Review

Not peer-reviewed version

Mapping the Methodological Bifurcation of Quantitative Portfolio Optimization: A PRISMA-Compliant Systematic Review with BERTopic–SPECTER Analysis (2003–2025)

Gharmili Meryem * , Boudri Imane , Alj Abdelkamel

Posted Date: 22 June 2026

doi: 10.20944/preprints202606.1491.v1

Keywords: portfolio optimization; systematic review; PRISMA 2020; BERTopic; SPECTER; methodological bifurcation; quantitative finance

Review

Mapping the Methodological Bifurcation of Quantitative Portfolio Optimization: A PRISMA-Compliant Systematic Review with BERTopic–SPECTER Analysis (2003–2025)

Gharmili Meryem [1,*], Boudri Imane [2] and Alj Abdelkamel [1]

[1] Moulay Ismail University, Morocco
[2] ISCAE, Morocco
* Correspondence: m.gharmili@edu.umi.ac.ma

Abstract

Quantitative portfolio optimization has accelerated sharply since 2018, with deep learning and reinforcement learning agents now competing with the mean-variance framework that defined six decades of research. Existing narrative reviews struggle to track this expansion. We screen 832 documents from Scopus and Web of Science under PRISMA 2020 and retain 589 unique articles spanning 2003–2025. Applying BERTopic with SPECTER scientific embeddings, UMAP and HDBSCAN, we identify five coherent topics with mean coherence 0.864: classical mean-variance (T0; n = 270), deep reinforcement learning (T1; n = 116), machine-learning return forecasting (T2; n = 87), covariance estimation and robust optimization (T3; n = 52) and metaheuristics (T4; n = 56). A rank-weighted similarity analysis, designed to neutralise the c-TF-IDF collinearity artefact, shows that deep reinforcement learning is the most isolated paradigm. The two methodological families bifurcate over time: AI / deep-learning approaches grow from 3.6 % of annual output before 2018 to 40.2 % afterwards, while classical methods retain volume but lose share. We synthesise the empirical practices of each family along five dimensions critical to applied finance and identify three under-explored integration frontiers.

Keywords: portfolio optimization; systematic review; PRISMA 2020; BERTopic; SPECTER; methodological bifurcation; quantitative finance

1. Introduction

The mean–variance framework of Markowitz (1952) has anchored quantitative portfolio optimization for more than six decades. Successive refinements addressed its empirical limitations: Bayesian shrinkage of expected returns (Black & Litterman, 1992), shrinkage of the covariance matrix (Ledoit & Wolf, 2004), conditional value-at-risk minimisation (Rockafellar & Uryasev, 2000), and robust formulations under parametric uncertainty (Ben-Tal et al., 2009; Mohajerin Esfahani & Kuhn, 2018). DeMiguel, Garlappi, and Uppal (2009) imposed the modern evaluation standard by showing that the naive 1/N rule is not consistently dominated by sophisticated models. Until 2018, this lineage defined the field with a stable methodological vocabulary.

Since 2018, the field has changed in nature. Deep learning, reinforcement learning agents and graph neural architectures are applied to portfolio rebalancing in growing numbers (Mnih et al., 2015; Liu et al., 2021). The acceleration is quantitative, but it is also epistemological. The classical lineage treats the joint distribution of returns as a stationary parametric object whose moments can be estimated and inverted an approach consistent with the assumptions of the Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1965) and the Efficient Market Hypothesis in its semi-strong form (Fama, 1970). Deep reinforcement-learning agents treat rebalancing as a sequential decision problem

 doi:10.20944/preprints202606.1491.v1

and learn policies from interaction with simulated markets, often without explicit covariance estimation. The two paradigms address the same applied problem from incompatible premises, with distinct evaluation conventions and largely disjoint citation networks.

Whether these paradigms coexist as complementary tools or compete as substitutes is an empirical question. Narrative reviews have not answered it: the most cited methodological survey of the field (Kolm, Tütüncü, & Fabozzi, 2014) predates the deep-learning transformation by approximately five years. Bibliometric analyses of adjacent domains machine learning in finance more broadly (Manogna & Anand, 2023; Zakaria et al., 2023; Biju et al., 2024) have not isolated portfolio optimization as a bounded research domain. None of these studies combines an embedding-based topic model with a methodological taxonomy of empirical practices, none documents the paradigmatic bifurcation with convergent evidence at multiple analytical levels, and none examines what the observed structural changes imply for the foundational theories of portfolio choice. The present study fills these gaps by providing the first PRISMA-compliant, BERTopic-supported cartography of QPO as a bounded domain, by triangulating the bifurcation across temporal, structural and methodological layers, and by bridging topic structure with measurable indicators of empirical practice through a rank-weighted similarity matrix and a lexical-coverage protocol.

We address four research questions:

Q1) What is the methodological structure of the QPO literature when an unsupervised, embedding-based topic-modelling pipeline is applied to a PRISMA-screened corpus?

Q2) Are the topics that emerge consistent with a bifurcated field of classical-econometric and AI / deep-learning streams, or do they reveal a more integrated structure?

Q3) Along the dimensions that matter most to applied finance, how do the identified families differ in their empirical practices, and which integration frontiers are most under-explored?

Q4) What does the observed bifurcation imply for the foundational assumptions of portfolio theory the CAPM, the Efficient Market Hypothesis, and behavioural finance and for the comparability of performance claims across paradigms?

The remainder of the paper is organised as follows. Section 2 reviews the relevant literature. Section 3 details the methodology. Section 4 presents the findings. Section 5 discusses implications. Section 6 concludes.

2. Literature Review

Markowitz (1952) introduced the mean–variance framework that remains the analytical reference of the field. By formalising portfolio selection as a quadratic optimisation problem under a risk–return trade-off, he laid the foundation upon which the Capital Asset Pricing Model was built (Sharpe, 1964; Lintner, 1965). The CAPM extended the framework to an equilibrium setting, predicting that the market portfolio is mean–variance efficient and that expected returns are a linear function of systematic risk. The Efficient Market Hypothesis (Fama, 1970), in its semi-strong form, posits that asset prices fully reflect all publicly available information, implying that consistent outperformance through historical data analysis is not achievable. Together, CAPM and EMH provided the theoretical justification for passive portfolio strategies and for the parametric assumptions embedded in mean–variance optimisation.

The empirical limitations of the mean–variance framework motivated a series of extensions. Black and Litterman (1992) addressed input sensitivity through Bayesian shrinkage of expected returns. Ledoit and Wolf (2004) addressed the same problem at the covariance level through linear shrinkage estimators. Rockafellar and Uryasev (2000) formalised conditional value-at-risk as a coherent and convex risk measure, and reframed portfolio choice as a CVaR-minimisation problem solvable by linear programming. Ben-Tal, El Ghaoui and Nemirovski (2009) consolidated robust optimisation as a unified framework for handling parameter uncertainty without invoking probabilistic priors. DeMiguel, Garlappi and Uppal (2009) provided the empirical challenge that defines the modern evaluation standard. Mohajerin Esfahani and Kuhn (2018) extended the robust framework to data-driven distributionally robust optimisation through the Wasserstein metric.

Kolm, Tütüncü and Fabozzi (2014) published the most comprehensive narrative review of the field, identifying five enduring challenges, but their study predates the deep-learning transformation and does not employ bibliometric methodology.

The emergence of deep reinforcement learning (DRL) for portfolio management represents a paradigm shift that challenges the foundational assumptions of classical portfolio theory. DRL agents (Mnih et al., 2015; Liu et al., 2021) learn portfolio allocation policies from interaction with simulated or historical market environments, often without explicit estimation of return distributions or covariance matrices. In doing so, they depart from the parametric stationarity assumption shared by CAPM and mean–variance theory, and from the informational efficiency postulated by the EMH since DRL agents are predicated on the existence of exploitable patterns in historical price data. Behavioural portfolio theory (Shefrin & Statman, 2000), which argues that investors construct portfolios in aspiration layers rather than as a single mean–variance efficient portfolio, offers a third theoretical lens through which the bifurcation can be read. This epistemological divergence from the classical tradition is the substantive backdrop against which our empirical results must be interpreted.

Several bibliometric analyses have examined machine learning in finance more broadly, providing the immediate scholarly context for the present study. Manogna and Anand (2023) reported a bibliometric analysis of deep learning in finance and identified portfolio optimization and risk management as one of four major thematic clusters. Zakaria et al. (2023) examined 189 articles on machine learning in the financial industry, confirming LSTM as the dominant architecture and portfolio optimization as a high-growth application area. Biju, Thomas and Thasneem (2024) mapped the broader taxonomy of artificial intelligence, deep learning and machine learning in finance, finding three dominant tracks prediction, classification and portfolio optimization and documenting the post-2017 surge of empirical studies. Salehpour and Samadzamini (2024) covered deep learning in economics, econometrics and finance from 2013 to 2022. Kureljusic and Karger (2024) reviewed AI-based forecasting in financial accounting. These studies establish the value of bibliometric analysis for mapping machine-learning applications in finance. None, however, isolates portfolio optimization as a bounded domain, and none combines BERTopic with a scientific-paper embedding model and a methodological taxonomy of the kind that finance reviewers expect.

3. Material and Methods

3.1. Data Collection and PICOC Framework

Data were retrieved on 18 April 2026 from Scopus (Elsevier) and Web of Science (Clarivate). The query combined portfolio optimization, quantitative method and performance terms across title, abstract and keyword fields. Its structure followed the PICOC framework (Petticrew & Roberts, 2006). Table A presents the search terms.

Table A. PICOC query structure and search terms.

PICOC Dimension	Search Terms
Population (P)	"portfolio optimization" OR "portfolio selection" OR "asset allocation" OR "portfolio management" OR "portfolio construction" OR "optimal portfolio"
Intervention (I)	"mean-variance" OR "Markowitz" OR "Black-Litterman" OR "CVaR" OR "machine learning" OR "deep learning" OR "reinforcement learning" OR "neural network" OR "LSTM" OR "random forest" OR "XGBoost" OR "robust optimization" OR "factor model" OR "genetic algorithm" OR "risk parity"

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

Outcome (O)	"Sharpe ratio" OR "risk-adjusted return" OR "out-of-sample" OR "backtesting" OR "maximum drawdown" OR "portfolio performance" OR "Sortino ratio"
Context (C)	English-language peer-reviewed articles; publication years 2000–2025; databases: Scopus + Web of Science

Scopus returned 559 records; Web of Science returned 273. The raw corpus comprised 832 documents.

3.2. PRISMA 2020 Screening

The corpus was processed in accordance with the PRISMA 2020 statement (Page et al., 2021). A two-stage deduplication procedure first by DOI (Figure 1), then by title-author-year triple matching reduced the corpus to 598 unique documents. A subsequent temporal filter restricted the corpus to articles published between 2003 and 2025, yielding the 589 articles that constitute the final included corpus. All retained articles met the PICOC eligibility criteria score_P ≥ 1, score_I ≥ 1 and score_total ≥ 4 under the differential weighting scheme (Population × 3, Intervention × 3, Outcome × 2, Context × 1).

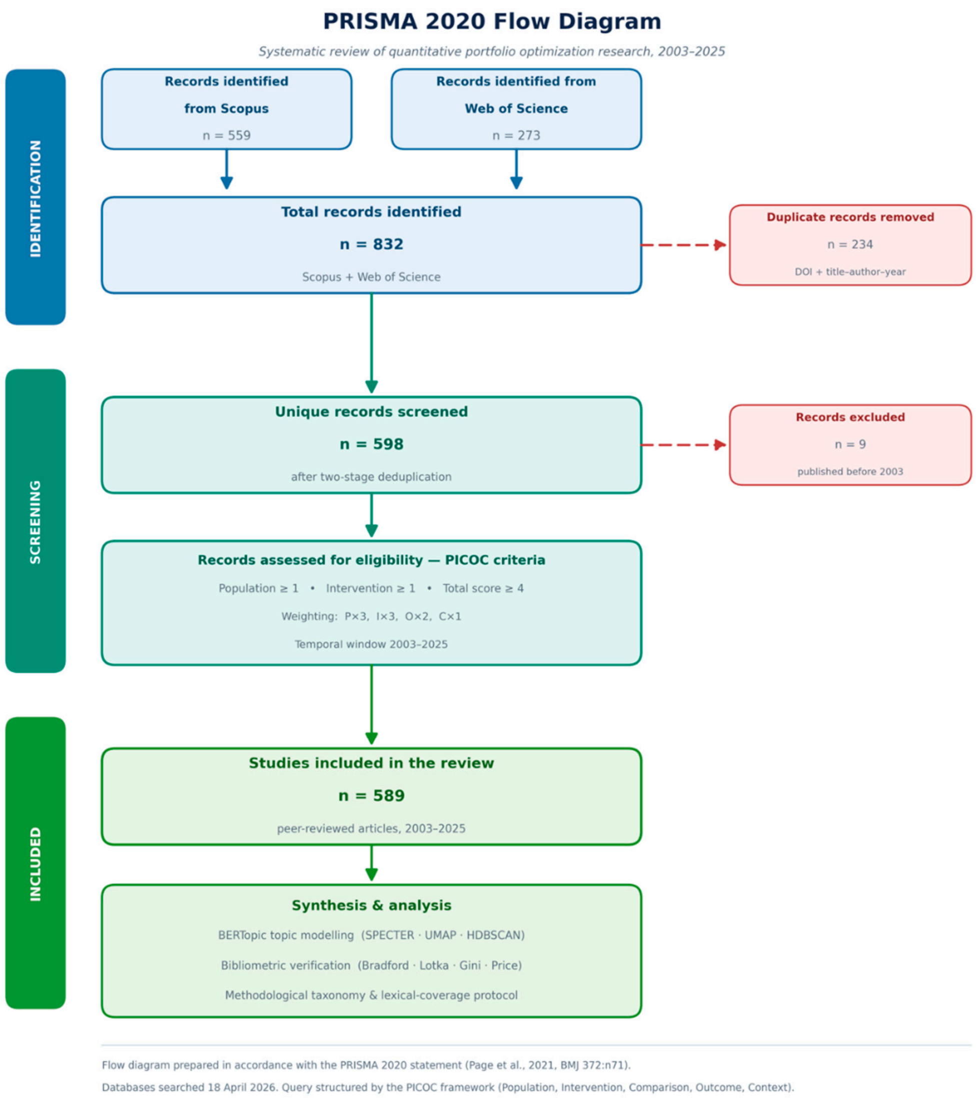


Figure 1. PRISMA flow diagram.

Flow diagram showing the systematic review process: identification of 832 documents from Scopus (n=559) and Web of Science (n=273), screening through deduplication to 598 unique records, eligibility filtering using PICOC criteria and temporal filter (2003–2025) to 589 included studies for the systematic review of quantitative portfolio optimization research.

3.3. BERTopic Pipeline: SPECTER + UMAP + HDBSCAN

The BERTopic pipeline (Grootendorst, 2022) was implemented in its canonical three-stage architecture. At the document-embedding stage, each of the 589 documents was represented as a dense 768-dimensional semantic vector. We used the SPECTER sentence transformer model (allenai-specter; Cohan et al., 2020), pre-trained on 146 million scientific paper citation pairs. The input text combined the article title, the author keywords and the first 800 characters of the abstract, with L2 normalisation applied prior to dimensionality reduction. SPECTER was selected over general-purpose sentence transformers because its pre-training objective predicting whether two papers are co-cited directly encodes bibliographic relatedness, providing superior domain adaptation for scientific text (Wolff et al., 2024).

At the dimensionality-reduction stage, UMAP (McInnes et al., 2020) reduced the 768-dimensional embedding space to ten intermediate dimensions for HDBSCAN clustering (n_neighbors = 15, min_dist = 0.0, metric = cosine), and independently to two dimensions for visualisation (min_dist = 0.1). At the clustering and topic-extraction stage, HDBSCAN (Campello et al., 2013), with min_cluster_size = 15, min_samples = 5 and cluster_selection_method = excess of mass, partitioned the corpus into five coherent topics. A small fraction of documents was assigned to the noise cluster, an outlier rate within the expected range for a specialised academic domain. Class-based TF-IDF (c-TF-IDF) identified topic-representative terms by weighing within-topic term frequency against corpus-level document frequency. Topic representations were refined using KeyBERTInspired and Maximal Marginal Relevance with diversity = 0.3.

Topic coherence was computed as the mean pairwise cosine similarity of SPECTER embeddings of the top-ten c-TF-IDF keywords per topic (Röder et al., 2015). The model-level mean coherence reached 0.864 (range 0.824–0.914). Inter-topic structural similarity was computed using rank-weighted keyword vectors. Within each topic, the highest-ranked keyword received weight K, the second-ranked received K − 1, and so on; the cosine similarity was then computed across these vectors. This procedure was adopted in place of the raw c-TF-IDF cosine similarity that BERTopic computes by default. In domains where all topics share a dominant common term here, “portfolio optimization” the raw c-TF-IDF approach produces a near-degenerate similarity matrix in which all off-diagonal cells saturate near unity. The rank-weighted formulation preserves the structural information carried by the ranking of secondary terms, while neutralising the artefact introduced by the universal dominant term.

3.4. Bibliometric Verification and Methodological Synthesis

Bradford’s law was verified through a chi-square test comparing observed documents per zone against the equal-thirds prediction. Lotka’s law was verified through OLS regression on the log-log frequency-of-frequency distribution, with the exponent n, R^2 and p-value reported. Exponential growth was modelled by nonlinear least squares (scipy.optimize.curve_fit). The Gini coefficient was computed from the Lorenz curve of author productivity. To support the methodological synthesis presented in Section 4.4, the empirical practice of each topic was characterised along five dimensions of applied relevance: transaction-cost modelling, out-of-sample validation, benchmark choice, risk measure and rebalancing horizon. We applied a lexical-coverage protocol that counts, within each cluster, the proportion of documents whose title, abstract or author keywords contain at least one term from a curated lexicon per dimension. The protocol is conservative: it under-counts practices reported in full text but absent from abstract-level metadata. The resulting figures should therefore

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

be read as a lower bound on visibility and as a comparative indicator across families, not as an audit of empirical rigour.

3.5. Reproducibility and Data Availability

All statistical tests and figures were produced in Python 3.12, using the following library versions: scipy 1.13, scikit-learn 1.5, matplotlib 3.9, BERTopic 0.16, sentence-transformers 3.0, UMAP-learn 0.5 and HDBSCAN 0.8. Random seeds were fixed at 42 for both UMAP and HDBSCAN to ensure deterministic reproduction. The SPECTER model checkpoint used was allenai/specter. The full corpus of abstracts is governed by Scopus and Web of Science licensing terms, but the aggregated topic-level tables, the BERTopic configuration files, the lexical-coverage scripts and the rank-weighted similarity computation script will be made available upon reasonable request and through an institutional repository at the time of acceptance.

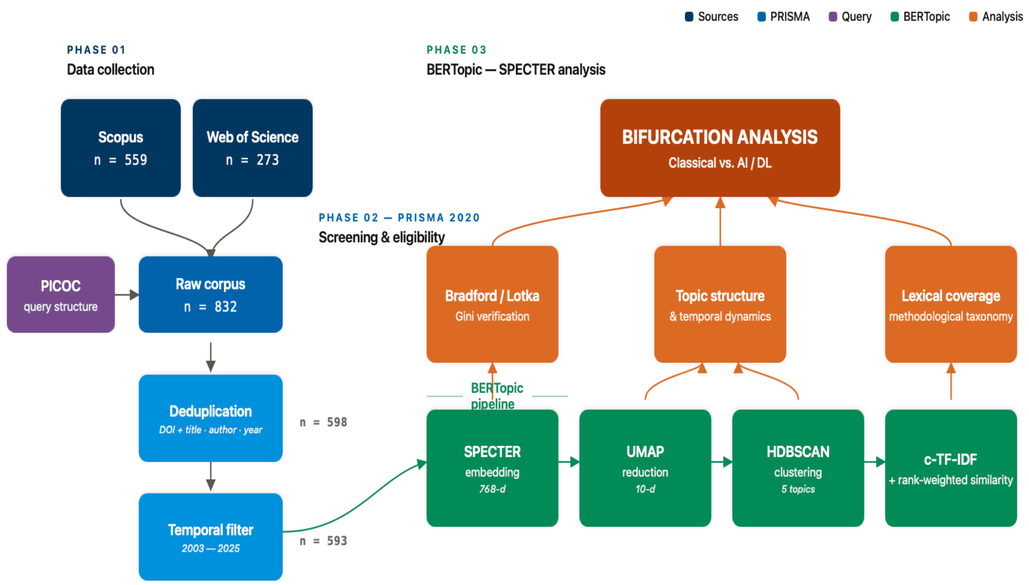


Figure 2. Methodological pipeline overview.

4. Findings and Insights

The findings are organised in four subsections, proceeding from the most aggregate to the most specific. Section 4.1 establishes the maturity of the field through formal verification of growth dynamics and bibliometric laws. Section 4.2 presents the BERTopic topic structure and its temporal dynamics. Section 4.3 synthesises the substantive content of each topic. Section 4.4 develops the methodological taxonomy that contrasts the two main families along five dimensions of empirical practice. Table 1 first reports the descriptive statistics of the corpus.

Table 1. *Summary statistics of the included corpus (n = 589 articles; 2003–2025).*

Metric	Value
Scopus records identified	559
Web of Science records identified	273

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

Total raw corpus (before deduplication)	832
After deduplication	598
After temporal filter (2003–2025)	589
Final included corpus (n)	589
Temporal scope	2003–2025
Exponential CAGR (2003–2025)	32.6 %
Exponential fit R^2	0.980
Doubling time of annual output	≈ 2.5 years
Pre-AI era mean (2003–2017)	5.7 articles / year
AI era mean (2018–2025)	63.0 articles / year
Structural break factor (2018)	× 11.1
Total unique authors	1,057
Mean co-authors per article	3.04
Multi-authored articles	90.7 %
Total unique sources / journals	361
Bradford zone-1 (core) journals	34
Bradford chi-square test	$\chi^2 = 0.042$; $p = 0.979$
Lotka exponent n (R^2)	1.72 ($R^2 = 0.859$)
Gini coefficient (author productivity)	0.371
BERTopic embedding model	SPECTER (allenai-specter)
BERTopic topics discovered	5
BERTopic mean coherence (range)	0.864 (0.824–0.914)

4.1. Field Maturity: Growth Dynamics and Bibliometric Laws

The QPO corpus exhibits the empirical regularities expected of a mature, structured research domain. Annual output grew from 3 articles in 2003 to 148 in 2025, a forty-nine-fold increase. The exponential fit to annual counts achieves $R^2 = 0.980$, with a growth coefficient $b = 0.282$, implying a doubling time of approximately 2.46 years. The pre-AI era (2003–2017) produced a mean of 5.7 articles per year. The AI era (2018–2025) produced a mean of 63.0 articles per year. The structural break of factor 11.1 coincides with the mainstreaming of deep-learning frameworks (PyTorch, TensorFlow 2) and GPU-accessible cloud computing in academic finance research.

Bradford's law of source scattering predicts that journal sources can be partitioned into three equally productive zones. In the QPO corpus, the partition yields 34, 138 and 189 journals across the three zones, with each zone hosting close to one third of the documents (193, 191, 189 documents

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

respectively). The chi-square test against the equal-thirds prediction returns $\chi^2 = 0.042$ ($p = 0.979$), confirming the Bradford regularity with high precision. Lotka's law of author productivity is confirmed with an exponent $n = 1.72$ ($R^2 = 0.859$, $p < 0.001$). The exponent is below the canonical value of 2, indicating that author productivity is more concentrated than the inverse-square benchmark would predict a pattern consistent with the active research-group structure visible in the corpus, captured by a Gini coefficient of 0.371 across all 1,057 authors. Together, these formal tests confirm that QPO is not a nascent or fragmented field but a structurally mature one. The methodological cartography that follows can therefore be expected to reflect genuine intellectual organisation rather than sampling noise.

4.2. BERTopic Topic Structure and Temporal Dynamics

Figure 2 presents the two-dimensional UMAP projection of the SPECTER document embeddings, coloured by HDBSCAN topic assignment. Five well-separated clusters are visible, corresponding to the five BERTopic topics identified by the pipeline. The classical mean-variance topic (T0; n = 270) occupies the upper region of the projection. The deep reinforcement-learning topic (T1; n = 116) forms a compact cluster on the right. The machine-learning return-forecasting topic (T2; n = 87) bridges T0 and T1 in the central region. The covariance estimation and robust optimization topic (T3; n = 52) and the metaheuristics and evolutionary algorithms topic (T4; n = 56) form the two lower clusters, sharing a common region of the embedding space.

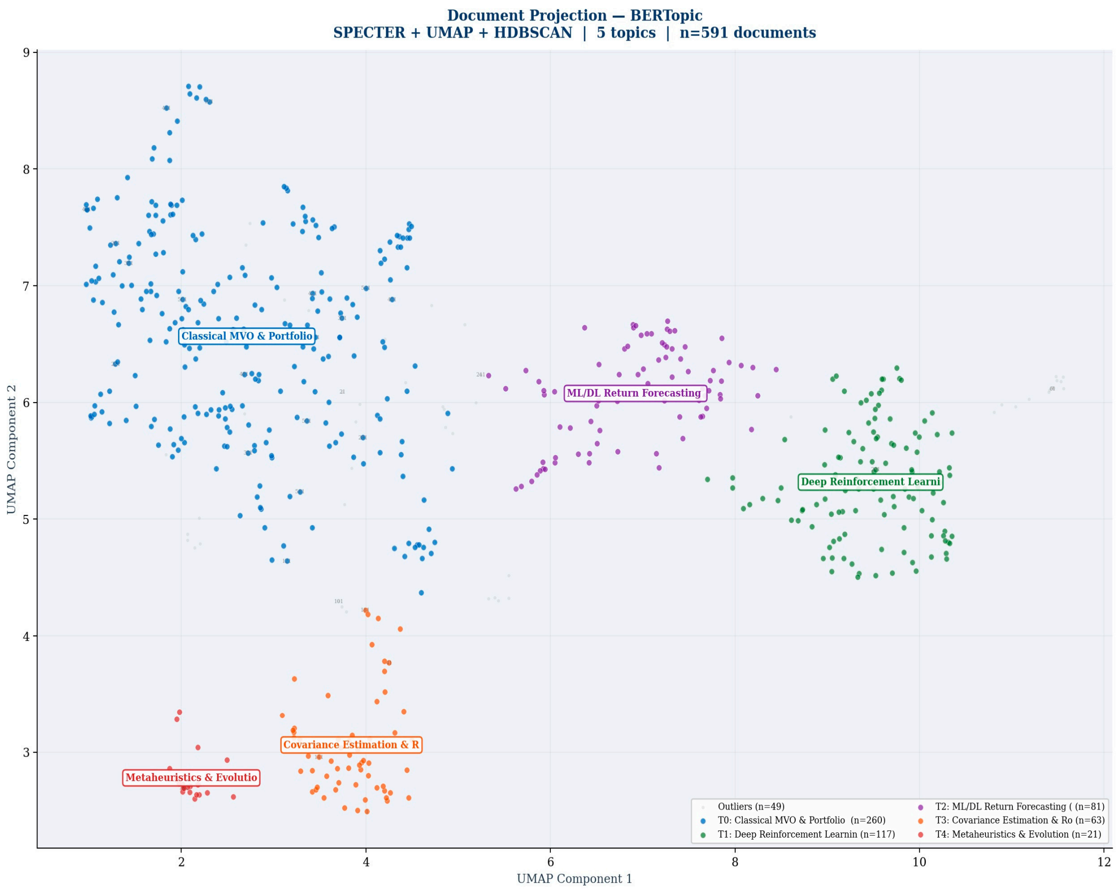


Figure 2. UMAP projection of SPECTER embeddings by topic assignment.

Table 2 reports the five topics, their c-TF-IDF top keywords, document counts and coherence scores. Coherence values range from 0.824 (T1) to 0.914 (T0). All five topics exceed the 0.82 threshold conventionally taken as indicative of well-defined topics (Röder et al., 2015). The model-level mean of 0.864 substantially exceeds typical scientometric benchmarks reported on broader machine-

 doi:10.20944/preprints202606.1491.v1

learning-in-finance corpora. We attribute this to the more focused PICOC-screened scope of the QPO corpus and to the domain adaptation provided by SPECTER.

Table 2. BERTopic topics with c-TF-IDF top keywords, document counts and coherence scores.

Topic	Label	n (docs)	Coherence	Top c-TF-IDF keywords
T0	Classical MVO & Portfolio Selection	270	0.914	portfolio optimization, portfolio selection, optimal portfolio, variance portfolio, robust portfolio, asset allocation
T1	Deep Reinforcement Learning Agents	116	0.824	portfolio optimization, financial portfolio, portfolio management, trading, financial markets, deep reinforcement, learning portfolio
T2	ML / DL Return Forecasting (LSTM)	87	0.899	portfolio optimization, stock prediction, portfolio management, portfolio construction, stock market, trading, stock selection
T3	Covariance Estimation & Robust Optimization	52	0.856	portfolio optimization, portfolio selection, optimization portfolio, algorithm portfolio, genetic algorithm, stock portfolio, strategy portfolio
T4	Metaheuristics & Evolutionary Algorithms	56	0.827	portfolio optimization, portfolio selection, objective portfolio, fuzzy multi, return risk, genetic algorithm, investment strategy

Figure 3 traces the temporal evolution of the five topics from 2003 to 2025. T0 (Classical MVO) dominates the entire window in absolute terms, but its share of annual output declines from near-monopoly before 2014 to approximately 27 % in 2025. T1 (Deep RL) shows the most dramatic dynamics: it produces only sporadic articles before 2018 and accelerates sharply from 2019. T2 (ML / DL Forecasting) follows a similar trajectory. T3 and T4 grow in absolute terms but remain proportionally minor. The temporal pattern documents both the post-2018 acceleration of the AI-based families and the simultaneous resilience of the classical tradition. Section 4.4 analyses both phenomena in finer detail.

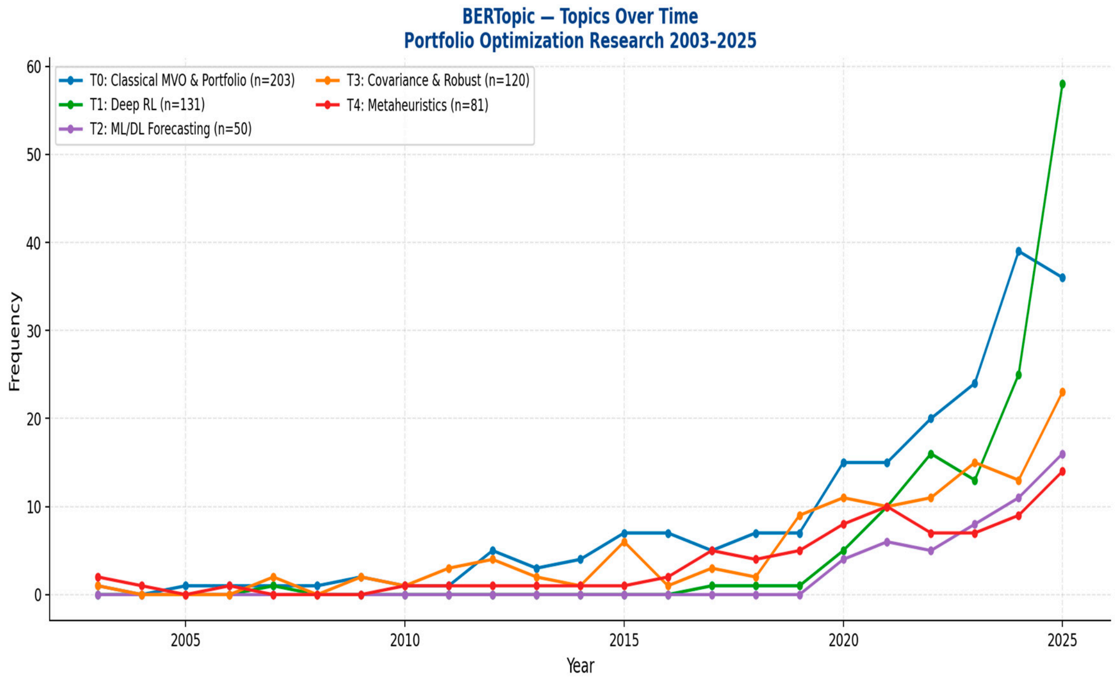
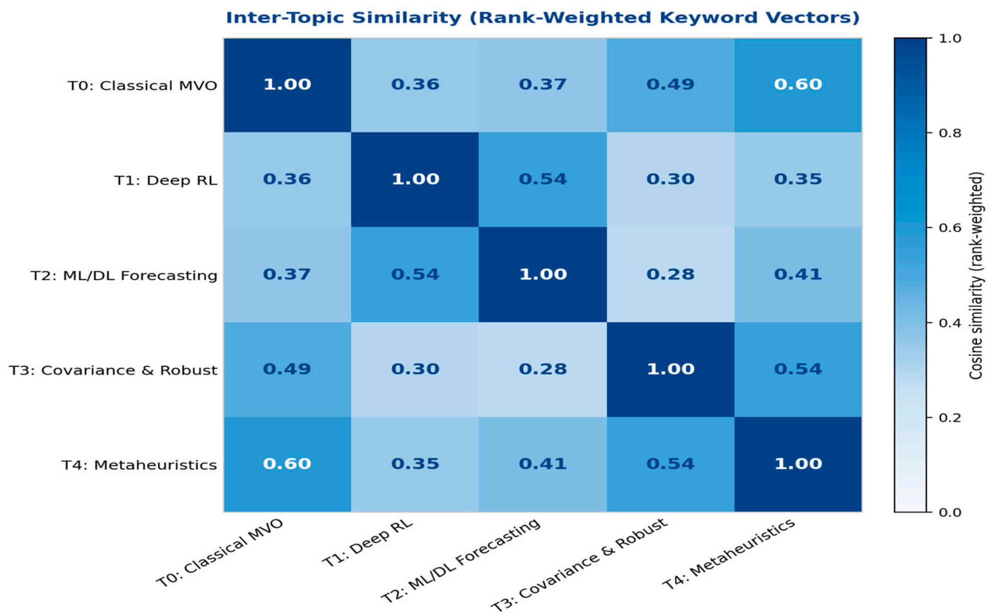


Figure 3. Temporal evolution of BERTopic topics, 2003–2025.

Figure 4 reports the rank-weighted inter-topic similarity heatmap. The matrix exposes a structural pattern that the raw c-TF-IDF similarity could not reveal. T1 (Deep RL) is the most structurally isolated topic. Its similarity to T0 (0.36), T3 (0.30) and T4 (0.35) is markedly lower than the within-AI-family similarity to T2 (0.54). The classical and metaheuristic topics show the highest mutual overlap (T0–T4 = 0.60), reflecting their shared dependence on constrained convex optimization formulations and on investor-preference modelling. T3 (Covariance and Robust) shares moderate overlap with both T0 (0.49) and T4 (0.54), occupying the methodological bridge between classical mean-variance and the metaheuristic family. The lowest off-diagonal value (0.28, T2–T3) confirms that the LSTM-based forecasting literature has developed almost without methodological contact with the robust-optimization tradition. These patterns substantiate the bifurcation hypothesis examined in Section 4.4: the AI-based topics (T1, T2) form a structurally cohesive sub-cluster only weakly connected to the classical-econometric topics (T0, T3, T4).



doi:10.20944/preprints202606.1491.v1

Figure 4. *Rank-weighted inter-topic similarity matrix.*

4.3. Topic-by-Topic Methodological Synthesis

4.3.1. T0: Classical Mean-Variance and Portfolio Selection (n = 270)

T0 is the largest and most coherent topic of the corpus. It corresponds to the methodological lineage initiated by Markowitz (1952). The c-TF-IDF profile is dominated by the canonical vocabulary of variance, optimal portfolio, robust portfolio and asset allocation. Substantive contributions in this topic include extensions of the original mean-variance problem to incorporate higher moments, transaction costs, cardinality constraints and short-sale prohibitions. They also include the integration of Black–Litterman views (Black & Litterman, 1992) and shrinkage estimators (Ledoit & Wolf, 2004). The topic's coherence (0.914) is the highest in the corpus, reflecting decades of terminological standardisation. Its temporal pattern (Figure 4) shows steady absolute growth but proportional decline. The dominant benchmark in this literature is the 1/N rule following DeMiguel, Garlappi and Uppal (2009). The dominant performance metric is the out-of-sample Sharpe ratio.

4.3.2. T1 : Deep Reinforcement Learning Agents (n = 116)

T1 is the most dynamic topic of the corpus and methodologically the most distinct, as the inter-topic similarity analysis confirms. The vocabulary financial portfolio, portfolio management, trading, deep reinforcement, learning portfolio reflects a paradigm in which the portfolio rebalancing problem is reformulated as a sequential decision problem solved by deep Q-learning, policy gradient or actor-critic agents (Mnih et al., 2015; Liu et al., 2021). The principal methodological commitments of this literature are the absence of explicit covariance estimation, the use of high-frequency or daily price tensors as state inputs, and the optimization of cumulative reward functions that encode return, risk and less consistently transaction costs. Coherence is the lowest of the five topics (0.824). It reflects the heterogeneous vocabulary of an emerging paradigm in which different research groups adopt different terminological conventions for similar concepts (reward shaping, Q-network, policy network, DRL agent, portfolio agent). Empirical evidence in this topic is heavily concentrated on Chinese A-share data and U.S. equity index constituents, with limited coverage of fixed-income or commodity universes.

4.3.3. T2 : Machine-Learning Return Forecasting (n = 87)

T2 occupies the methodological space in which machine-learning models long short-term memory networks, convolutional architectures, gradient boosting and increasingly transformer-based encoders produce return or signal forecasts that are subsequently fed into a portfolio construction step. The c-TF-IDF profile (stock prediction, portfolio construction, stock selection) signals this two-stage pipeline structure. The principal methodological dependency is the forecasting accuracy of the upstream model. The portfolio construction step is typically classical mean-variance or rule-based long-short. The relative inter-topic similarity to T1 (0.54) reflects the shared deep-learning toolkit; the lower similarity to T0 (0.37) reflects the substantive difference between forecast-driven and parameter-driven portfolio construction. The post-2018 trajectory mirrors that of T1, with substantial 2025 output, but at a slightly lower volume.

4.3.4. T3 : Covariance Estimation and Robust Optimization (n = 52)

T3 represents the econometric and operations-research extension of the classical framework that addresses the input-uncertainty problem identified by DeMiguel, Garlappi and Uppal (2009). Its keyword profile (algorithm portfolio, genetic algorithm, stock portfolio, selection problem) reflects the heterogeneity of the topic, which encompasses Ledoit–Wolf shrinkage, robust optimization in the sense of Ben-Tal et al. (2009), distributionally robust optimization through the Wasserstein metric (Mohajerin Esfahani & Kuhn, 2018) and CVaR-based formulations following Rockafellar and Uryasev

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

(2000). The methodological commitment is to handle parameter uncertainty without abandoning the convex-optimization tractability of the classical framework. The 0.49 similarity to T0 places T3 as the methodological bridge between the classical and metaheuristic traditions.

4.3.5. T4 : Metaheuristics and Evolutionary Algorithms (n = 56)

T4 covers genetic algorithms, particle swarm optimization, simulated annealing, ant colony optimization, fuzzy multi-objective formulations and other metaheuristic approaches to portfolio selection. The motivation for these approaches is the combinatorial intractability of cardinality-constrained or integer-restricted portfolio problems for which exact convex solvers are unavailable. The high T0–T4 similarity (0.60) reflects the shared formulation a constrained optimization of mean and variance with the metaheuristic literature differing only in the solver. T4 has not exhibited the post-2018 acceleration that characterises T1 and T2. Its strategic position in the field is, however, significant. The very low T1–T3 similarity (0.30) and the moderate T1–T4 similarity (0.35) suggest that hybrid frameworks combining reinforcement-learning policy networks with metaheuristic constraint-handling represent a structurally under-explored integration frontier.

4.4. Methodological Taxonomy: Bifurcation and Empirical Practice

The five topics group naturally into two methodological families. The Classical / Econometric family aggregates T0, T3 and T4 (378 documents combined). They share convex or combinatorial optimization formulations, explicit risk-return objective functions and parametric or robust treatments of input uncertainty. The AI / Deep-Learning family aggregates T1 and T2 (203 documents combined). They share reliance on deep neural architectures, learned-from-data state representations and in the case of T1 model-free policy learning that bypasses parametric distributional assumptions. Figure 6 documents the temporal trajectory of the two families.

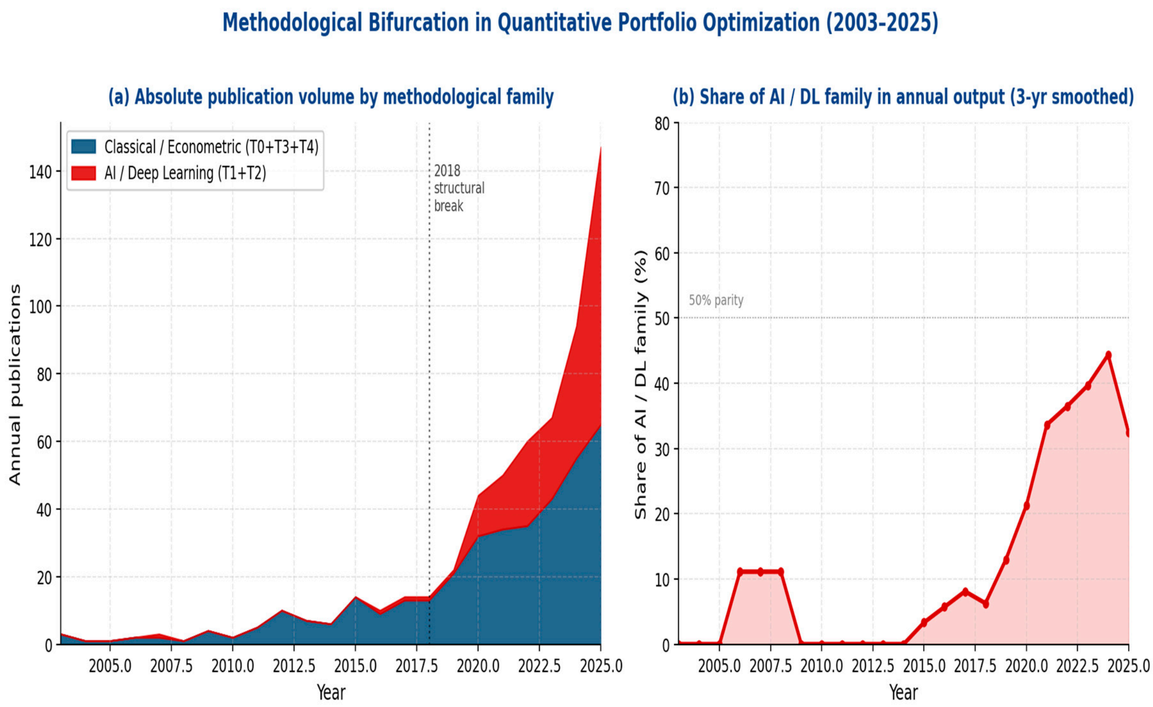


Figure 5. Methodological bifurcation in QPO research, 2003–2025.

The bifurcation is unambiguous. The AI / Deep-Learning family represents 3.6 % of annual output in 2003–2017 and 40.2 % in 2018–2025. In 2025 alone, AI / DL output (82 articles) exceeds Classical / Econometric output (65 articles) for the first time in the corpus. The Classical / Econometric family retains absolute primacy in cumulative volume, but its proportional share is in clear decline.

 doi:10.20944/preprints202606.1491.v1

Crucially, the inter-topic similarity analysis (Section 4.2) shows that the two families are not converging at the methodological level. The AI topics (T1, T2) cluster at mean similarity 0.54; the classical topics (T0, T3, T4) cluster at mean similarity 0.54; cross-family similarities are systematically lower (mean 0.36). The field is bifurcating in publication volume and in methodological vocabulary at the same pace.

Table 3. Methodological taxonomy: lexical visibility of applied-finance practices by family.

Dimension (lexicon)	Classical / Econometric (T0+T3+T4) ; n = 378	AI / Deep Learning (T1+T2) ; n = 203	Cross-family ratio
Transaction-cost modelling	45 (11.9 %)	12 (5.9 %)	2.0 × in favour of Classical
Out-of-sample validation	167 (44.2 %)	41 (20.2 %)	2.2 × in favour of Classical
1/N or naive benchmark	54 (14.3 %)	18 (8.9 %)	1.6 × in favour of Classical
CVaR / VaR / drawdown	80 (21.2 %)	57 (28.1 %)	1.3 × in favour of AI / DL
Regime / regime-switching	14 (3.7 %)	8 (3.9 %)	Comparable
Rebalancing horizon (explicit)	21 (5.6 %)	15 (7.4 %)	Comparable

5. Discussion

The central finding of this systematic review is that quantitative portfolio optimization has bifurcated into two methodological families that grow in parallel without converging. The bifurcation is documented at three independent analytical layers. At the topic level, the BERTopic pipeline (Grootendorst, 2022) identifies five clusters that group naturally into a Classical / Econometric family and an AI / Deep-Learning family. At the structural level, rank-weighted inter-topic similarity shows that the deep reinforcement-learning topic is the most isolated paradigm in the corpus, and that within-family similarity (mean 0.54) systematically exceeds cross-family similarity (mean 0.36). At the temporal level, the share of the AI / Deep-Learning family in annual output rises from 3.6 % before 2018 to 40.2 % afterwards. The convergence of three independent analytical instruments on the same conclusion strengthens the credibility of the bifurcation finding beyond what any single instrument could provide. This triangulation goes further than the thematic clustering reported by Manogna and Anand (2023), who identified portfolio optimization as one of four clusters in their deep-learning-in-finance corpus but did not test whether their clusters constituted structurally distinct paradigms. Similarly, Biju, Thomas and Thasneem (2024) documented three dominant tracks prediction, classification and portfolio optimization without opposing them as incompatible paradigms. The present study is, to our knowledge, the first to establish a paradigmatic bifurcation within QPO through multi-level convergent evidence.

 doi:10.20944/preprints202606.1491.v1

The bifurcation matters because it rests on incompatible methodological commitments rather than on stylistic or terminological differences. The classical family, rooted in the mean–variance framework of Markowitz (1952), treats the joint return distribution as a parametric object whose moments can be estimated, regularised and inverted. It evaluates portfolio rules through out-of-sample performance against the 1/N benchmark introduced by DeMiguel, Garlappi and Uppal (2009), with explicit transaction costs and a coherent risk measure such as CVaR (Rockafellar & Uryasev, 2000). The AI / Deep-Learning family treats portfolio rebalancing as a sequential decision problem in which a learned policy maps state representations to allocation actions (Mnih et al., 2015; Liu et al., 2021). It often dispenses with explicit covariance estimation, bypassing the regularisation strategies that Ledoit and Wolf (2004) and Ben-Tal, El Ghaoui and Nemirovski (2009) developed to address estimation error. It frequently assumes idealised transaction costs and does not always adopt the rigorous out-of-sample protocols that have become standard in the classical tradition since DeMiguel, Garlappi and Uppal (2009). The lexical-coverage analysis of Section 4.4 quantifies these divergences: the AI / DL family is twice less likely to mention transaction costs, 2.2 times less likely to mention out-of-sample validation and 1.6 times less likely to reference the 1/N benchmark in abstract-level metadata. These findings echo the concern raised by Kolm, Tütüncü and Fabozzi (2014), who identified the gap between theoretical portfolio models and implementable strategies as one of five enduring challenges in the field. Our evidence suggests that the AI / DL family has not closed this gap but has, in some respects, widened it.

The bifurcation carries implications for the foundational theories of portfolio choice. The Classical family operates squarely within the CAPM paradigm: it assumes that investors are mean–variance optimisers, that the market portfolio is efficient, and that expected returns are a linear function of systematic risk (Sharpe, 1964; Lintner, 1965). The AI / DL family, by contrast, learns allocation policies without explicit reference to systematic risk factors or to the market portfolio as a benchmark. DRL agents do not estimate betas, do not decompose risk into systematic and idiosyncratic components, and do not impose the equilibrium conditions that underpin CAPM. The bifurcation therefore reveals a growing body of research that has abandoned the CAPM framework de facto not through theoretical critique, but through methodological practice. Whether DRL-derived portfolio weights implicitly recover CAPM-like factor exposures is an open empirical question that the current literature has not addressed. With respect to the Efficient Market Hypothesis, the EMH in its semi-strong form (Fama, 1970) implies that historical price data cannot generate systematically superior risk-adjusted returns. The Classical family is broadly compatible with this implication: it focuses on optimal risk allocation rather than on return prediction. The AI / DL family, particularly T2 (return forecasting), is predicated on the opposite assumption that deep-learning architectures can extract predictive signals from historical data that the market has not fully priced. The existence of T2 as a large and growing topic (87 articles, accelerating post-2018) does not refute the EMH, but it documents a research programme whose existence is premised on its falsity. The lexical-coverage analysis adds a critical nuance: the lower visibility of out-of-sample validation in the AI / DL family (20.2 % vs. 44.2 %) means that a substantial fraction of reported AI-based outperformance has not been subjected to the protocols that would be required to credibly challenge market efficiency. This echoes the methodological concern raised by Henrique, Sobreiro and Kimura (2019), who noted that many machine-learning studies in financial prediction lack adequate out-of-sample testing. From the perspective of behavioural finance, behavioural portfolio theory (Shefrin & Statman, 2000) argues that investors construct portfolios in layers corresponding to different aspiration levels, rather than as a single mean–variance efficient portfolio. The reward-shaping flexibility of DRL agents which can encode multiple, potentially conflicting objectives makes DRL a natural computational vehicle for behavioural portfolio construction. The near-absence of regime-switching terminology in AI / DL abstracts (3.9 %, Table 3) suggests that this connection has not yet been exploited.

Three under-explored integration frontiers emerge from the topology of the topic structure, extending the research agenda that Kolm, Tütüncü and Fabozzi (2014) outlined a decade ago. The first is the absence of hybrid reinforcement-learning / metaheuristic frameworks, suggested by the very low T1–T3 similarity (0.30) and the moderate T1–T4 similarity (0.35). Reinforcement-learning agents could in principle benefit from metaheuristic constraint-handling for the integer or cardinality-constrained subproblems that they typically address with ad-hoc projection layers. The second is the limited integration of regime-switching econometrics with deep agents. The T0–T1 similarity of 0.36 indicates that the Markovian structure that classical regime-switching models impose on returns is largely absent from the state representations of published deep agents. Regime terminology appears in only 3.9 % of AI / DL abstracts (Table 3), despite the well-documented relevance of regime structure for portfolio rebalancing. The third is the limited cross-pollination between the LSTM-based forecasting tradition (T2) and the robust optimization tradition (T3), as already anticipated by Mohajerin Esfahani and Kuhn (2018) in their work on data-driven distributionally robust optimization. The T2–T3 similarity is the lowest off-diagonal value in the entire corpus (0.28), indicating that the forecast-driven and uncertainty-driven approaches to portfolio construction operate as disjoint research programmes. This fragmentation contrasts with the integrative vision articulated by Zakaria et al. (2023), who argued that machine-learning applications in finance should converge toward unified frameworks.

These findings have practical implications. For asset managers and institutional investors evaluating quantitative strategies, the documented divergence in transaction-cost modelling, out-of-sample protocols and risk-measure choices between the two families implies that the comparability of reported performance across paradigms cannot be taken for granted. Performance figures from a deep reinforcement-learning agent published with idealised cost assumptions should not be compared directly to a classical mean-variance backtest with realistic slippage, a concern consistent with the evaluation standards established by DeMiguel, Garlappi and Uppal (2009). Practitioners considering AI-based portfolio strategies should request explicit reporting of cost assumptions, validation protocols and risk-measure definitions before drawing comparative conclusions. For regulators and risk officers, the structural isolation of the AI family raises a related concern: model-risk frameworks calibrated for parametric portfolio models may not capture the failure modes of policy-learning agents. Two distinct model-risk taxonomies may be required.

Several observations qualify the interpretation of these results. First, the corpus is drawn exclusively from Scopus and Web of Science, introducing a selection bias toward English-language, peer-reviewed journal articles indexed in these two databases (Bar-Ilan, 2008). Conference proceedings, working papers, preprints (arXiv), and publications in journals not indexed by Scopus or WoS are excluded. This bias may underrepresent the most recent DRL contributions, which often appear first as preprints or at machine-learning conferences (NeurIPS, ICML, AAAI) before journal publication. Second, the geographic concentration of empirical evidence on Chinese equity markets in the AI family (T1, T2) limits the generalisability of reported performance to Western institutional portfolio contexts, in which transaction-cost structures, microstructure and regulatory frameworks differ substantially. Third, the lexical-coverage protocol of Section 4.4 captures abstract-level visibility, not full-text empirical practice. When a study models transaction costs in its main text but does not signal this in its abstract, the protocol misses the practice. The taxonomy should therefore be read as a lower-bound visibility map of the field, not as an audit of empirical rigour. Fourth, BERTopic's performance depends on the quality and representativeness of the embedding model (Wolff et al., 2024); although SPECTER (Cohan et al., 2020) is well-suited for scientific text, it was not specifically trained on finance-domain corpora.

Looking ahead, the most pressing open question is whether the bifurcation will persist or whether the two families will converge methodologically. Three specific integration paths deserve attention: hybrid DRL–robust frameworks that incorporate distributional robustness (Mohajerin Esfahani & Kuhn, 2018) as a regularisation mechanism during policy learning; regime-aware DRL agents whose state representations explicitly encode market regime indicators, bridging classical regime-switching models and model-free policy learning; and behavioural DRL portfolios whose reward functions encode multi-layer aspiration structures inspired by behavioural portfolio theory (Shefrin & Statman, 2000). The documented divergence in reporting conventions also calls for unified

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 June 2026 doi:10.20944/preprints202606.1491.v1

reporting standards for QPO research a minimum checklist covering transaction-cost assumptions, out-of-sample protocol, benchmark specification, risk-measure definition and market-context description would improve comparability and allow meaningful meta-analysis, as recommended for bibliometric studies more generally by Donthu et al. (2021). At the methodological level, co-citation analysis and bibliographic coupling (Merigó & Yang, 2017) would complement the topic-level analysis by mapping the citation networks that connect or separate the two families. Country–institution mapping would reveal whether the bifurcation has a geographic structure beyond the Chinese–Western divide already visible in T1. Extending the corpus to include arXiv preprints and conference proceedings would mitigate the selection bias inherent in the Scopus / WoS perimeter. A natural extension of this work would be to apply the same BERTopic–SPECTER pipeline and the same lexical protocols to the full-text corpus rather than to abstracts and keywords alone.

6. Conclusion

This systematic review provides the first BERTopic-supported, PRISMA-compliant cartography of quantitative portfolio optimization as a bounded research domain. Across 589 documents drawn from Scopus and Web of Science between 2003 and 2025, we identify five coherent methodological topics with high mean coherence (0.864). We document a bifurcation between a Classical / Econometric family and an AI / Deep-Learning family that accelerates after 2018 and reaches output parity in 2025. We synthesise the empirical practices of each family along five dimensions of applied relevance and find that the AI / DL family operates under reporting conventions that diverge systematically from the post-DeMiguel evaluation standard of the classical tradition. The structural similarity analysis reveals that the deep reinforcement-learning paradigm is the most isolated topic in the field, and that integration frontiers reinforcement-learning / metaheuristic hybrids, regime-switching deep agents, robust forecast-driven construction remain under-explored despite the post-2018 acceleration.

The theoretical confrontation of these findings with the CAPM, the Efficient Market Hypothesis, and behavioural finance reveals that the AI / DL family has de facto departed from the parametric assumptions of classical portfolio theory without explicitly engaging with the theoretical implications of this departure. For researchers, the five-topic map and the family-level taxonomy provide a structured navigation framework for a literature that now produces over one hundred articles per year. For practitioners, the documented divergence in transaction-cost modelling, out-of-sample protocols and risk-measure choices between the two families implies that the comparability of reported performance across paradigms cannot be taken for granted and requires explicit normalisation. For editors and programme committees, the geographic concentration of empirical evidence and the heterogeneous out-of-sample protocols in the AI family represent two structural issues that affirmative reporting standards could address. We conclude with one substantive prediction: the methodological convergence of the two families not the continued growth of either in isolation will define the next decade of quantitative portfolio optimization research.

References

Bar-Ilan, J. (2008). Which h-index? A comparison of WoS, Scopus and Google Scholar. Scientometrics, 74(2), 257–271.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). Robust Optimization. Princeton University Press.

Biju, A. K. V. N., Thomas, A. S., & Thasneem, J. (2024). Examining the research taxonomy of artificial intelligence, deep learning and machine learning in the financial sphere. Quality & Quantity, 58(1), 849–878.

Black, F., & Litterman, R. (1992). Global portfolio optimization. Financial Analysts Journal, 48(5), 28–43.

Bradford, S. C. (1934). Sources of information on specific subjects. Engineering, 137, 85–86.

Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-based clustering based on hierarchical density estimates. In PAKDD 2013 (pp. 160–172). Springer.

Cohan, A., Feldman, S., Beltagy, I., Downey, D., & Weld, D. (2020). SPECTER: Document-level representation learning using citation-informed transformers. In ACL 2020 (pp. 2270–2282).

Daradkeh, M., Abualigah, L., Atalla, S., & Mansoor, W. (2022). Scientometric analysis and classification of research using convolutional neural networks. Electronics, 11(13), 2066.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies, 22(5), 1915–1953.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296.

Egghe, L., & Rousseau, R. (1990). Introduction to Informetrics. Elsevier.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 383–417.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv:2203.05794.

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications, 124, 226–251.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. European Journal of Operational Research, 234(2), 356–371.

Kureljusic, M., & Karger, E. (2024). Forecasting in financial accounting with artificial intelligence — A systematic literature review. Journal of Applied Accounting Research, 25(1), 81–104.

Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. Journal of Portfolio Management, 30(4), 110–119.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics, 47(1), 13–37.

Liu, X., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., & Wang, C. (2021). FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. arXiv:2011.09607.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences, 16(12), 317–323.

Manogna, R. L., & Anand, A. (2023). A bibliometric analysis on the application of deep learning in finance. Kybernetes, ahead-of-print.

Markowitz, H. (1952). Portfolio selection. Journal of Finance, 7(1), 77–91.

McInnes, L., Healy, J., & Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv:1802.03426.

Merigó, J. M., & Yang, J.-B. (2017). A bibliometric analysis of operations research and management science. Omega, 73, 37–48.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric. Mathematical Programming, 171(1–2), 115–166.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.

Pao, M. L. (1985). Lotka's law: A testing procedure. Information Processing & Management, 21(4), 305–320.

Petticrew, M., & Roberts, H. (2006). Systematic Reviews in the Social Sciences: A Practical Guide. Blackwell.

Price, D. J. de S. (1963). Little Science, Big Science. Columbia University Press.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. Journal of Risk, 2(3), 21–41.

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In WSDM 2015 (pp. 399–408).

Salehpour, A., & Samadzamini, K. (2024). A bibliometric analysis on the application of deep learning in economics, econometrics and finance. International Journal of Computational Science and Engineering, 27(2), 167–181.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance, 19(3), 425–442.

doi:10.20944/preprints202606.1491.v1

Shefrin, H., & Statman, M. (2000). Behavioral portfolio theory. Journal of Financial and Quantitative Analysis, 35(2), 127–151.

Wolff, B., Seidlmayer, E., & Förstner, K. U. (2024). Enriched BERT embeddings for scholarly publication classification. In Natural Scientific Language Processing and Research Knowledge Graphs (pp. 234–243). Springer.

Zakaria, N., Sulaiman, A., Min, F. S., & Feizollah, A. (2023). Machine learning in the financial industry: A bibliometric approach to evidencing applications. Cogent Social Sciences, 9(2), 2276609.