčovič, "Dynamic intertemporal utility optimization by means of Riccati transformation of Hamilton–Jacobi–Bellman equation," *Jpn. J. Ind. Appl. Math.*, vol. 36, no. 2, pp. 497–519, 2019, doi: 10.1007/s13160-019-00349-3.
- [20] W. C. Kim, J. H. Kim, and F. J. Fabozzi, "Robust Optimization Software," 2015, Wiley. doi: 10.1002/9781118797358.ch11.
- [21] H. Buehler, L. Gonon, J. Teichmann, and B. Wood, "Deep hedging," *Quant. Financ.*, vol. 19, no. 8, pp. 1271–1291, 2019, doi: 10.1080/14697688.2019.1571683.
- [22] G. Rosenberg, P. Haghnegahdar, P. Goddard, P. Carr, K. Wu, and M. L. De Prado, "Solving the Optimal Trading Trajectory Problem Using a Quantum Annealer," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 6, pp. 1053–1060, 2016, doi: 10.1109/JSTSP.2016.2574703.
- [23] D. Venturelli and A. Kondratyev, "Reverse quantum annealing approach to portfolio optimization problems," *Quantum Mach. Intell.*, vol. 1, no. 1–2, pp. 17–30, 2019, doi: 10.1007/s42484-019-00001-w.