# Artificial Intelligence in Supply Chain Optimization: A Systematic Review of Machine Learning Models, Methods, and Applications

December 1, 2025

Nipun Galande[1], Kimiya Mohammadi Jozani[1], İ. Esra Büyüktahtakın[1*]

[1]Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, 24061, USA.

*Corresponding author(s). E-mail(s): esratoy@vt.edu;

**Abstract**

Modern supply chains face mounting uncertainty and scale, motivating the integration of Artificial Intelligence (AI) and Machine Learning (ML) with mathematical optimization to enable robust and adaptive decisions. We present a systematic review of 199 articles on tangible supply chains, categorizing how ML is used—primarily for parameter estimation and for solution generation—and proposing a taxonomy that links ML roles to problem types and optimization formulations. The review surfaces consistent patterns (e.g., reinforcement learning in logistics), identifies underexplored areas (e.g., ML-aided reformulation and learned uncertainty for robust/DRO), and introduces a research framework to orient future studies. A dedicated subsection examines how these integrations relate to supply chain viability and resilience.

**Keywords:** Supply chain, Optimization, Machine learning, Logistics

## 1 Introduction

The design and optimization of supply chains are fundamental pillars of the modern global economy, directly influencing business competitiveness, sustainability, and social resilience. In an increasingly complex and volatile world, the ability to make optimal decisions regarding production, inventory, and logistics is no longer a competitive advantage but a necessity for survival [63]. For decades, the field of Operations Research has provided the foundational tools for this endeavor through mathematical optimization, enabling organizations to solve complex planning and scheduling problems. The challenges in this domain are multi-faceted, involving multi-scale, multi-objective [1], and multi-player considerations. Addressing these challenges through rigorous optimization is vital for achieving operational excellence and navigating the intricate trade-offs between cost, service, and sustainability. However, having the right tool does not guarantee the successful optimization of supply chain operations. It is important to know how to use the ML tool according to the best practices. Through the strategic implementation of ML algorithms in supply chain optimization, a supplier's competitive advantage can be determined.

In parallel to mathematical optimization, the recent boom in data has led to the introduction of a new direction for solving supply chain problems. The modern supply chain produces streams of information from sensors, transactions, and external sources. To make the best use of these data, organizations are turning to ML models for every aspect, ranging from prediction to improving solutions. ML algorithms can identify the underlying patterns and break down complex data for prediction, such as the prediction of demand [184], without being explicitly programmed. Similarly, ML is also being used to solve large-scale supply chain problems [2] that take days to solve using traditional methods. In more advanced applications, ML models can even help in the formulation of optimization problems themselves by learning complex data-driven constraints or objective functions. The rapid involvement of ML models suggests a paradigm shift, moving supply chain problems from traditional models to incorporating learning methods.

While optimization provides the framework for decision-making and machine learning offers data-driven insights, the true potential for innovation lies in their integration. Individually, both methodologies address only one part of the challenge. The true power of these methods can be unlocked when the predictive power of ML models is used to enhance the mathematical optimization formulations. This relationship, often called the "predict-then-optimize" paradigm [45], allows decision-makers to build optimization models that are not based on static historical averages but instead are powered by accurate real-time predictions about uncertain variables. ML techniques such as reinforcement learning can learn policies to solve sequential decision problems that are too complex for traditional optimization solvers [215]. This symbiotic relationship between ML and mathematical optimization allows the creation of more dynamic, resilient, and effective supply chain solutions that can adapt to the complexities of the real world.

The growing interest in this integrated approach has led to a surge in academic research. Although several reviews have surveyed the applications of ML in supply chain management [140, 195] or focused on specific optimization problems [12, 173], there remains a significant gap. Even recent comprehensive reviews, such as Vlachos and Reddy [200], which analyzed ML applications and barriers through 2023, primarily focus on management frameworks such as the 4Vs (Volume, Variety, Variation, Visibility) rather than the technical integration with optimization models. The current literature lacks a comprehensive and systematic review that focuses specifically on the intersection of machine learning and mathematical optimization for tangible supply chains—that is, supply chain networks concerned with physical production, storage, and transportation of goods and materials. We adopt this specific focus to clearly distinguish our work from the growing body of research on intangible supply chains, such as those in software delivery, cloud computing, or digital service provision [209]. In addition to the importance of these digital domains, the challenges in Supply Chains (SC) are unique, involving physical constraints such as warehouse capacity, vehicle routing, production scheduling, and inventory spoilage. By focusing on tangible systems, this review aims to provide targeted insight for industries such as manufacturing, logistics, and healthcare, where the optimization of physical assets is paramount. As our analysis of publication trends reveals, the volume of research in this specific integrated domain has grown exponentially in recent years, as seen in Figure 1, making a structured synthesis of this work timely.

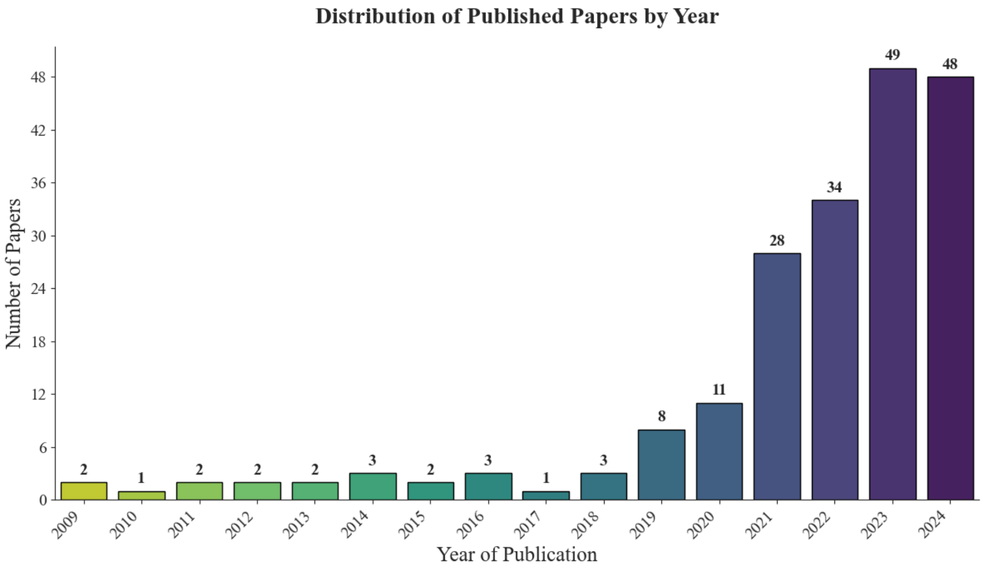


**Fig. 1** Annual distribution of the 199 journal publications included in this review

This paper fills this gap by systematically mapping and analyzing research at the intersection of machine learning, mathematical optimization, and SC. Our primary research objective is to provide a comprehensive overview of how these fields are being integrated to solve complex physical problems. To achieve this, we classify the relevant literature across several key dimensions: the type of academic contribution (e.g., model vs. method), the specific supply chain domain (e.g., logistics, manufacturing), the mathematical formulation (e.g., MIP, stochastic), different problem categories (e.g., *VRP*, *lot-sizing*), machine learning algorithms used (e.g., random forest, reinforcement learning), the specific role of ML within the optimization framework (e.g., parameter estimation, solution generation), and the operational scale (e.g., small-scale vs. large-scale).

The main contributions of this paper are as follows.

1. We provide a systematic review of 199 articles published between 2009 and 2024, focusing exclusively on the integration of ML and optimization in tangible supply chains.
2. We offer a detailed, multi-faceted classification of the literature, identifying dominant trends such as the prevalence of Reinforcement Learning for solving Vehicle Routing Problems and the extensive use of ML for demand forecasting.
3. We analyze the relationships between different classifications to uncover deep insights, such as the strong correlation between dynamic routing problems and learning-based solvers.
4. We identify critical research gaps, particularly in the under-explored areas of model reformulation and the application of these techniques to humanitarian and circular supply chains.
5. We synthesize how integrated ML and optimization enable supply chain viability and resilience, through anticipation, adaptive control, structural redesign, recovery planning, and robustness, and provide a compact capability to method mapping to guide practice and research.
6. We propose a structured research framework to guide future studies, offering a clear pathway for researchers to define problems, formulate models, and select appropriate ML integration strategies in supply chain management.

The remainder of this paper is organized as follows. Section 2 details the systematic review methodology employed to identify and screen the relevant literature. Section 3 presents the results of our analysis, structured by our classification scheme. Section 4 provides a discussion of the results, where we explore the deeper connections between these classifications. We analyze the relationships between the ML models used and the types of optimization problems they solve, the supply chain domains to which they are applied, and the purpose of their integration. This discussion is designed to highlight key trends, identify research gaps, and set the stage for Section 5, where we propose a research framework to guide future work. The paper concludes with Section 6, which summarizes our key findings and outlines promising future research directions. For completeness, an acronym glossary is provided in Appendix A.

## 2 Methodology

In this section, we provide a brief overview of the methodology used to implement a systematic review of the literature. We started with a set of keywords and, after an initial review, we finalized the keywords used for the search. This systematic review adheres to the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement [162]. PRISMA is a standardized reporting framework designed to enhance the transparency, completeness, and reproducibility of systematic reviews and meta-analyses. It provides a 27-item checklist that covers all major components of a review, including the title, abstract, rationale, objectives, eligibility criteria, information sources, search strategy, selection process, data collection, risk of bias, synthesis methods, and limitations.

All studies included in this review were identified by a comprehensive search of the electronic databases Scopus and Web of Science using a predefined search strategy, followed by screening and eligibility evaluation based on defined inclusion and exclusion criteria. Duplicates were removed and all screening was conducted independently by two reviewers with disagreements resolved by consensus.

A central feature of the PRISMA framework is the flow diagram, which depicts the study identification, screening, eligibility assessment, and the final inclusion process. This four-phase diagram ensures rigorous documentation of the selection process, including the number of records retrieved from databases, records excluded after title/abstract screening, full-text articles assessed for eligibility, and the final number of studies included in the qualitative and/or quantitative synthesis.

Traditional narrative reviews, while valuable, can fail to provide a comprehensive and reproducible overview of a rapidly evolving field. They may be influenced by the ' existing knowledge of the authors and can unintentionally overlook important clusters of research. To ensure a thorough, transparent, and unbiased analysis, we adopted a systematic literature review (SLR) methodology. The primary goal of this review is to systematically investigate and map the research landscape at the intersection of machine learning (ML), mathematical optimization, and their applications within SC.

To provide a clear answer to this broad research question, we broke it down into the following eight sub-research questions (SRQs). (SRQ 1) What types of academic contribution are the most common in this field? Are they focused on proposing new models, developing novel solution methods, or synthesizing existing knowledge? (SRQ 2) In what specific SC domains, such as logistics, manufacturing, or healthcare, are these integrated techniques—methods that combine both a machine learning algorithm and a mathematical optimization model— most frequently applied? (SRQ 3) On what operational scale, small, large, or generalizable "any" scale, are these SC problems being addressed? (SRQ 4) What are the dominant mathematical optimization formulations, such as Integer Programming or Combinatorial Optimization, that are being employed by these papers? (SRQ 5) Which specific classes of optimization problems, such as the *Vehicle Routing Problem (VRP)* or *Inventory Control*, are receiving the most attention? (SRQ 6) What machine learning models and techniques are most commonly used in these integrated frameworks, where a learning component is used to inform, solve, or define an optimization problem? (SRQ 7) How is machine learning integrated into the optimization models? Is it used for parameter estimation, to accelerate the solution process, or in another capacity? (SRQ 8) In what ways do integrated ML–optimization approaches enable supply chain viability and resilience? The purpose of these SRQs is to analyze the existing literature to provide deep insight into the current state of knowledge and to identify potential gaps and opportunities for future research.

Our search and selection process was designed to be rigorous and replicable, and we defined our search window from 2009 to September 2024 to capture the literature from the last 15 years, a period marked by significant advances in both machine learning algorithms and computational optimization capabilities. The iterative refinement of our search query, conducted in the Scopus database, is illustrated in Figure 3. We began with a broad query that yielded an unmanageable number of 42,087 papers. To focus the search, we narrowed the scope to more specific supply chain keywords, which reduced the set of results to 5,145 articles. However, a preliminary screening revealed a high concentration of irrelevant papers on non-tangible, IT-related topics. Consequently, our final query utilized the Scopus feature to exclude these keywords. During this process, we also made a methodological decision to refine our scope regarding nature-inspired optimization. We excluded papers focused on traditional metaheuristics, such as genetic algorithms or particle swarm optimization, when presented without an explicit data-driven learning component. Most of these papers often use the term “AI” superficially without a true learning component. This final targeted search strategy was executed in September 2024, resulting in a definitive set of 2,253 articles for screening.

The complete screening and filtering protocol, which follows the PRISMA guidelines, is detailed in the flow diagram shown in Figure 2. The 2,253 papers identified from our final query were subjected to a complete title and abstract selection. To mitigate bias, each article was independently reviewed by the authors, and inclusion was voted on based on our predefined criteria. This process narrowed the selection to 396 articles that were deemed potentially relevant. These articles were then subjected to a full-text review, in which we assessed their eligibility in detail. This final stage of screening excluded another 197 articles, resulting in a final corpus of 199 studies for our in-depth analysis.

In subsequent sections of this paper, we present the results of our synthesis. The analysis is structured to directly answer the seven sub-research questions posed earlier. We classify the 199 selected papers according to their type of paper, type of SC, type of optimization formulation, category of optimization problem, ML algorithm utilized, scale of SC problems, and the role of ML in the optimization model. This multi-faceted classification provides a comprehensive overview of the field and allows us to highlight key trends, gaps, and promising directions for future research.

# 3 Analysis of Results

In this section, we present the results of our analysis from different perspectives, from bibliographic analysis to optimization-focused analysis and AI-focused analysis. The analyzes are organized in six main subsections.

## 3.1 Types of Papers

We categorized the reviewed literature by its primary contribution (Model, Method, and Review). For a complete classification table and a detailed analysis of these paper types, please refer to Appendix B.

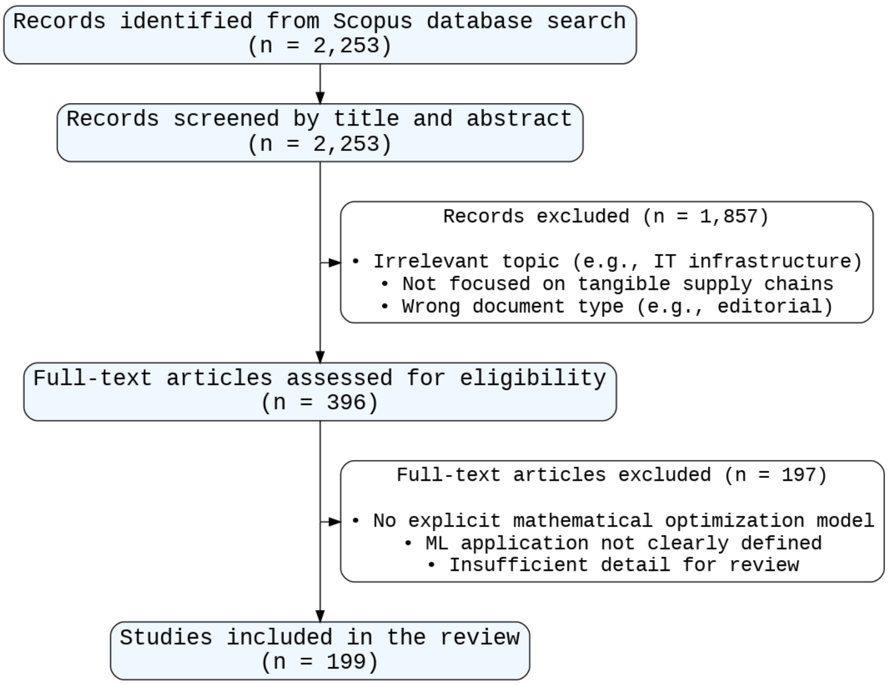


**Fig. 2** Flowchart representing the PRISMA approach followed for the survey

## 3.2 Types of Tangible Supply Chain Domains

The synergy between machine learning and mathematical optimization is not limited to a single industry. Its principles are being applied across a wide range of SC domains, each with unique characteristics, constraints, and objectives. This subsection classifies the reviewed literature according to the specific type of SC to which the models are applied. Our analysis shows that the field is heavily dominated by applications in logistics and transportation, which serve as a foundational domain for routing and network design problems. However, there is a growing and significant body of research in specialized sectors such as Food, Manufacturing, and Healthcare, demonstrating the versatility of these integrated methods. A summary of the analysis of the type of SC is presented in Table 1.

Transportation & Logistics is the most prominent category, with 117 papers dedicated to solving problems related to the movement and storage of goods. The central focus of this domain is the *Vehicle Routing Problem (VRP)* and its many variations. A large portion of the research is dedicated to developing new ML-based solution methodologies, including hybrid neural frameworks [138], learning-based heuristics [74, 212], and automated algorithm design [231]. Many studies tackle specific variants of *VRP*, such as capacitated *VRP* [46, 47], *VRP* with time windows [31, 252], multidepot *VRP* [6, 7, 115], and heterogeneous fleet *VRP* [121, 165]. Modern logistics challenges are also addressed, including last-mile and urban delivery [36, 174], drone integration [21, 100, 108, 119], and the route of electric vehicles [152, 169]. Other research focuses

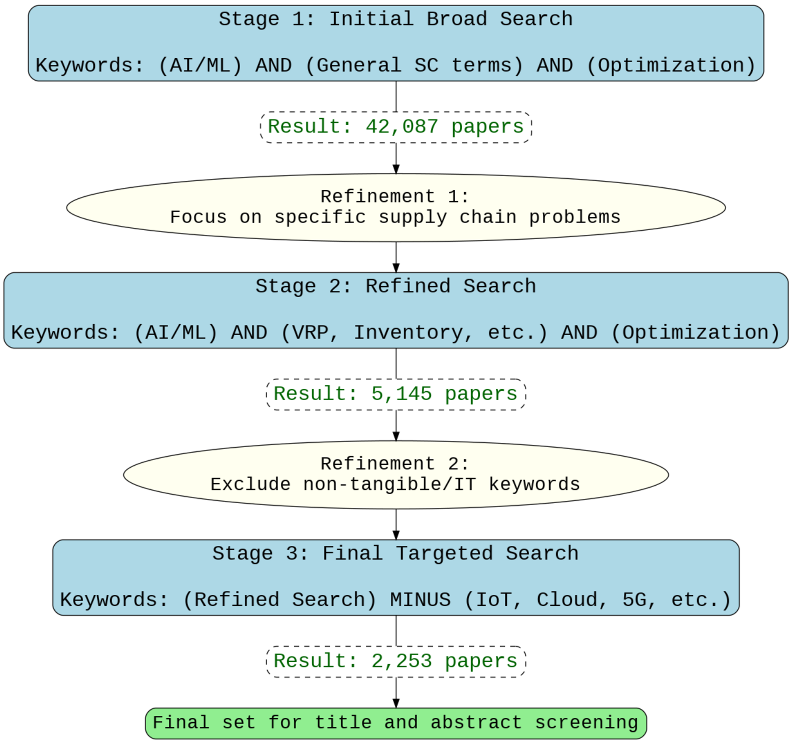


**Fig. 3** Flowchart depicting the query refinement process

on specific logistics hubs such as ports [81, 85, 124] and airports [245]. Beyond routing, papers in this category also cover inventory management [153, 181, 199], network design [59, 157], and cross-docking [116].

The Food & Water supply chain is the second largest category with 24 articles. This sector is characterized by unique challenges, including the perishability of products, the volatility of demand, and the need for sustainability. Several articles tackle these issues head-on, for example, modeling sustainable agri-food production under water scarcity [142], optimizing sustainable lime production [156], and managing perishable inventory with deep reinforcement learning [210]. Other studies focus on different stages of the food supply chain, including demand forecasting and price optimization [110], supplier selection and order allocation [77, 78], and integrated production and distribution *scheduling* [53]. The rise of on-demand services is also reflected in studies on optimizing customer and driver dispatch areas for food delivery [228] and data-driven optimization for last-mile food delivery [36]. In parallel to food, a distinct subset of papers addresses the management of water resources, a critical input for agriculture and human consumption. The research here focuses on optimizing reservoir operations [39, 180], allocating irrigation water during droughts [40, 67], and managing conjunctive surface-groundwater systems [73, 102]. Strategic studies also explore regional challenges, such as water allocation between basins [94] and managing the risks of water shortages in urban areas [34, 69].

With 16 papers, Manufacturing is another significant domain where ML and optimization are combined to improve efficiency across the production lifecycle. Research in this area addresses strategic problems such as the design of distributed manufacturing platforms [24], the optimization of sustainable automotive supply chains [43], and modular manufacturing frameworks [17]. A major focus is on scheduling, with articles on integrated supply chain *scheduling* [129], multi-factory *scheduling* with batch delivery [137], and integrated production and distribution *scheduling* [249]. Other key problems include *inventory control* [13, 72, 160], supplier selection [98], and collaborative planning in multi-tier supply chains [55].

The Healthcare domain, with 15 papers, involves high-stakes problems where optimization directly impacts public health. A significant portion of this research focuses on emergency and critical care logistics, including the distribution of COVID-19 vaccines [179, 236] and the management of blood supply chains [2, 9, 109]. Other articles address hospital operations, such as preventing drug shortages [240], optimizing inventory management [60], and routing mobile medical units [141]. Strategic healthcare problems are also explored, such as the design of facility location [193], emergency resource allocation [33, 244], and designing resilient supply chains for disaster response [243]. The management of pharmaceutical supply chains is another key topic, with studies on uncertainty-aware planning [101] and remanufacturing medical equipment [183].

Several other specialized domains and general frameworks are also represented in the literature. Broad/General Frameworks, with 10 papers, includes studies that propose methodologies applicable across various supply chain contexts. This includes reviews on the role of AI in supply chain management [117, 139, 143], frameworks for managing supply chain risk [23], models for general inventory management [204], and methods for partner selection in agile supply chains [202]. The Fuel category, with 6 papers, focuses on energy supply chains, including crude oil *scheduling* [148], biomass supply chain clustering [30], and the design of sustainable bioethanol networks [97, 133, 134]. The Port category (4 papers) addresses logistics specific to maritime hubs, such as hinterland logistics [81] and truck dispatching [85]. The Retail sector, with 3 articles, tackles challenges such as dynamic omni-channel pricing [123] and *inventory control* for neighborhood stores [65]. Finally, several smaller but important domains are covered, including Humanitarian logistics (2 papers) [66, 99], Plastic Recycling (1 paper) that focuses on closed-loop supply chains [11], and Waste management (1 paper) [155].

Table 1: SC application domains in the reviewed studies, sorted by frequency of use.

| Tangible Supply Chain Domain | Count | References |
|---|---|---|
| Transportation & Logistics | 117 | [3, 5–8, 10, 18, 21, 25, 29, 31, 32, 35, 37, 41, 42, 44, 46–51, 56–59, 61, 64, 70, 74–76, 82, 89, 100, 104, 106–108, 112, 114–116, 118–122, 128, 130–132, 135, 136, 138, 144, 146, 147, 149–154, 157, 159, 161, 163–166, 168–171, 174, 175, 181, 185–188, 190–192, 196, 199, 201, 203, 205–208, 211–214, 216, 217, 219, 220, 229, 231–235, 237–239, 241, 242, 248, 250–252] |
| Food & Water | 24 | [34, 36, 39, 40, 52–54, 62, 67–69, 73, 77, 78, 94, 95, 102, 110, 142, 156, 180, 194, 210, 228] |
| Manufacturing | 16 | [13, 17, 24, 43, 55, 71, 72, 98, 125, 129, 137, 160, 172, 218, 246, 249] |
| Healthcare | 15 | [2, 9, 33, 60, 101, 109, 141, 179, 183, 193, 197, 236, 240, 243, 244] |
| Broad/General Frameworks | 10 | [23, 96, 117, 127, 139, 143, 167, 202, 204, 247] |
| Fuel | 6 | [30, 97, 133, 134, 148, 230] |
| Port | 4 | [81, 85, 124, 245] |
| Retail | 3 | [65, 123, 145] |
| Humanitarian | 2 | [66, 99] |
| Plastic Recycling | 1 | [11] |
| Waste | 1 | [155] |

## 3.3 Mathematical Optimization in Modeling Supply Chain Operations

### 3.3.1 Types of Optimization Problem Formulations

A crucial aspect of integrating machine learning into mathematical optimization is understanding the structure of the problem being solved. The choice of formulation dictates the problem's complexity and shapes how ML can be effectively applied. Our review categorizes the articles surveyed according to their formulations of optimization problems to reveal dominant trends. The classifications show a strong prevalence of Mixed-Integer Programming (MIP) and its variants, which is characteristic of the logistics, routing, and scheduling problems that dominate this research area. For readers unfamiliar with these concepts, detailed definitions of each optimization problem type are provided in Appendix C.

The largest category by a significant margin is Mixed-Integer Programming (MIP), with 134 papers. This broad class of problems requires some or all decision variables to be discrete numbers, making it ideal for yes/no or counting decisions common in supply chains. Many of these papers are also described as Combinatorial Optimization problems, reflecting a focus on discrete structures such as graphs and sequences. The prevalence of this category highlights that the core challenge in the field is the handling of discrete, often large-scale decision problems. MIP and its sub-types are applied extensively across all domains, especially for routing problems such as the *VRP* and its many variants [6, 18, 41, 108]. It is also the go-to formulation for strategic decisions such as network design [11, 25, 43] and *Facility Location* [24], as well as for operational problems such as production *scheduling* [129] and inventory management [147, 153, 181]. One of the most complex sub-types, Mixed-Integer Non-Linear Programming (MINLP), appears in works tackling intricate problems with both discrete variables and non-linear relationships, such as in vaccine supply chains [179] and humanitarian logistics [234].

The next most frequent category is Linear Programming (LP), with 56 papers. These models are used when all relationships between variables are linear, and decisions can be fractional. LP serves as the foundation for many more complex models and is often a component of MILP. It is widely used for resource allocation problems, such as in crude oil *scheduling* [148] and water distribution [34, 40], as well as for network design [59] and *inventory control* [145]. Problems formulated purely as Integer Linear Programming (ILP), where all variables are integers but relationships are linear, were found in 11 studies, often for classic combinatorial problems such as the capacitated *VRP* [47] and multi-depot *VRP* [7].

Non-Linear Programming (NLP) formulations, identified in 28 papers, are essential when the objective function or constraints involve non-linear relationships. This approach is frequently applied in the management of natural resources, for example, in optimizing water systems [94, 102] and planning sustainable agri-food production [142]. Other applications include large-scale inventory optimization, where costs are non-linear [207] and modeling bi-level optimization problems [57].

To handle real-world uncertainty, 22 papers utilize Stochastic Programming. These models aim to find solutions that are optimal on average over a set of possible future scenarios. This includes Two-Stage Stochastic Programming, as seen in models for multi-compartment *VRP*s [32] and healthcare inventory management [60]. Other applications include data-driven *News-vendor* models [96] and frameworks for agricultural systems [52]. A related approach, Robust Optimization, was found in 6 papers. It seeks solutions that are immune to worst-case uncertainty, making it suitable to build resilience against disruptions in *VRP* [75, 251] and manage supply chain risk [23, 151].

Beyond articles presenting novel methodologies, our survey identified 10 Review Papers. These articles provide structured overviews of the existing literature to synthesize the state-of-the-art, identify the prevailing trends, and highlight critical research gaps. Their scope varies significantly, from broad ML surveys in logistics and supply chain management [3, 139] to focused analyses on specific sub-fields such as last-mile logistics [58], *inventory control* [61], and freight transportation [196]. Others center on particular ML paradigms, such as the use of unsupervised learning [167] or emerging technologies, such as ChatGPT [51], providing a valuable meta-perspective on the field's trajectory.

Finally, several other specific formulations were identified. Multi-objective Optimization, with 15 papers, is used when there are multiple conflicting goals to optimize simultaneously, such as minimizing cost while maximizing service level or sustainability [73, 155, 248]. Simulation-Optimization, found in 10 papers, is a powerful approach for complex, dynamic systems where a simulation model mimics the environment and an ML algorithm learns the best operational policies [72, 81, 207]. Dynamic Programming (DP) (5 papers) is used for sequential decision-making, such as in dynamic pricing [123] and dynamic routing [161]. The remaining papers use highly specialized formulations such as Prediction-Optimization [33] and Goal Programming [230].

Table 2: Types of optimization problems in reviewed SC studies, sorted by frequency of use.

| Problem Formulation Type | Count | References |
|---|---|---|
| MIP | 134 | [6, 8, 10, 11, 17, 18, 21, 24, 25, 29–33, 35–37, 41–43, 48–50, 53, 55, 56, 59, 62, 64, 71, 74–78, 81, 82, 89, 95, 97–101, 106–110, 112, 114–116, 118, 119, 121, 122, 124, 125, 127–138, 141, 144, 147, 149, 150, 152–154, 156, 157, 159, 165, 166, 168–172, 174, 175, 181, 183, 185–188, 191–193, 199, 204–206, 210–212, 214, 216–220, 228–231, 233–239, 241, 242, 244–246, 248–252] |
| LP | 56 | [2, 5, 6, 9, 21, 25, 34, 40, 43, 49, 53–55, 59, 64, 68–70, 77, 78, 95, 99, 100, 107–109, 112, 125, 129, 133–135, 137, 141, 145–148, 150, 153, 156, 159, 160, 163, 165, 172, 181, 185, 206, 211, 220, 231, 240, 241, 244, 245] |
| NLP | 28 | [8, 13, 17, 30, 44, 57, 67, 85, 89, 94, 101, 102, 116, 130, 142, 149, 151, 164, 179, 180, 188, 194, 204, 207, 228, 232, 234, 243] |
| Stochastic Programming | 22 | [2, 5, 23, 32, 39, 52, 54, 60, 66–69, 72, 96, 97, 151, 160, 161, 163, 179, 190, 208] |
| Multi-objective | 15 | [10, 11, 33, 34, 62, 73, 94, 142, 155, 192, 229, 243, 248, 249] |
| ILP | 11 | [7, 46, 47, 65, 104, 120, 197, 201–203, 213] |
| Review Paper | 10 | [3, 51, 58, 61, 117, 139, 143, 167, 196, 247] |

Continued on next page

Table 2: Types of optimization problems in reviewed SC studies, sorted by frequency of use. (Continued)

| | | |
|---|---|---|
| Simulation-Optimization | 10 | [5, 32, 72, 73, 81, 95, 102, 194, 207, 228] |
| Robust Optimization | 6 | [23, 56, 75, 242, 250, 251] |
| Dynamic Programming | 5 | [39, 71, 123, 161, 180] |
| Goal Programming | 1 | [230] |
| Prediction-Optimization | 1 | [33] |

### 3.3.2 Problem Categories in Reviewed Studies

Now, we move from the type of optimization problem formulations to the specific category of problems to which each study belongs. These categories are determined based on the structure of optimization problems in terms of variable type, constraint groups, and objective functions. A comprehensive definition of the categories used is provided in Appendix D.

Table 3 presents the frequency of these problem categories. As shown, the Vehicle Routing Problem (VRP) is the most dominant category, followed by the classic supply chain network problems and inventory control. For a detailed discussion on the specific trends within each problem category (such as VRP, TSP, and Lot Sizing) and how ML is applied to them, please refer to Appendix E.

Table 3: Category of optimization models in the reviewed studies, sorted by frequency of use.

| Category of Optimization Problems | Count | References |
|---|---|---|
| *VRP* | 95 | [6, 7, 9, 18, 24, 29–32, 35–37, 40–42, 46–51, 53, 64, 74, 75, 81, 82, 85, 89, 95, 100, 104, 106–108, 114, 115, 118, 120–122, 124, 127, 128, 131, 132, 135, 136, 138, 141, 144, 150, 152, 154, 159, 161, 165, 166, 169, 170, 174, 185–187, 191, 192, 196, 201, 203, 205, 206, 211–214, 216, 217, 219, 220, 228, 229, 231, 233, 235–239, 241, 242, 245, 246, 248, 251, 252] |
| *Classical Supply Chain* | 39 | [2, 10, 11, 25, 33, 34, 39, 43, 44, 52, 54, 56, 62, 67–69, 71, 73, 94, 97–99, 102, 109, 133, 134, 139, 142, 155–157, 179, 202, 204, 208, 218, 230, 234, 247, 250] |

Continued on next page

Table 3: Category of optimization models in the reviewed studies, sorted by frequency of use. (Continued)

| | | |
|---|---|---|
| *Inventory Control* | 27 | [2, 13, 60, 61, 65, 72, 99, 101, 109, 116, 117, 123, 145, 147, 153, 160, 181, 188, 190, 194, 197, 204, 207, 208, 210, 232, 240] |
| *TSP* | 16 | [21, 108, 119, 120, 122, 127, 130, 138, 141, 206, 211, 212, 216, 217, 238, 248] |
| *Scheduling* | 9 | [53, 76, 129, 137, 148, 172, 246, 247, 249] |
| *Lot-sizing* | 7 | [149, 168, 175, 183, 188, 199, 210] |
| *Supplier Selection* | 7 | [5, 66, 77, 78, 99, 202, 218] |
| *Collaborative Supply Chain* | 6 | [55, 59, 125, 151, 180, 244] |
| *Facility Location* | 6 | [17, 24, 112, 193, 243] |
| *Assignment* | 4 | [8, 77, 116, 171] |
| *Bi-level* | 4 | [57, 70, 71, 146] |
| *Bin Packing* | 4 | [118, 127, 128, 246] |
| *News-vendor* | 4 | [96, 163, 164, 181] |
| *Price Optimization* | 2 | [110, 123] |
| *Knapsack* | 1 | [240] |

### 3.4 Machine Learning Algorithms Utilized

Machine learning algorithms are traditionally classified into supervised, unsupervised, semi-supervised, and reinforcement learning approaches [176]. Within each of these classes, numerous algorithms have been developed, some of which include classification, clustering, deep learning, and neural networks. In this section, we extract the ML algorithms used in each of the papers. A review of the definitions of ML models is presented in Appendix F. Machine learning can be broadly categorized into four main paradigms: Supervised learning, Reinforcement learning (RL), Deep learning, Unsupervised learning, and Generative AI. **Supervised learning** refers to training a model using labeled data, where each input is paired with a known output. The objective is to learn a mapping from inputs to outputs that generalizes well to unseen data [22]. **Reinforcement learning** involves an agent that interacts with an environment by taking actions and receiving feedback in the form of rewards or penalties. The goal is to learn a policy that maximizes long-term cumulative reward [177]. RL operates to maximize the value in an MDP. According to the type of optimization and the structure of the environment, agents, and reward function, different types of RL have been developed, including but not limited to: Q-Learning (QL), Deep Q-Network (DQN), Proximal Policy Optimization (PPO), Deep Reinforcement Learning (DRL), Actor-Critic (AC), Multi-agent Reinforcement Learning (MARL), Self-Supervised Reinforcement Learning (SSRL), Hierarchical Reinforcement Learning (HRL), Neuroevolutionary Methods (NERL), and Multi-Armed Bandit.

**Deep learning** is a subfield of machine learning that automatically learns hierarchical representations of data. Instead of relying on hand-crafted features, deep learning uncovers increasingly abstract patterns through stacked nonlinear transformations. Its scalability, representational power, and generalizability have made deep learning the foundation of modern artificial intelligence [111]. Finally, **unsupervised learning** deals with unlabeled data, with the objective of discovering hidden structures such as clusters or lower-dimensional representations [88]. **Generative AI** focuses on AI to create new, original content, such as text and images, by learning patterns from large existing datasets.

With the definitions of the algorithms established, we next turn to their applications in the SC optimization literature. Table 4 presents the machine learning techniques identified, the number of articles employing each, and the corresponding citations. The spectrum of ML algorithms applied is broad; however, certain techniques appear more frequently and indicate clear trends.

Supervised learning methods, especially those under the umbrella of deep learning, are the most prevalent. In particular, various forms of neural networks appear in a large portion of the papers. When consolidating synonyms (ANN, DNN, CNN, etc.), we find that roughly a third to one-half of the studies employed neural network models. These range from basic feed-forward networks to specialized architectures (e.g., convolutional neural networks for spatial data, recurrent neural networks/LSTMs for time-series demand, and graph neural networks for network-structured problems) [2, 32, 136]. Neural networks are popular due to their ability to approximate complex functions; for instance, some researchers train neural nets to approximate the objective function or constraints of an optimization model, or to directly predict decisions (as in learning to construct routes or schedules). Despite their complexity in training, neural networks have shown promise, especially as larger datasets become available in supply chains [32, 40, 56].

In addition to neural networks, there is considerable use of tree-based and kernel-based ML methods. Support Vector Machines (SVM) are used, often for predictive tasks such as estimating demand or risk, or for classification problems within supply chain decisions. Decision tree ensembles such as random forests and boost algorithms such as XGBoost [109, 133] are also common. These models are valued for their interpretability and strong performance on tabular data. In some cases, gradient-boosting models are used to capture nonlinear relationships in supply chain data, e.g. to predict lead times or disruption risks, which then feed into optimization decisions.

Another highly popular approach is reinforcement learning (RL). About 92 papers applied reinforcement learning techniques (often deep reinforcement learning) to dynamic or sequential decision problems. RL is well-suited for problems such as *inventory control*, fleet management, or adaptive routing, where decisions at one stage affect future states [39, 148, 205]. Many studies specifically leverage deep RL, combining neural networks with RL algorithms (deep Q-Networks, policy gradient methods, etc.) to handle high-dimensional state spaces [76, 161, 240]. In fact, the use of deep reinforcement learning has grown so much that it almost rivals classical RL in frequency in our dataset. For example, some previous work uses RL to train agents that make inventory replenishment decisions, achieving performance comparable to or better than traditional OR policies under uncertainty [61, 72, 160]. Q-learning is another popular category in RL, providing tractable solutions using a model-free approach [99, 116, 137, 183, 192, 208].

Deep learning methods have been used frequently with RL [6, 76, 121, 219, 229, 234]. Deep learning algorithms are very fast in learning complex functional representations between input and output, making them a suitable tool for learning the value or policy functions. In addition, they have been used to develop efficient heuristics to solve complex, large-scale problems [145, 152, 185, 242]. LSTM algorithms, suitable for learning sequential relationships, have been used to solve dynamic problems such as *inventory control* and *VRP* [61, 114, 123].

Unsupervised learning methods are also reported in the literature, with approximately 19 studies employing clustering techniques. Clustering is typically used to segment customers, products, or scenarios (for example, clustering demand patterns to simplify stochastic optimization models [101]or grouping facilities by similar profiles [230]). Some works also utilize principal component analysis (PCA) or dimensionality reduction to preprocess data [78].

In addition to the articles mentioned above, [3, 58, 117, 167, 186, 196, 246] provides a review of machine learning in logistics and supply chain management, optimization, and machine learning in last-mile logistics, impacts of AI on supply chain management, unsupervised learning algorithms in supply chain management, machine learning in freight transportation, and solving *VRP* using ML. Our paper advances the literature on ML in supply chains by investigating how ML is used to solve optimization problems in tangible supply chains.

In particular, a handful of cutting-edge or specialized ML approaches are present in the literature: two studies explore graph neural networks (GNN) for routing problems (since *VRP* can be naturally represented on graphs) [37], some use transformers [47, 48] or advanced sequence models for time-series forecasting in supply chains [110, 150], and one study reports a quantum SVM for faster computation in a research context [131]. However, these areas constitute a growing body of literature with ample potential for investigation. In general, the most frequently used ML algorithms are neural-network-based (including deep learning), followed by reinforcement learning methods, and then classical ML models such as SVMs, decision trees, and clustering. This indicates a strong interest in leveraging the power of deep learning and RL in operations research problems, while also employing conventional ML for structured prediction tasks in supply chain contexts.

A particularly noteworthy and recent trend is the emergence of Generative AI, which focuses on creating new original content, such as text and images, by learning patterns from large existing datasets. Although our systematic review captured limited instances of its use, reflecting its novelty, Generative AI is poised to become a transformative tool in supply chain optimization. As highlighted by recent work, such as [51], initial evidence for the application of models, such as ChatGPT, in supply chains is already surfacing. The potential applications in optimization are vast, including the generation of realistic synthetic data to train other ML models, the assistance of human planners in formulating complex optimization problems through natural language, or the creation of dynamic scenarios for robust testing of supply chain strategies. Its ability to bridge the gap between human language and mathematical models presents a significant avenue for future research.

Table 4: Machine learning models and techniques used in the reviewed studies, sorted by frequency of use.

| AI/ML Category | ML Model | Count | References |
|---|---|---|---|
| Supervised Learning | Neural Networks | 75 | [2, 8–10, 18, 21, 23, 24, 32, 34, 37, 39–41, 41–44, 53, 54, 56, 57, 59, 61, 67–70, 73, 74, 82, 89, 94, 98, 100, 104, 112, 118–120, 135, 136, 142, 143, 146, 147, 149, 151, 152, 154, 157, 164, 166, 175, 179, 180, 183, 190, 193, 194, 201, 202, 205–207, 211, 214, 229, 231, 233, 235, 240, 243, 245, 250] |
| | Regression | 14 | [36, 50, 65, 77, 95, 97, 110, 128, 156, 159, 191, 194, 228, 232] |
| | Random Forest | 13 | [2, 21, 60, 61, 97, 110, 128, 133, 134, 155, 156, 172, 194] |
| | SVM | 12 | [17, 21, 23, 52, 61, 65, 97, 102, 128, 132, 156, 172] |
| | Classification | 11 | [2, 21, 60, 97, 110, 124, 150, 172, 228, 229, 236] |
| | Boosting | 6 | [60, 61, 109, 110, 133, 134] |
| | Decision Tree | 4 | [25, 144, 186, 228] |
| | Ensemble Learning | 2 | [133, 134] |
| | Linear Machine Learning | 1 | [96] |
| | Logistic Regression | 1 | [11] |
| | Quantum Supervised ML | 1 | [131] |

Continued on next page

Table 4: Machine learning models and techniques used in the reviewed studies, sorted by frequency of use. (Continued)

| | | | |
|---|---|---|---|
| Reinforcement Learning | DRL | 45 | [6, 9, 29, 35, 37, 39, 42, 47, 48, 61, 72, 75, 85, 100, 107, 108, 119–121, 123, 125, 138, 161, 171, 185, 186, 188, 197, 201, 204, 210–214, 219, 231, 233, 239, 240, 244, 247, 248, 251, 252] |
| | QL | 15 | [5, 39, 76, 99, 116, 137, 183, 192, 208, 234, 237, 242, 246, 249, 251] |
| | PPO | 13 | [7, 31, 82, 127, 139, 141, 153, 159, 160, 165, 181, 199, 238] |
| | MARL | 10 | [6, 55, 81, 122, 145, 147, 170, 187, 205, 241] |
| | AC | 6 | [71, 119, 129, 154, 203, 206] |
| | DQN | 4 | [139, 152, 183, 229] |
| | Multi-armed bandit | 2 | [106, 169] |
| | HRL | 1 | [148] |
| | NERL | 1 | [168] |
| | SSRL | 1 | [217] |
| Deep Learning | Deep Learning | 61 | [6, 29, 33, 35, 37, 41–43, 48, 71, 72, 75, 76, 78, 85, 100, 107, 108, 115, 116, 119, 121–123, 127, 138, 145, 152, 161, 163, 165, 183, 185, 186, 188, 193, 197, 203, 204, 206, 208, 210, 211, 214, 216, 219, 229, 232–234, 239–242, 244, 246–248, 250–252] |
| | LSTM | 6 | [61, 114, 123, 157, 190, 218] |
| Unsupervised Learning | Clustering | 20 | [13, 30, 46, 49, 56, 62, 64, 66, 101, 118, 124, 130, 174, 186, 191, 192, 197, 220, 230, 249] |
| | PCA | 1 | [78] |
| Generative AI | Generative AI | 1 | [51] |

## 3.5 Purpose of Machine Learning in Optimization Models

Table 5 presents a classification of the reviewed studies according to the specific role or purpose of machine learning (ML) techniques within the optimization framework. This categorization enables a structured understanding of how ML is used in decision support systems and optimization workflows, providing both a snapshot of current research practices and directions for future research.

The classification schema includes five main categories: Solution, Formulation, Parameter Estimation, Reformulation, and Review Paper.

The first category is the **Solution** category; this is the most populated category in the reviewed literature. Here, ML is used to enhance or replace traditional solution methods, such as exact solvers, heuristics, or metaheuristics. Reinforcement learning, neural networks, and hybrid models are frequently applied to learn optimal or near-optimal policies, especially in complex, dynamic, or large-scale environments. These approaches often enable online decision-making and adaptive optimization in real time.

The second category is **Formulation**; in this group, ML techniques are employed to influence the mathematical structure or logic of the optimization model itself. This includes learning constraints, identifying hidden relationships among decision variables, or discovering surrogate functions that approximate complex system behaviors. In particular, deep learning models have been used to replace black-box physical models, thereby enabling tractable optimization formulations.

The third category is **Parameter Estimation**; many optimization models rely on parameters such as demand forecasts, production costs, travel times, or probability distributions. In this category, ML is used to estimate such parameters from historical or real-time data. Techniques such as supervised learning, clustering, and probabilistic modeling are commonly applied. This category reflects a data-driven approach to optimization where model precision depends on the quality of parameter inference.

The next category is **Reformulation**; this smaller but growing category includes studies in which ML is used to modify or transform an existing optimization problem. For example, ML can help detect redundant constraints, approximate infeasible regions, or suggest alternative variable encoding. These reformulations can lead to significant improvements in the computational efficiency or robustness of the optimization process.

Finally, we have the **Review Paper** category; this category encompasses review and survey papers that provide a high-level overview of ML applications in optimization without necessarily implementing the ML methods themselves. These articles contribute to knowledge consolidation, taxonomy development, and gap identification in the literature. Examples include comprehensive reviews of ML-augmented supply chain optimization or reinforcement learning in resource allocation.

The analysis reveals that the majority of studies apply ML in the solution stage, indicating a clear trend toward integrating learning-based solvers into optimization pipelines. This is consistent with the larger shift from static optimization to intelligent and adaptive decision systems. The significant presence of studies in the formulation and parameter estimation categories further highlights the central role of data in shaping modern optimization models. However, the relatively small number of studies in the reformulation category suggests that this remains an underexplored area with considerable potential for future contributions, especially in improving scalability and solver efficiency.

In general, this categorization not only clarifies the functional roles of ML in optimization models but also reflects the evolving research landscape. It underscores the multifaceted integration of ML into both the design and execution of optimization, bridging data science with operations research and systems engineering.

Table 5: Purpose of using machine learning in the optimization model of the paper reviewed, sorted by frequency of use.

| Purpose of Using ML | Count | References |
|---|---|---|
| Solution | 97 | [2, 5–8, 10, 13, 17, 18, 24, 25, 30, 31, 33–35, 37, 39, 41, 44, 50, 54, 55, 57, 59, 64, 67–70, 73, 74, 82, 89, 94, 97, 99, 100, 102, 104, 106, 116, 118–120, 122, 124, 128, 130, 131, 135–138, 144, 146, 149, 151, 153, 154, 160, 161, 163, 164, 168, 169, 171, 175, 180, 185, 188, 191–193, 197, 199, 201–203, 205, 207, 212, 213, 219, 228, 229, 233–236, 238, 239, 241, 243, 245, 250, 251] |
| Formulation | 76 | [6, 9, 17, 23, 29, 33, 39, 42, 47, 48, 54, 67, 68, 71, 72, 75, 76, 81, 85, 95, 96, 107, 108, 114–116, 119, 121–123, 125, 129, 132, 141, 145, 147, 148, 152, 154, 157, 159–161, 165, 166, 168, 170, 180, 181, 183, 185, 202, 204, 206–208, 210, 211, 214, 216–219, 231, 233, 237, 239, 240, 242, 245–250, 252] |
| Parameter Estimation | 58 | [5, 11, 21, 23, 32, 36, 40, 43, 46, 49, 52, 53, 56, 60, 62, 64–67, 77, 78, 89, 94, 95, 98, 101, 109, 110, 112, 117, 127, 130, 133, 134, 139, 142, 146, 150, 155–157, 164, 165, 172, 174, 179, 187, 190–192, 194, 196, 210, 220, 228–230, 232] |
| Review Paper | 7 | [3, 51, 58, 61, 143, 167, 186] |
| Reformulation | 2 | [53, 244] |

### 3.6 Geographical Scale of Tangible Supply Chain Problems

The operational scale of a supply chain (Local, National, or Scale-Agnostic) significantly influences optimization challenges. For a comprehensive classification of papers by geographical scale and a discussion of the findings, please refer to Appendix G.

## 4 Discussion and Synthesis of Findings

This section discusses our main findings on how ML and mathematical optimization are integrated to address SC problems. Figure 4 presents a three-dimensional data cube that captures the co-occurrence of ML models, optimization problem categories, and purposes of ML usage. The landscape it reveals shows that specific ML methods are deliberately chosen for particular optimization tasks, rather than being applied randomly. To enhance readability, we also include a Sankey diagram (Fig. 5), which illustrates the flow of connections between problem categories, ML models, and purposes, as well as complementary two-dimensional heatmaps (Figs. 6 - 12, presented in Sections 4.1 - 4.7) that decompose the cube into pairwise relationships and highlight dominant patterns.

We found that the choice of an ML model depends on a few key factors: the category of the optimization problem (such as *VRP* or *inventory control*), the purpose of using ML (e.g., to help set up the problem or to find the final solution) and the operational scale of the problem.

For example, we consistently saw that logistical problems such as the *VRP* are often tackled with Reinforcement Learning and Neural Networks. This is a logical pairing because these problems involve making a series of decisions, which is exactly what these ML models are good at. We also found that what the ML model is used for is very specific. Flexible models, such as Neural Networks, are used for almost everything, while Reinforcement Learning is mostly used to directly find a solution to an optimization problem. The following subsections will explore these connections in more detail, providing a closer look at the trends shaping this field.

## 4.1 ML Model vs Type of Optimization Problems

This subsection explores the relationship between the type of ML model and the mathematical formulation of the optimization problem it is paired with. The analysis observed in Figure 6 reveals that certain ML models are strongly preferred for specific types of formulations, reflecting a natural synergy between their capabilities.

A dominant trend is the deep integration of ML with MIP. This formulation is heavily paired with Reinforcement Learning (42 instances) and Neural Networks (39 instances), as well as Deep Reinforcement Learning (18 instances). This strong connection is logical, as MIP is the standard for modeling problems involving discrete choices, such as *facility location* or *scheduling*. Reinforcement Learning excels at learning decision policies for such environments, while Neural Networks can approximate complex functions or predict parameters within the MIP framework.

Beyond MIP, other formulations also show distinct patterns. LP is frequently combined with Neural Networks (21 instances) and Reinforcement Learning (19 instances). Similarly, NLP shows a strong link with Neural Networks (15 instances), likely leveraging NNs to handle or approximate complex non-linear relationships. Furthermore, for problems involving uncertainty, Stochastic Programming is most often paired with Neural Networks (12 instances), which are well-suited for forecasting the random variables central to these models.

The key insight is that the research community has established a powerful and practical partnership between learning-based models and discrete optimization, particularly MIP. The benefit of this focus is the accelerated progress in solving widespread and computationally demanding industrial problems. However, this concentration also reveals a potential limitation. The sparse application of ML to other valuable formulations, such as Dynamic or Robust Programming, suggests that the field may be under-exploring novel integrations that could unlock new solutions.

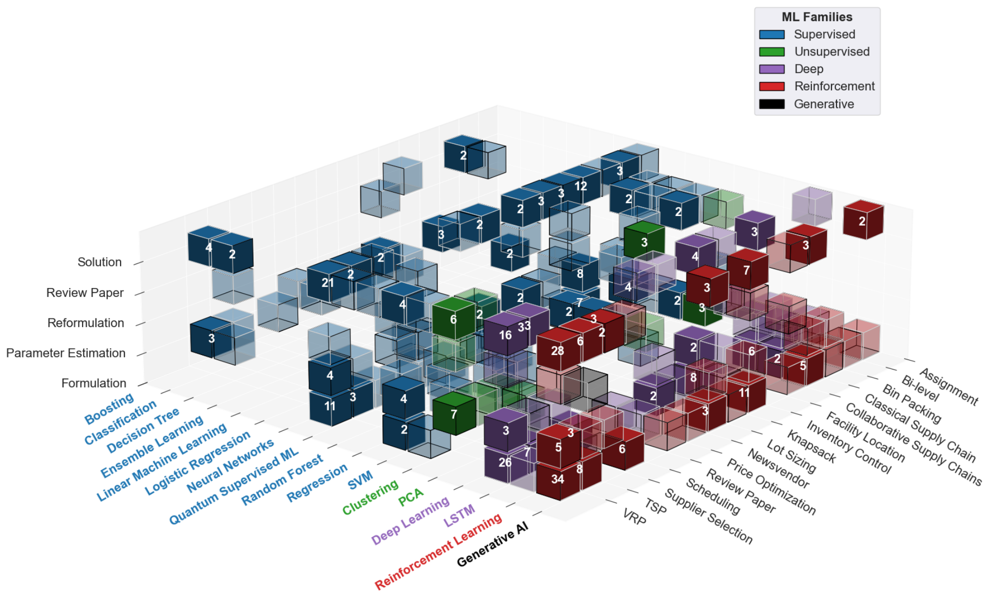


**Fig. 4** Three-dimensional data cube illustrating how machine learning methods, optimization problem categories, and purposes of ML intersect. Cubes without numerals indicate a single input. High-frequency clusters (e.g., reinforcement learning for solving vehicle routing problems) emerge, highlighting research hotspots, while sparse regions reveal an underexplored area.

Figure 5 presents a Sankey diagram illustrating the relationships among the three main dimensions identified in the review of the literature. (i) **Types of optimization problems** (left column), (ii) **Machine learning techniques** (center column), (iii) **Application purposes of ML in optimization** (right column). Sankey diagrams are a type of flow diagram in which the width of each connection is proportional to the quantity of flow it represents. They are widely used to visualize resource, energy, and information flows [178]. Each flow in the diagram represents a co-occurrence in the reviewed studies, where a given machine learning technique was applied to a specific optimization problem for a particular purpose. The width of each link is proportional to the number of papers that exhibit that combination.

The left column demonstrates the diversity of optimization problems addressed in the literature, with particular concentration in areas such as the *VRP* and *lot-sizing*. The central column shows a rich variety of machine learning techniques employed, such as Deep Neural Networks, Reinforcement Learning, and Support Vector Machines

(SVM). The right column captures the purpose of ML usage, which is predominantly for *Solution generation*, followed by *Parameter Estimation*, and in some cases for *Formulation* and *Reformulation*. This Sankey diagram visually summarizes the methodological landscape and identifies which ML techniques are the most prevalent for specific optimization challenges and tasks.

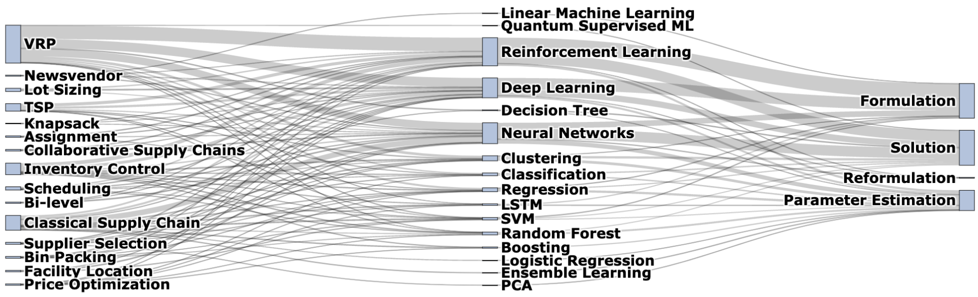


**Fig. 5** Sankey diagram showing connections between types of optimization problems (left), machine learning techniques (middle), and their purposes in the optimization workflow (right). Flow width represents the frequency of co-occurrence across the reviewed literature.

## 4.2 ML Model vs Category of Problems

This analysis examines how specific categories of optimization problems are paired with different ML models, revealing that the choice of model is highly dependent on the inherent structure of the problem. A clear and dominant pattern emerges in logistical and network problems, as seen in Figure 7. The *VRP* shows an overwhelming preference for Deep Reinforcement Learning (39 instances), Reinforcement Learning (36 instances), and Neural Networks (35 instances). A similar, though less frequent, trend is visible in the *TSP*, which also heavily utilizes these three models. This strong correlation is driven by the nature of these problems, which involve sequential decision-making in complex environments where RL is ideal for learning optimal policies and NNs are effective at approximating the required value or policy functions.

Other significant patterns are also evident. The broad category of *Classical Supply Chain* problems shows a strong reliance on Neural Networks (24 instances), highlighting their versatility in modeling general supply chain dynamics. In *Inventory Control*, Deep Reinforcement Learning (13 instances) is the most prominent model. This is a logical fit, as managing inventory is a classic sequential decision problem requiring the balancing of stock levels over time under uncertain demand, a task well-suited for DRL. Additionally, Clustering is used significantly in *VRP* (16 instances), typically as a preprocessing step to partition customers into manageable groups before a routing algorithm is applied, demonstrating a different but crucial role for ML in simplifying complex problems.

The key insight here is the natural fit between learning-based models, such as RL and NNs, and dynamic problems, such as *VRP*. The strength of this approach is the

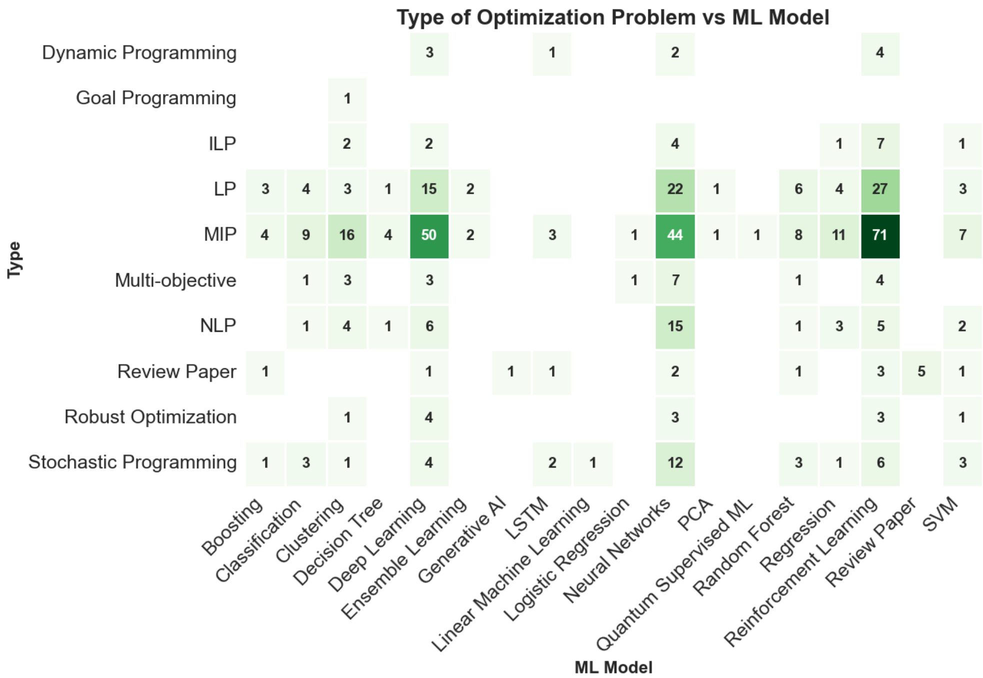


**Fig. 6** Heatmap of Co-occurrence of Machine Learning Models and Type of Optimization Problems

ability to find effective solutions in complex real-world scenarios where traditional methods might struggle with scale or uncertainty. However, this often involves a trade-off: these ML methods provide fast, high-quality answers but may not guarantee the single best, mathematically optimal solution. The use of Clustering as a first step highlights a practical benefit, making huge problems more manageable, though the final outcome depends heavily on the quality of this initial simplification.

## 4.3 ML Model vs Purpose of Use

This analysis examines the functional role that different ML models play within optimization frameworks (Figure 8), revealing that models are chosen for specific purposes that align with their core strengths.

The data show that Neural Networks are exceptionally versatile, with high usage across all primary stages: finding a Solution (50 instances), Formulation (26 instances) and Parameter Estimation (21 instances). This highlights their flexibility in handling a wide range of tasks from prediction to direct control.

In contrast, Reinforcement Learning models show more specialized, yet powerful, applications. Deep Reinforcement Learning is overwhelmingly the dominant model for problem Formulation (47 instances). Similarly, traditional Reinforcement Learning is also heavily applied to both Formulation (35 instances) and finding the final Solution

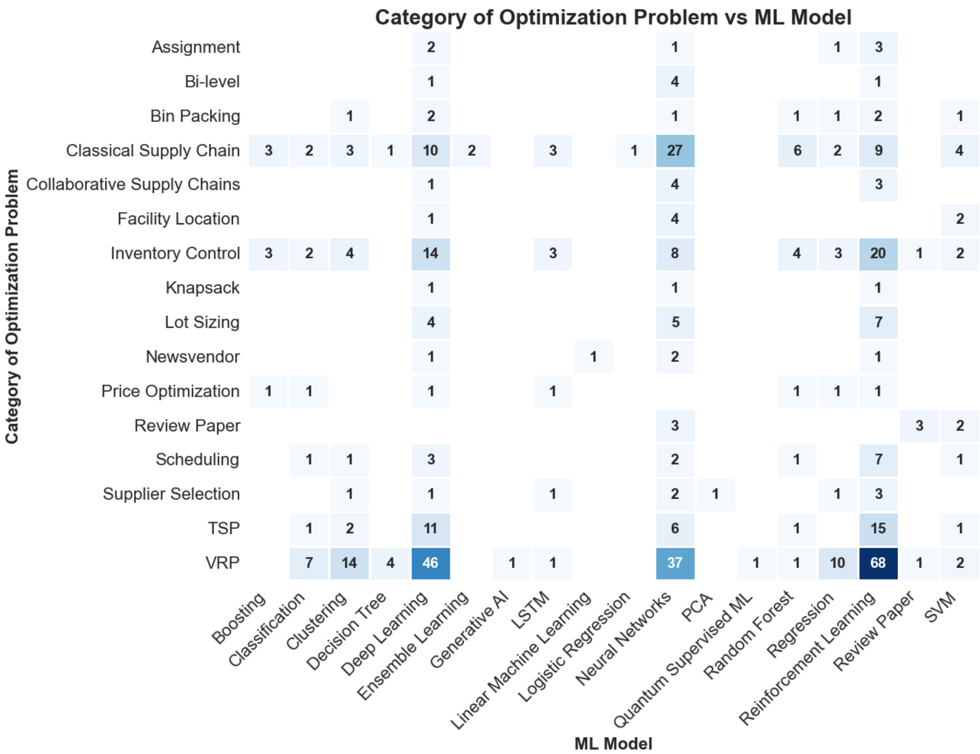


**Fig. 7** Heatmap of co-occurrence of category of optimization problem and ML model

(36 instances), underscoring its dual strength in defining a problem and then solving it.

For Parameter Estimation, Neural Networks (21 instances) are the most prominent, followed by traditional statistical methods such as Clustering (15 instances) and Regression (11 instances). This is intuitive, as NNs can capture complex relationships for prediction, while clustering segments data and regression predicts numerical inputs, such as demand or cost. Finally, the analysis reinforces a significant research gap: the purpose of Reformulation remains almost entirely unexplored, with negligible application across all ML models.

The key insight is that researchers consistently apply a ”right tool for the job” principle, selecting methods deliberately based on the task at hand. The advantage of a versatile model, such as a Neural Network, is its ability to handle many different tasks, though it may not always be the most efficient or easy-to-interpret choice. In contrast, specialized models such as Reinforcement Learning are extremely powerful for their intended purposes, defining problems and finding policies, but this strength comes with higher implementation complexity. The notable gap in Reformulation suggests a missed opportunity in the field: using ML to make problems simpler before solving them.

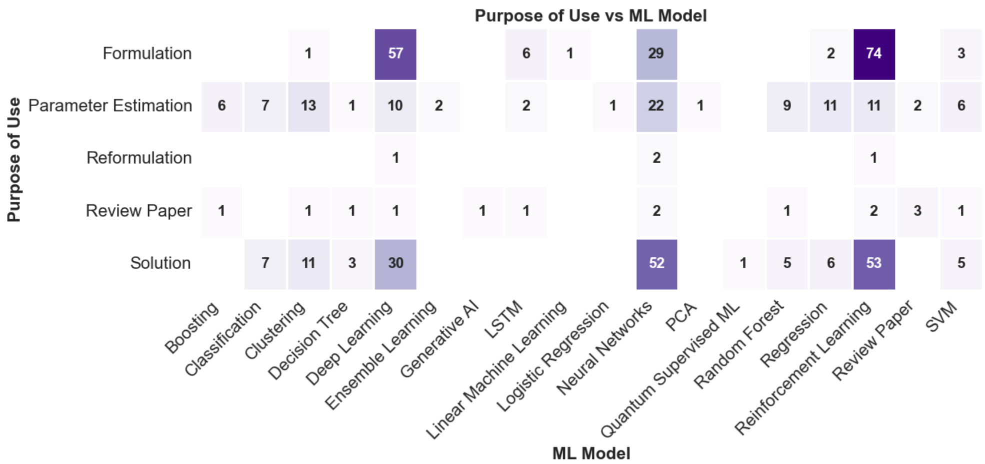


**Fig. 8** Heatmap of co-occurrence of purpose of ML and ML model

## 4.4 Tangible Supply Chain Type vs Type of Optimization Formulation

This subsection analyzes the relationship between specific SC domains and the choice of mathematical optimization formulation. The results, as seen in Figure 9, indicate that the formulation is strongly influenced by the intrinsic nature of the problems within a given domain.

The most prominent trend is observed in the Transportation & Logistics domain, which is overwhelmingly based on MIP with 91 instances. This strong preference is logical, as transportation problems are fundamentally concerned with making discrete choices, such as which routes to take or which facilities to use. The same domain also frequently employs LP (27 instances) and NLP (13 instances) for problems involving continuous variables.

The versatility of MIP is further underscored by its frequent application in other key domains, including Manufacturing (13 instances), Food & Water (10 instances), and Healthcare (9 instances). This suggests that discrete decision-making is a common challenge in areas such as production planning, distribution, and resource scheduling. In contrast, formulations designed to handle uncertainty, such as Stochastic Programming, are used more sparingly, but appear across multiple domains, signaling their targeted use for problems with unpredictable variables, such as customer demand.

In contrast to these well-established areas, emerging and specialized domains such as Humanitarian logistics, Plastic Recycling, and Waste management show significantly less research activity. Although they also utilize foundational formulations such as MIP and LP, the low frequency of publications suggests that the unique challenges within these fields are underexplored. This presents a clear opportunity for future research. The sophisticated optimization techniques successfully applied in commercial logistics could be adapted to address critical issues, such as efficient resource distribution in humanitarian crises or the design of optimal circular supply chains for recycling, enhancing sustainable supply chain management.

This highlights a core principle: the real-world problem dictates the mathematical language used to describe it. The advantage of using MIP is its ability to precisely model discrete, yes-or-no decisions common in supply chains, though this often comes with a high computational cost for large-scale problems. The concentration of research in established commercial domains demonstrates the maturity of these methods. However, the sparse application in emerging critical areas such as humanitarian and circular supply chains points to a major opportunity: adapting these proven, powerful formulations to solve new high-impact societal challenges.

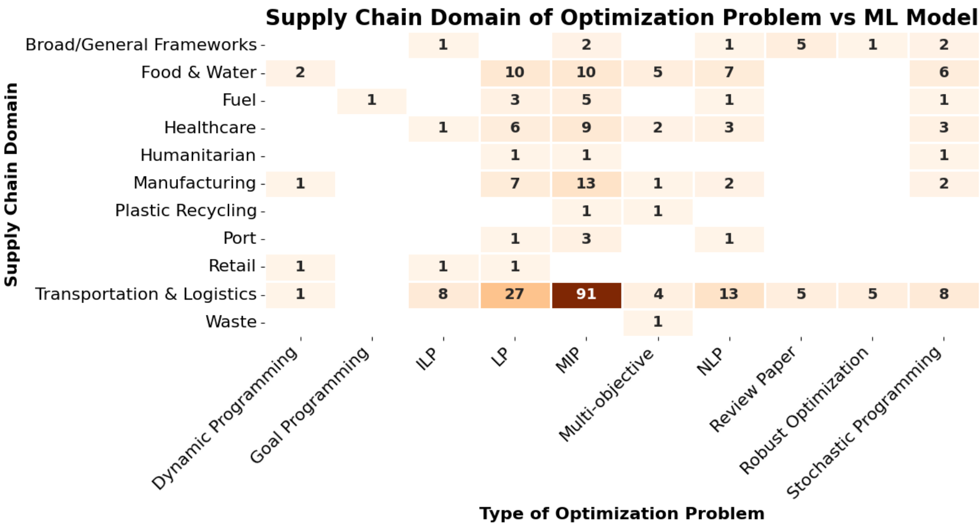


**Fig. 9** Heatmap of co-occurrence of type of SC and type of optimization problem

## 4.5 Category of Problem vs Purpose of Use of ML

This analysis reveals that the functional role of machine learning is not uniform, but is strategically adapted to the specific structural demands of different categories of optimization problems. A clear pattern (Figure 10) distinguishes between the problems in which ML is used for preparatory stages versus those in which it is used for the generation of direct solutions.

The most dominant trend is seen in complex logistical problems. For the *VRP*, ML is overwhelmingly used for direct Solution generation (72 instances) and problem Formulation (62 instances). This indicates that for routing challenges, ML is integral to both defining the problem's dynamic elements and finding the final operational solution. A similar pattern is observed in the *TSP*, where Formulation (17 instances) and Solution (10 instances) are the primary applications.

In contrast, for the broad category of *Classical Supply Chain* problems, the most significant application of ML is for Parameter Estimation (26 instances). This is a critical insight, suggesting that for general supply chain models, the greatest value of ML is seen in providing accurate data-driven input (such as demand forecasts, lead times, or cost estimates) before the optimization model is solved. This focus on parameter estimation is also prominent in problems such as the Knapsack problem (16 instances), where accurately defining the value or weight of items is essential. Parameter estimation is a classical application of ML in optimization, especially when some prior data is available. Depending on the type of problem, the ML tool used can be modified. For example, if we are basically predicting future demand based on available data, regression-based methods can be used. If the data are not abundant, semi-supervised algorithms are the most suitable. If we have numerous data points and we want to reduce them in order to make the optimization model smaller but representative of the entire population, clustering algorithms might be used.

A striking counterpoint to these well-defined roles is the near-total absence of ML for Reformulation across all problem types, highlighting a significant and consistent gap in current research practices.

The insight here is that the role of ML shifts from a preparatory tool to an end-to-end solver depending on the nature of the problem. For highly dynamic problems such as *VRP*, using ML to find the solution directly is powerful, yet these systems can be difficult to interpret, making it challenging to understand why a particular solution was chosen. For broader, more structured problems, the primary benefit of ML is to improve the model's input through accurate parameter estimation, leading to more reliable optimization outcomes. A potential limitation of this latter approach is that it separates the learning and solving steps, which may not be globally optimal compared to a more integrated method.

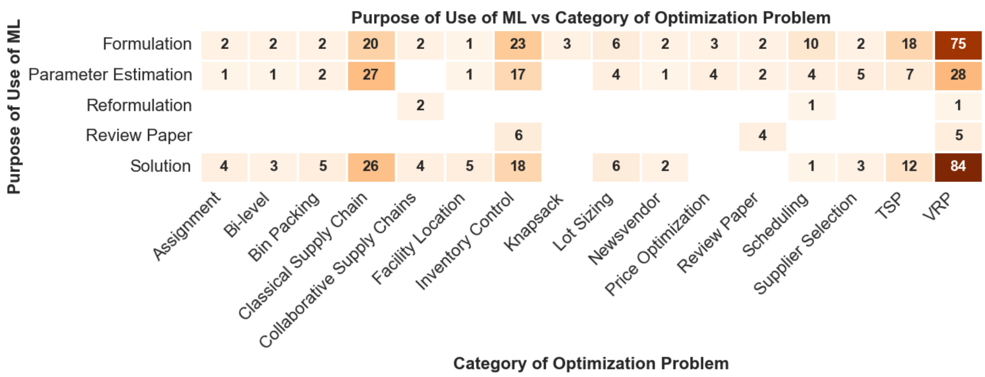


**Fig. 10** Heatmap of Co-occurrence of Purpose of Using ML and Category of Optimization Problems

## 4.6 ML Model/Method vs Scale of Problem

This subsection examines how the operational scale of an SC problem correlates with the choice of ML model (Figure 11). A key finding is that the most popular models are most frequently cited in papers that do not specify a scale (Any). This category is dominated by Neural Networks (47 instances), Reinforcement Learning (42 instances), and Deep Reinforcement Learning (38 instances). This suggests that a significant portion of the research focuses on developing general frameworks rather than tailoring models to a specific operational size.

For problems with a defined scale, Neural Networks remain a versatile choice, appearing frequently in both Large-scale (12 instances) and Small-scale (11 instances) contexts, likely due to their ability to model complex, non-linear relationships found in systems of any size. Interestingly, traditional Reinforcement Learning shows more use in small-scale problems (9 instances), while other models, such as Clustering, are more common in large-scale applications (5 instances), where they can be used to segment large datasets.

A key insight is that scale matters. Much of the current research does not specify the scale of the problem, which has significant implications. On the one hand, developing scale-agnostic models is valuable to advance theory and ensure generality. However, models designed without a specific scale in mind may not perform well in practice, whether applied to a small factory or a city-wide network. This highlights a gap between theoretical research and practical implementation, where scale is often a critical determinant of performance.

*Scale is a design variable.* The dominance of "Any-scale" studies signals that many papers abstract away from instance size, which limits operational guidance. Where scale is specified, clearer patterns emerge. Neural networks remain the most versatile choice across sizes because they learn non-linear structure and can be right-sized (from shallow MLPs to deep architectures). Methods with high sample or compute demand—especially model-free RL and deep RL—tend to work best on *small to medium* instances where exploration and long training runs are feasible; on larger instances, they benefit from action-space restriction, imitation learning from strong heuristics, or hierarchical policies. For *large-scale* deployments, representation/segmentation steps (clustering, dimensionality reduction, graph embeddings) paired with optimization are prevalent: they compress the state/action space before learning and integrate naturally with decomposition (Benders/Lagrangian, column generation, and CTDE in MARL). In practice, match the method to scale: for small testbeds, favor model-free RL or policy-gradient/actor–critic with rich simulation; for growing instances, move to hybrid pipelines that learn forecasts, costs, or cuts and solve with decomposition or matheuristics; for network- or city-scale problems, prioritize scalable surrogates and OR-hybrids (learning-aided routing/scheduling, learned pricing/branching, or policy learning over restricted neighborhoods), and report instance sizes, wall-clock/compute budgets, and latency to make scalability transparent.

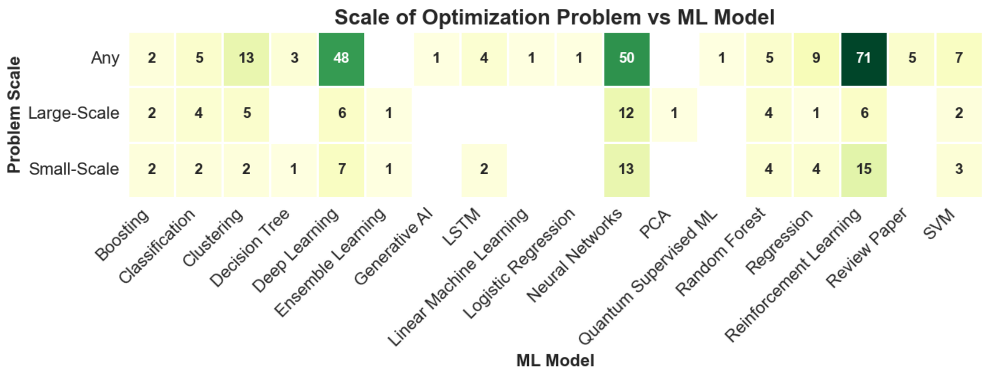


**Fig. 11** Heatmap of Co-occurrence of Machine Learning Models and Scale of Optimization Problems

## 4.7 Type of Optimization Problem vs Purpose of ML

This analysis investigates the connection between the functional purpose of an ML application and the underlying mathematical formulation with which it is paired. The results, shown in Figure 12, reveal that while some formulations are deeply integrated with ML, others represent significant and underexplored opportunities.

As expected, the most common formulations, MIP and LP, show deep integration with ML across all stages. MIP, in particular, is heavily supported by ML for finding a Solution (64 instances), defining the Formulation (56 instances), and providing Parameter Estimation (38 instances). This confirms that current research is heavily focused on using ML to enhance these foundational discrete and continuous optimization methods.

However, the heatmap reveals significant research gaps in which ML is sparsely applied. The most striking is Reformulation, with only a handful of instances, indicating that using ML to simplify or restructure optimization problems before solving them is a largely untapped strategy. Beyond this functional gap, several important optimization types are overlooked. For instance, Dynamic Programming shows minimal ML integration (e.g., 5 for Formulation), despite its sequential nature being a natural fit for techniques such as Reinforcement Learning. Similarly, Robust Optimization and Multi-objective Programming show very limited ML applications. This is a notable gap, as ML could be a powerful tool for learning the data-driven uncertainty sets required for robust models, or for approximating the complex trade-offs between conflicting goals in multi-objective problems.

In SC applications, the choice of RL method should follow the action structure and information flow rather than a rigid mapping. Value-based RL (e.g., Q-learning, DQN) works well for many *discrete-action* problems such as certain inventory policies, though very large combinatorial action spaces often require action factorization or other structure. Policy-gradient and actor–critic methods (e.g., REINFORCE, PPO, SAC, DDPG) naturally handle *continuous* controls and can also parameterize large *discrete* spaces; importantly, most routing and scheduling decisions are discrete rather than continuous. When decision-making is decentralized or collaborative, multi-agent

RL is appropriate, commonly using centralized training with decentralized execution to mitigate non-stationarity and support coordination. For long-horizon problems, model-based RL can improve foresight and sample efficiency when a credible dynamics model exists, while planning-augmented model-free approaches (e.g., lookahead rollouts, tree search, hierarchical RL) are also effective. In practice, the best method is problem-specific and often combines RL with optimization or domain structure.

The key insight is that while the partnership between ML and mainstream formulations, such as MIP, is mature and effective, this focus has created a research landscape in which other powerful techniques are neglected. The advantage of the current approach is the significant progress made on widely known, difficult problems. The missed opportunity, however, lies in not exploring how ML could innovate in areas such as robust decision-making or multi-criteria optimization, which are equally critical for modern supply chains.

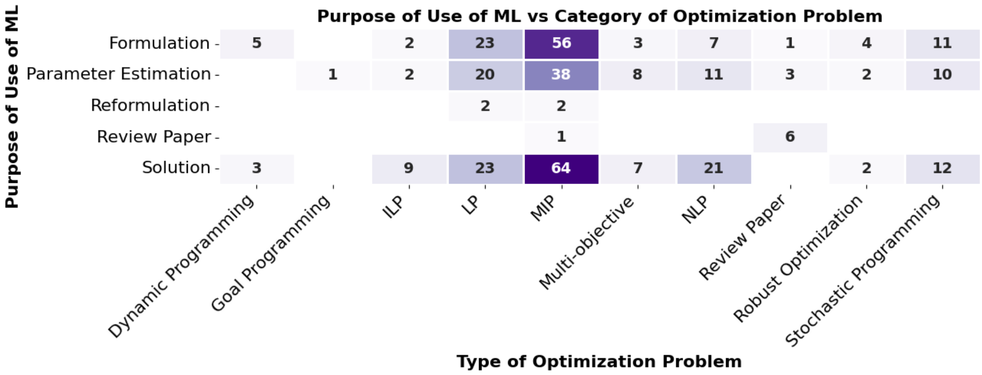


**Fig. 12** Heatmap of Co-occurrence of Type of Optimization Problems and Purpose of ML

## 4.8 Analysis of Supply Chain Parameters Predicted by Machine Learning

We observe that one of the most significant applications of machine learning in supply chain optimization models is to help estimate parameters. Whether this estimation is used in deterministic settings to estimate a value or in stochastic settings to estimate and generate samples from a probability distribution, machine learning models are primarily based on data [87]. However, more often than not, the data required to train these models is either unavailable (e.g., estimating the demand for a new product), scarce (e.g., estimating the preference of individuals on certain products), or expensive to acquire (e.g., the number of customers using a high-tech facility). In this case, insufficient data might lead to over-fitted models with unreliable estimates [38]. One of the primary methods to handle this issue is by data augmentation; inspired by Bayesian learning, data augmentation generates new data points with the same marginal distribution as current data points, given certain estimators [198]. In addition, self-supervised techniques can be used to derive labels for unlabeled data [126].

A primary role of machine learning in integrated optimization frameworks is to provide accurate, data-driven estimates for the uncertain parameters that govern supply chain models. Our review analyzed the specific parameters that researchers are predicting using ML, revealing a strong focus on certain key variables while others remain underexplored. As shown in the breakdown in Figure 13, a substantial portion of the literature, nearly half of all parameter estimation studies, focuses on predicting demand. However, a wide range of other critical parameters are also being estimated, as detailed in Table 6.

The most commonly predicted parameter is undoubtedly demand, which is a fundamental source of uncertainty in supply chain planning. Numerous studies apply a variety of ML models to forecast customer demand, which then serves as a direct input for *inventory control* [60, 65], production planning [43], and pricing optimization models [110]. Following demand, several other categories have received notable attention. Travel Time / Routing Cost is predicted in studies to inform *VRP* models, often accounting for dynamic conditions [36, 95, 191]. Another key area is Customer / Data Grouping, where unsupervised learning is often used as a preprocessing step to segment customers or data points before optimization [46, 130, 174]. Resource Availability / Yield is also a focus, particularly in natural resource contexts, such as water management [40, 94].

In contrast, several critical parameters appear to be significantly understudied. While important, parameters related to Supplier Performance / Risk [66, 172], Production/Sourcing Costs [49, 194], and Lead Time [179] are estimated in only a handful of articles. This analysis highlights a major research gap: while the field has matured in using ML for demand forecasting, there is a substantial opportunity to apply these techniques to a wider range of uncertainty drivers, particularly those related to supply-side risks, internal process variability, and reverse logistics.

Table 6: Type of supply chain parameters predicted with machine learning.

| Parameter Type | Count | Reference |
|---|---|---|
| Demand | 22 | [5, 11, 21, 32, 43, 53, 56, 60, 65, 77, 78, 101, 109, 110, 112, 117, 133, 134, 157, 164, 179, 190, 196, 228, 232] |
| Travel Time / Routing Cost | 5 | [36, 49, 89, 95, 191] |
| Customer / Data Grouping | 5 | [46, 64, 130, 174, 187, 192] |
| Resource Availability / Yield | 5 | [40, 52, 67, 94, 142] |
| Supplier Performance / Risk | 4 | [66, 98, 172, 210] |
| Costs (Production/Sourcing) | 3 | [49, 56, 194] |
| Quality / State of Product | 2 | [155, 156] |

Continued on next page

Table 6: Type of supply chain parameters predicted with machine learning. (Continued)

| Parameter Type | Count | Reference |
|---|---|---|
| Order Quantity | 2 | [210, 232] |
| Lead Time | 2 | [56, 179] |
| Asset Performance / Capacity | 1 | [56] |
| Real-Time System State | 1 | [62] |
| Service Level | 1 | [228] |
| Location | 1 | [196] |

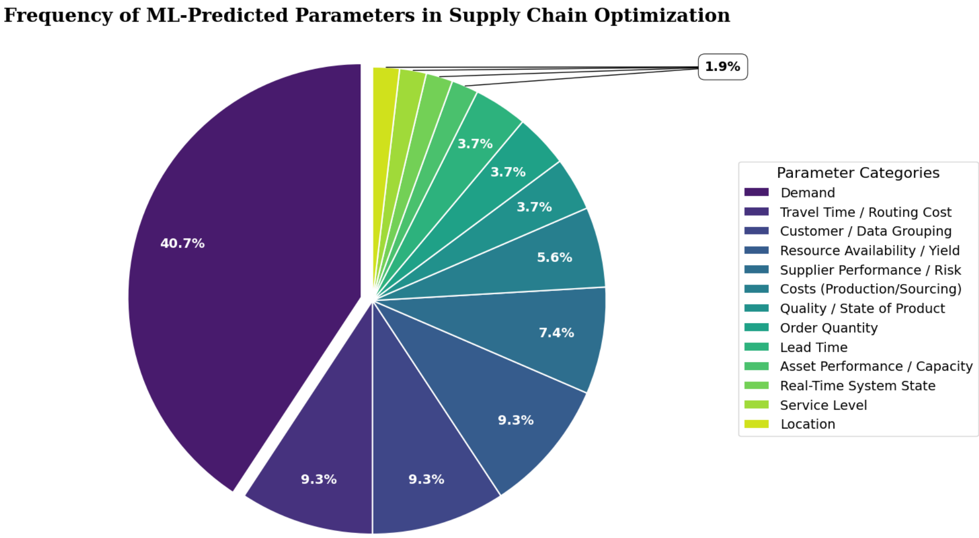


**Fig. 13** Frequency distribution of machine learning-predicted parameters in supply chain optimization.

## 4.9 Machine Learning and Optimization in Real-Time Supply Chains

A significant and growing trend identified in our review is the application of integrated ML and optimization techniques to Real-Time Supply Chains. Unlike traditional planning, which operates on static, pre-collected data, real-time supply chains are characterized by a continuous flow of streaming data from sources such as GPS, IoT sensors, and live order systems. This environment introduces unique challenges, including the need for online or rolling-horizon optimization, strict latency constraints where decisions must be made in seconds, and the ability to adapt to dynamically arriving information. Our analysis reveals that the synergy between ML and optimization is particularly powerful in addressing these challenges.

The most prominent application in this area is dynamic online vehicle routing. In these problems, customer requests are not known in advance but are revealed over time, requiring routes to be constantly re-optimized. For example, [161] develops a deep reinforcement learning model specifically for dynamic and uncertain *VRP*, while [233] proposes a neural combinatorial optimization approach for online routing. Similarly, [18] enriches machine learning with combinatorial optimization to solve dynamic *VRP*s with time windows. This real-time routing challenge extends to specialized domains, such as the dynamic pickup and allocation problem with fairness considerations studied in [154], and route planning in dynamic road networks where traffic conditions change in real-time [114]. In these contexts, RL and DRL are the dominant ML techniques, as they are naturally suited for learning adaptive decision policies that can react instantly to new information without having to re-solve a complex optimization model from scratch.

Beyond routing, real-time decision-making is also critical in other supply chain functions. For example, [62] focuses on real-time decision-making for dynamic distribution in the grocery supply chain, where demand and inventory levels are constantly changing. In manufacturing, [183] addresses the need for real-time decision-making in the remanufacturing of medical equipment, a problem that became critical during the pandemic. In these applications, ML models are often used to rapidly process streaming data and predict the state of the system, while optimization algorithms (or RL policies) make immediate operational decisions. This highlights a clear trend: as supply chains become more digitized and data-rich, the integration of ML and optimization is the key enabler for shifting from periodic, offline planning to continuous, real-time operational control.

## 4.10 AI–Optimization Synergies for Viable and Resilient Supply Chains

Recent work shows that AI and optimization jointly enable supply chain viability, survival, and adaptation under structural change [28, 84, 158, 182, 189]. Consistent with our taxonomy of the role, purpose, and formulation of ML, the reviewed studies focus on five recurring capabilities. First, *anticipation* uses ML to learn uncertainty and disruption patterns (e.g., neural networks, generative models) for parameter estimation; the resulting inputs inform stochastic or robust LP/MIP models to produce proactive inventory, capacity, and routing plans [86, 113]. Second, *adaptive control* relies on policy learning and state estimation via RL/DRL, often with encoder–decoder or Transformer controllers and digital twins, to generate solutions in real time, while optimization runs on a rolling horizon with feasibility and constraint enforcement (LP/MIP/NLP) to deliver responsive and valid actions [20, 83]. Adaptive control also includes an adaptive multi-echelon supply chain framework that handles long-horizon spatially shifting shocks under unknown dynamics, strengthening

**Table 7** Synergy map of capabilities, role and purpose of ML, and formulations.

| **Capability** | **Primary ML Role (methods)** | **Purpose (this review)** | **Typical Formulation** | **Joint Advantage** |
|---|---|---|---|---|
| Anticipation | Learn uncertainty/disruptions; forecasting (neural networks, generative models) | Parameter estimation | Stochastic/robust; LP/MIP with learned inputs | Proactive plans that hedge data-driven risks |
| Adaptive Control | Policy learning and state estimation (RL/DRL; encoder–decoder/Transformers) | Solution | Rolling-horizon LP/MIP/NLP/RL with feasibility checks and constraint enforcement | Real-time responsiveness with constrained feasibility |
| Structural Redesign | Segmentation; scenario/model generation (clustering, GenAI) | Formulation; parameter estimation | Network design (MIP); facility location; capacity allocation | Agile reconfiguration under structural shifts |
| Recovery Planning | Impact prediction; time-to-recovery estimation (neural networks; regression) | Parameter estimation; solution | Multi-period scheduling; resource reallocation (MIP) | Recovery trajectories minimizing downtime and cost |
| Robustness & Risk | Scenario synthesis; anomaly/outlier detection (GANs, clustering) | Formulation | Min–max/robust; chance-constrained; distributionally robust optimization (DRO) | Resilient decisions under data perturbations and shocks |

viability without probabilistic uncertainty models [15, 26, 90, 226, 227]. Third, *structural redesign* applies clustering and generative methods to propose network scenarios for formulation and parameter estimation. The MIP network design, facility location and capacity models then evaluate the alternatives, enabling agile restructuring under structural changes [80]. Fourth, *recovery planning* combines impact and time-to-recovery prediction with multi-period MIP scheduling and resource reallocation to compute cost-aware trajectories back to steady state. Finally, *robustness and risk* are based on scenario synthesis and anomaly or outlier detection to shape min–max, chance-constrained, and distributionally robust models, making decisions resilient to perturbations and shocks. These patterns align with our quantitative evidence (e.g., RL with MIP for solution generation in logistics; neural networks for parameter estimation) [16, 90, 92, 221, 226]. Table 7 maps each capability to the dominant role and purpose of ML, the typical formulations, and the resulting joint advantage for viable and resilient supply chains.

# 5 Proposed Research Framework

Based on a comprehensive analysis of the literature, we propose a structured research framework to guide future work at the intersection of machine learning and mathematical optimization in tangible supply chains. This framework, illustrated as a flowchart in Figure 14, is designed to help researchers navigate the key decisions involved in developing a well-defined research project, from initial ideation to final validation. The framework is divided into four primary stages: Problem Identification, Model Formulation, Machine Learning Integration, and Solution Evaluation.

The process begins with Stage 1: Problem Identification, where a researcher moves from a "Broad Research Area" to a well-defined problem. The first step is to "Define the Application Domain" (Step 1) by selecting a specific supply chain context, such as logistics, healthcare, or manufacturing. This choice is critical because it frames the unique challenges of the problem. Following this, the researcher must "Identify the Specific Problem" (Step 2) within that domain, such as a *Vehicle Routing Problem (VRP)*, *inventory control*, or *scheduling* task. The final step in this stage is to "Characterize the Problem Nature" (Step 3) by identifying its core properties—whether it involves discrete decisions, significant uncertainty, or non-linear relationships—which directly informs the subsequent modeling choices.

Stage 2: Model Formulation translates the defined problem into a formal mathematical structure. As shown in Figure 14, this stage is guided by two critical decision nodes. The first decision asks whether there is significant uncertainty. If "Yes," the framework recommends formulating the problem using "Stochastic or Robust Optimization" (Step 4a). If "No," a "Deterministic Optimization" model (Step 4b) is appropriate. The second decision node asks if there are discrete variables. If "Yes," the framework points to "Integer Programming (IP)" formulations (Step 5a). If the decisions are purely continuous, "Continuous Programming" models such as LP or NLP are suitable (Step 5b). Following one of these paths results in a "Complete Mathematical Model", which serves as the foundation for the integration stage.

Stage 3: Machine Learning Integration is the central and most detailed stage, determining how ML will be used to enhance the optimization model. This stage begins with the question: "What is the primary purpose of ML?". The framework presents three distinct pathways. The "Prediction" path leads to "Parameter Estimation" (Step 6a), where supervised learning models (such as Neural Networks, Regression, and LSTMs for time-series) or unsupervised methods (such as Clustering) are used to generate data-driven inputs for the optimization model, such as demand forecasts or travel times. The "Solution" path leads to "Solution Generation" (Step 6b), where techniques such as Reinforcement Learning (RL/DRL) or neural combinatorial solvers (e.g., Transformers) are trained to act as heuristics or end-to-end solvers for dynamic problems such as routing and scheduling. Finally, the "Formulation" path leads to "Model Formulation & Reformulation" (Step 6c), where ML is used to learn parts of the model itself, for instance by creating neural network surrogates for complex constraints or learning effective branching rules. As our analysis indicates, reformulation remains a significant research opportunity. The selection of one of these strategies culminates in an "Integrated ML-Optimization Framework".

The final step, Stage 4: Solution Evaluation, emphasizes the critical need for rigorous validation of the integrated framework. Once a solution is generated, it must be evaluated through a comprehensive process (Step 7). This includes comparing the "Solution Quality vs. Baselines" of traditional solvers or heuristics, assessing computational metrics such as "Runtime" and "Latency", and verifying the "Feasibility" of the generated solutions. Furthermore, for models dealing with uncertainty, performing "Robustness / Stress Tests" is crucial. Finally, the "Scalability" of the approach should be tested by evaluating its performance on instances of varying sizes to ensure its practical applicability. This final stage ensures that the proposed framework is not only novel, but also effective and reliable for real-world deployment.

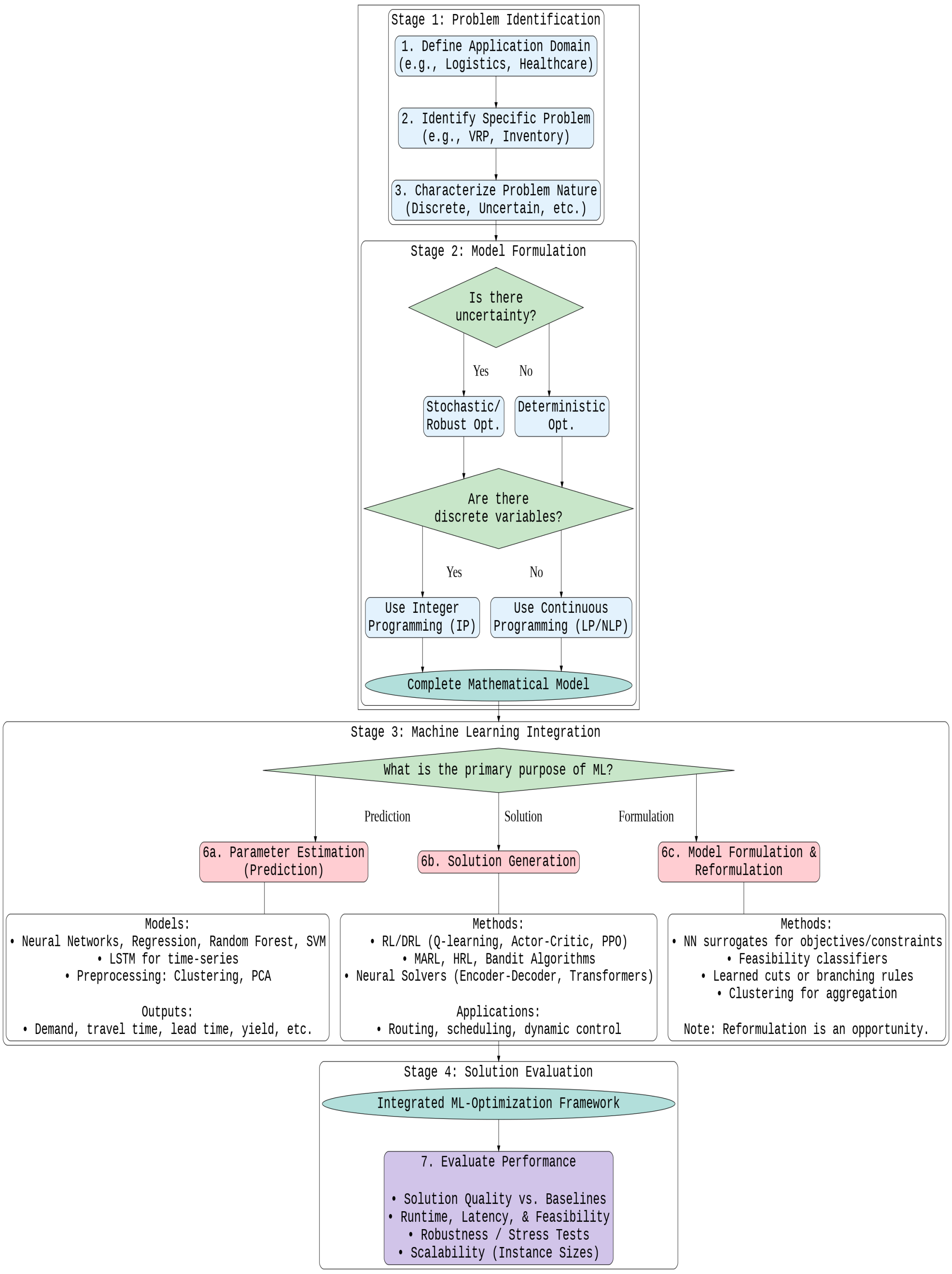


**Fig. 14** A proposed research framework to guide the development of integrated Machine Learning and Mathematical Optimization models in supply chain management

# 6 Conclusion and Future Directions

In this paper, we systematically reviewed the current literature on applications of machine learning in SC optimization. After a detailed review of the titles and abstracts of 2,253 identified papers, we ended up with 199 articles for full-text review. We constructed our research questions to assess the applicability of machine learning techniques in different parts of optimization problems used to model SC operations, based on the characteristics of the problem. These different categorizations provide insight into how ML is operationalized in optimization models across various domains, highlighting both mature applications and areas for future research. In the end, we suggest a general framework for future researchers to be able to apply machine learning algorithms in SC optimization. Through this research, we ensured that this paper serves as a comprehensive reference that captures the majority of publications that use machine learning methods alongside mathematical programming in SC management in the last 18 years.

Our findings indicate that ML methods such as neural networks and deep learning can scale effectively in data-rich environments and support optimization by learning complex relationships [225]. Architectures such as graph neural networks can leverage underlying network structures, although performance is problem dependent. Neural networks are also used as heuristic tools in integer and combinatorial optimization, for example to predict variable fixings or guide branching, which helps reduce computational effort. Reinforcement learning performs well in solving sequential and dynamic optimization problems, especially when explicit modeling is difficult, though the results vary by problem and tuning.

The most popular areas of research are the application of reinforcement learning and deep reinforcement learning in formulating or solving *VRP* problems, and the use of neural networks in solving problems with network structures. *VRP* problems are the most studied problems, followed by *Classical Supply Chain*, *Inventory Control*, and *TSP* problems.

These fields can be further extended to include broader sets of problems in the supply chain. Interestingly, *knapsack*, *bi-level*, *lot-sizing*, and general dynamic programming problems are not frequently observed in modeling supply chains through ML-optimization, despite their proper structure for ML modeling. Despite the great promise of integrating ML models into supply chain optimization problems, one should be cautious about the appropriate use of these techniques. We highly recommend the specification of the characteristics of the problem to ensure the fit of the ML model for the problem. In addition, some ML models are prone to memorizing the data and overfitting. Exhaustive validation techniques are suggested to not only ensure the proper performance of ML models in the problem of interest, but also enable the generalization of the use case to a wider set of problems.

In the reviewed literature, the central tension in the implementation of machine learning (ML) for SC optimization is not conceptual promise, but operational reliability. Data-related frictions remain primary: many supply networks exhibit fragmented, noisy, or delayed data, which propagates uncertainty through learned models and undermines prescriptive quality [23]. Even when accurate predictions are achievable, limited interpretability reduces managerial trust and hampers auditability of high-stakes decisions; recent explainable-AI (XAI) and neurosymbolic approaches offer remedies, but are not yet turnkey for complex, multi-echelon contexts [103]. Distribution shift is a second structural challenge: models calibrated to historical regimes degrade under shocks (pandemics, geopolitical disruptions), urging firms to design for viability and survivability rather than steady-state efficiency [79]. A third concern is the coupling of prediction and optimization: when decisions influence future uncertainty, naive pipelines risk endogeneity and mis-specification; prescriptive analytics emphasizes causal reasoning and integrated learning-to-optimize frameworks [19, 222, 225]. Finally, deployment at operational cadence is nontrivial: hybrid ML–OR stacks can be computationally intensive, integration-heavy and sensitive to latency constraints, especially in transportation, *inventory control* and real-time pricing [105]. Collectively, these issues argue for a design position that privileges data governance, explainability, robustness to shift, and decision-aware learning.

The literature converges on a pragmatic view: ML can materially enhance forecasting, sensing, and decision support in supply chains, but unqualified adoption risks brittleness. Effective practice requires moving beyond prediction toward prescriptive, interpretable, and robust pipelines that respect data realities and operational constraints. Integrating causal structure, uncertainty modeling, and computationally efficient optimization is required for reliable value creation [19, 23]. As supply chains face increasingly non-stationary environments, the target is not merely higher average performance, but resilient, auditable systems that maintain decision quality under stress [79]. Ultimately, the tight integration of ML and optimization is the key enabler of supply chain viability, allowing systems to effectively anticipate disruptions, adapt controls in real-time, and recover through intelligent structural redesign [84].

Our review highlights that scale is a key factor in determining the effectiveness of the method. Although general-purpose learning models can adapt across problem sizes, more complex approaches, such as deep reinforcement learning, tend to perform best in small to medium instances unless paired with simplification or decomposition strategies. For large-scale problems, hybrid approaches that combine learned representations with classical optimization techniques remain the most scalable [223, 224]. Hybrid models that integrate deep reinforcement learning with clustering and agent-based simulation offer promising avenues for addressing the large-scale complexity of supply chains. These approaches combine the scalability of unsupervised learning with the adaptability of reinforcement learning to optimize dynamic, data-driven decisions to logistics problems such as vaccine distribution [26] and multi-dimensional resource allocation [14], highlighting their potential for broader applications in supply chain logistics.

Generative AI is also emerging as a promising direction in supply chain optimization. Although still limited in current applications, it shows strong potential for generating synthetic data, formulating models via natural language, and simulating complex decision scenarios. Recent work suggests that large language models (LLMs) can support the design of effective heuristics and algorithms for solving large-scale combinatorial supply chain problems, including machine scheduling [27]. By bridging human language and mathematical models, generative AI creates new opportunities for data generation, model formulation, and decision support. Recent systems such as OptiMUS [4] show that LLMs can model and solve complex optimization tasks from natural language descriptions, suggesting a strong potential of LLMs to formulate and address various supply chain problems. Intelligent optimization methods that integrate predictive models, such as epidemic dynamics or biological growth, with logistics optimization demonstrate the promise of simultaneous prediction and optimization in health-care facility location [15, 221], spatio-temporal resource allocation and cancer treatment [91, 93, 226], and equitable vaccine allocation during pandemics [90].

Future work should target viability- and adaptability-oriented gaps: ML-aided reformulation (e.g., generative surrogates and learned cuts) that preserves feasibility and robustness; data-driven uncertainty sets for robust and DRO with stress-tested performance under non-stationarity; and scalable policy–optimization hybrids for network-level control and recovery. Additionally, we propose the following areas as promising directions for future research according to the current state-of-the-art: i Developing methods that detect, quantify, and adapt to changes (covariate and label shift) while preserving feasibility and service levels. This can be achieved by combining distributionally robust optimization with online and continual learning for closed-loop control [79]. ii Creating decision-aware and causal pipelines by embedding causal identification and counterfactual prediction within prescriptive workflows to mitigate endogeneity and feedback loops in pricing, assortment, and capacity allocation [19]. iii Providing explainable prescriptions by providing tailored to optimization outputs, producing human-interpretable rationales, sensitivity attributions, and policy stability certificates that meet governance requirements [103]. iv Improving scalability of models by including latency in the models' architectures and integrating ML and OR components, which leads to inference, re-optimization, and data updates in real-time, suitable for transportation and fulfillment applications [105, 225]. v Ensuring the reproducibility of the model by creating open benchmarks with realistic data imperfections (missingness, delays, structural breaks) and standardized metrics that jointly evaluate predictive accuracy, optimization quality, robustness, and interpretability [23].

## Conflict of Interest

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

## References

[1] Aslam, T., Amos, H.C.N.: Multi-objective optimization for supply chain management: A literature review and new development. In: 2010 8th International Conference on Supply Chain Management and Information, pp. 1–8 (2010)

[2] Abbasi, B., Babaei, T., Hosseinifard, Z., Smith-Miles, K., Dehghani, M.: Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management. Computers and Operations Research **119** (2020) https://doi.org/10.1016/j.cor.2020.104941

[3] Akbari, M., Do, T.N.A.: A systematic review of machine learning in logistics and supply chain management: current trends and future directions. Benchmarking **28**(10), 2977–3005 (2021) https://doi.org/10.1108/BIJ-10-2020-0514

[4] AhmadiTeshnizi, A., Gao, W., Brunborg, H., Talaei, S., Lawless, C., Udell, M.: Optimus-0.3: Using large language models to model and solve optimization problems at scale. arXiv preprint arXiv:2407.19633 (2024)

[5] Aghaie, A., Heidary, M.H.: Simulation-based optimization of a stochastic supply chain considering supplier disruption: Agent-based modeling and reinforcement learning. Scientia Iranica **26**(6), 3780–3795 (2019) https://doi.org/10.24200/sci.2018.20789

[6] Arishi, A., Krishnan, K.: A multi-agent deep reinforcement learning approach for solving the multi-depot vehicle routing problem. Journal of Management Analytics **10**(3), 493–515 (2023) https://doi.org/10.1080/23270012.2023.2229842

[7] Anuar, W.K., Lee, L.S., Seow, H.V., Pickl, S.: A multi-depot vehicle routing problem with stochastic road capacity and reduced two-stage stochastic integer linear programming models for rollout algorithm. Mathematics **9**(13) (2021) https://doi.org/10.3390/math9131572

[8] Anderson, T., Martínez, S.: Distributed resource allocation with binary decisions via Newton-like Neural Network dynamics. Automatica **128** (2021) https://doi.org/10.1016/j.automatica.2021.109564

[9] Altaf, M.M., Roshdy, A.S., AlSagri, H.S.: Deep reinforcement learning model for blood bank vehicle routing multi-objective optimization. Computers, Materials and Continua **70**(2), 3955–3967 (2022) https://doi.org/10.32604/cmc.2022.019448

[10] Akbarian-Saravi, N., Mobini, M., Rabbani, M.: Development of a comprehensive decision support tool for strategic and tactical planning of a sustainable bioethanol supply chain: Real case study, discussions and policy implications. Journal of Cleaner Production **244** (2020) https://doi.org/10.1016/j.jclepro.2019.118871

[11] Ashtab, S., Tosarkani, B.M.: Scenario-based multi-objective optimisation model based on supervised machine learning to configure a plastic closed-loop supply chain network. International Journal of Business Performance and Supply Chain Modelling **14**(1), 106–128 (2023) https://doi.org/10.1504/IJBPSCM.2023.130469

[12] Albayrak Ünal, Ö., Erkayman, B., Usanmaz, B.: Applications of artificial intelligence in inventory management: A systematic review of the literature. Archives of computational methods in engineering **30**(4), 2605–2625 (2023) https://doi.org/10.1007/s11831-022-09879-5

[13] Bandaru, S., Aslam, T., Ng, A.H.C., Deb, K.: Generalized higher-level automated innovization with application to inventory management. European Journal of Operational Research **243**(2), 480–496 (2015) https://doi.org/10.1016/j.ejor.2014.11.015

[14] Bushaj, S., Büyüktahtakın, İ.E.: A k-means supported reinforcement learning framework to multi-dimensional knapsack. Journal of Global Optimization **89**(3), 655–685 (2024) https://doi.org/10.1007/s10898-024-01364-6

[15] Büyüktahtakın, .I., des-Bordes, E., Kıbış, E.: A new epidemics–logistics model: Insights into controlling the ebola virus disease in west africa. European Journal of Operational Research **265**(3), 1046–1063 (2018) https://doi.org/10.1016/j.ejor.2017.08.037

[16] Bertsimas, D., Digalakis Jr, V., Jacquillat, A., Li, M.L., Previero, A.: Where to locate covid-19 mass vaccination facilities? Naval Research Logistics (NRL) **69**(2), 179–200 (2022)

[17] Bhosekar, A., Ierapetritou, M.: A framework for supply chain optimization for modular manufacturing with production feasibility analysis. Computers and Chemical Engineering **145** (2021) https://doi.org/10.1016/j.compchemeng.2020.107175

[18] Baty, L., Jungel, K., Klein, P.S., Parmentier, A., Schiffer, M.: Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows. Transportation Science **58**(4), 708–725 (2024) https://doi.org/10.1287/trsc.2023.0107

[19] Bertsimas, D., Kallus, N.: From predictive to prescriptive analytics. Management Science **66**(3), 1025–1044 (2020) https://doi.org/10.1287/mnsc.2018.3253

[20] Bensoussan, A., Li, Y., Nguyen, D.P.C., Tran, M.-B., Yam, S.C.P., Zhou, X.: Chapter 16 - machine learning and control theory. In: Trélat, E., Zuazua, E. (eds.) Numerical Control: Part A. Handbook of Numerical Analysis, vol. 23, pp. 531–558. Elsevier, Amsterdam (2022). https://doi.org/10.1016/bs.hna.2021.12.016

[21] Boccia, M., Mancuso, A., Masone, A., Murino, T., Sterle, C.: New features for customer classification in the Flying Sidekick Traveling Salesman Problem. Expert Systems with Applications **247** (2024) https://doi.org/10.1016/j.eswa.2023.123106

[22] Bishop, C.M., Nasrabadi, N.M.: Pattern Recognition and Machine Learning vol. 4. Springer, New York (2006)

[23] Baryannis, G., Validi, S., Dani, S., Antoniou, G.: Supply chain risk management and artificial intelligence: state of the art and future research directions.

International Journal of Production Research **57**(7), 2179–2202 (2019) https://doi.org/10.1080/00207543.2018.1530476

[24] Bootaki, B., Zhang, G.: A location-production-routing problem for distributed manufacturing platforms: A neural genetic algorithm solution methodology. International Journal of Production Economics **275** (2024) https://doi.org/10.1016/j.ijpe.2024.109325

[25] Chobar, A.P., Adibi, M.A., Kazemi, A.: Multi-objective hub-spoke network design of perishable tourism products using combination machine learning and meta-heuristic algorithms. Environment, Development and Sustainability (2022) https://doi.org/10.1007/s10668-022-02350-2

[26] Cibaku, E., Büyüktahtakın, İ.E.: An adaptive k-means and reinforcement learning (rl) algorithm to effective vaccine distribution. Computers & Operations Research, 107275 (2025) https://doi.org/10.1016/j.cor.2025.107275

[27] Çetinkaya, İ.O., Büyüktahtakın, İ.E., Shojaee, P., Reddy, C.K.: Discovering heuristics with large language models (LLMs) for mixed-integer programs: Single-machine scheduling. Computers & Operations Research, 107325 (2025) https://doi.org/10.1016/j.cor.2025.107325

[28] Chen, W., Choi, T.-M., Dolgui, A., Ivanov, D., Pesch, E.: Digital manufacturing and supply chain: creating benefits through operations research and artificial intelligence. Annals of Operations Research **344**(2), 569–574 (2025) https://doi.org/10.1007/s10479-024-06450-2

[29] Chi, C.Y., Chen, D.F., Doan, H.P., Kuo, C.H.: UAV parcel Delivery System with Deep Reinforcement Learning Based Collision Avoidance and Route Optimization. International Journal of Automation and Smart Technology **14**(1), 8–18 (2024) https://doi.org/10.5875/P25FHT10

[30] Christou, I.T., Psathas, F., Rentizelas, A., Papadakis, A., Georgiou, P.N., Anastasopoulos, D., Lappas, P.: A heuristic for improving clustering in biomass supply chains. International Journal of Systems Science: Operations and Logistics **11**(1) (2024) https://doi.org/10.1080/23302674.2024.2378859

[31] Chen, B., Qu, R., Bai, R., Laesanklang, W.: A variable neighborhood search algorithm with reinforcement learning for a real-life periodic vehicle routing problem with time windows and open routes. RAIRO - Operations Research **54**(5), 1467–1494 (2020) https://doi.org/10.1051/ro/2019080

[32] Chamurally, S., Rieck, J.: A practical and robust approach for solving the multi-compartment vehicle routing problem under demand uncertainty using machine learning. Networks **84**(3), 300–325 (2024) https://doi.org/10.1002/net.22231

[33] Chen, Z.Y., Sun, M., Han, X.X.: Prediction-driven collaborative emergency medical resource allocation with deep learning and optimization. Journal of the Operational Research Society **74**(2), 590–603 (2023) https://doi.org/10.1080/01605682.2022.2101953

[34] Cai, Y., Wang, H., Yue, W., Xie, Y., Liang, Q.: An integrated approach for reducing spatially coupled water-shortage risks of Beijing-Tianjin-Hebei urban agglomeration in China. Journal of Hydrology **603** (2021) https://doi.org/10.1016/j.jhydrol.2021.127123

[35] Cai, H., Xu, P., Tang, X., Lin, G.: Solving the Vehicle Routing Problem with Stochastic Travel Cost Using Deep Reinforcement Learning. Electronics (Switzerland) **13**(16) (2024) https://doi.org/10.3390/electronics13163242

[36] Chu, H., Zhang, W., Bai, P., Chen, Y.: Data-driven optimization for last-mile delivery. Complex and Intelligent Systems **9**(3), 2271–2284 (2023) https://doi.org/10.1007/s40747-021-00293-1

[37] Chen, J., Zong, Z., Zhuang, Y., Yan, H., Jin, D., Li, Y.: Reinforcement Learning for Practical Express Systems with Mixed Deliveries and Pickups. ACM Transactions on Knowledge Discovery from Data **17**(3) (2023) https://doi.org/10.1145/3546952

[38] Dietterich, T.: Overfitting and undercomputing in machine learning. ACM computing surveys (CSUR) **27**(3), 326–327 (1995) https://doi.org/10.1145/212094.212114

[39] Dariane, A.B., Moradi, A.M.: Comparative analysis of evolving artificial neural network and reinforcement learning in stochastic optimization of multireservoir systems. Hydrological Sciences Journal **61**(6), 1141–1156 (2016) https://doi.org/10.1080/02626667.2014.986485

[40] Delavar, M., Moghadasi, M., Morid, S.: Real-Time Model for Optimal Water Allocation in Irrigation Systems during Droughts. Journal of Irrigation and Drainage Engineering **138**(6), 517–524 (2012) https://doi.org/10.1061/(ASCE)IR.1943-4774.0000440

[41] Dornemann, J.: Solving the capacitated vehicle routing problem with time windows via graph convolutional network assisted tree search and quantum-inspired computing. Frontiers in Applied Mathematics and Statistics **9** (2023) https://doi.org/10.3389/fams.2023.1155356

[42] Deng, J., Wang, J., Wang, X., Cai, Y., Liu, P.: Multi-Task Multi-Objective Evolutionary Search Based on Deep Reinforcement Learning for Multi-Objective Vehicle Routing Problems with Time Windows. Symmetry **16**(8) (2024) https://doi.org/10.3390/sym16081030

[43] Dong, W., Zhang, X., Guo, H., Qu, J., Dong, L., Zhang, H., Jiang, Y.: Optimization of Sustainable Automotive Manufacturing Supply Chain with Demand Prediction: A Combinatorial Framework. International Journal of Industrial Engineering : Theory Applications and Practice **30**(4), 1042–1051 (2023) https://doi.org/10.23055/ijietap.2023.30.4.9035

[44] Ezazipour, S., Golbabai, A.: A globally convergent neurodynamics optimization model for mathematical programming with equilibrium constraints. Kybernetika **56**(3), 383–409 (2020) https://doi.org/10.14736/kyb-2020-3-0383

[45] Elmachtoub, A.N., Grigas, P.: Smart "predict, then optimize". Management Science **68**(1), 9–26 (2022) https://doi.org/10.1287/mnsc.2020.3922

[46] Ewbank, H., Wanke, P., Hadi-Vencheh, A.: An unsupervised fuzzy clustering approach to the capacitated vehicle routing problem. Neural Computing and Applications **27**(4), 857–867 (2016) https://doi.org/10.1007/s00521-015-1901-4

[47] Fitzpatrick, J., Ajwani, D., Carroll, P.: A scalable learning approach for the capacitated vehicle routing problem. Computers and Operations Research **171** (2024) https://doi.org/10.1016/j.cor.2024.106787

[48] Fellek, G., Farid, A., Gebreyesus, G., Fujimura, S., Yoshie, O.: Graph Transformer with Reinforcement Learning for Vehicle Routing Problem. IEEJ Transactions on Electrical and Electronic Engineering **18**(5), 701–713 (2023) https://doi.org/10.1002/tee.23771

[49] Frias, N., Johnson, F., Valle, C.: Hybrid Algorithms for Energy Minimizing Vehicle Routing Problem: Integrating Clusterization and Ant Colony Optimization. IEEE Access **11**, 125800–125821 (2023) https://doi.org/10.1109/ACCESS.2023.3325787

[50] Furian, N., O'Sullivan, M., Walker, C., Cela, E.: A machine learning-based branch and price algorithm for a sampled vehicle routing problem. OR Spectrum **43**(3), 693–732 (2021) https://doi.org/10.1007/s00291-020-00615-8

[51] Frederico, G.F.: ChatGPT in Supply Chains: Initial Evidence of Applications and Potential Research Agenda. Logistics **7**(2) (2023) https://doi.org/10.3390/logistics7020026

[52] Flores, H., Villalobos, J.R.: A stochastic planning framework for the discovery of complementary, agricultural systems. European Journal of Operational Research **280**(2), 707–729 (2020) https://doi.org/10.1016/j.ejor.2019.07.053

[53] Fu, Y., Wang, Y., Gao, K., Suganthan, P.N., Huang, M.: Integrated scheduling of multi-constraint open shop and vehicle routing: Mathematical model and learning-driven brain storm optimization algorithm. Applied Soft Computing **163** (2024) https://doi.org/10.1016/j.asoc.2024.111943

[54] Fu, J., Yang, Y.: Construction of an intelligent optimization model for water resource allocation under drought conditions. Applied Mathematics and Nonlinear Sciences **9**(1) (2024) https://doi.org/10.2478/amns-2024-1162

[55] Guo, Y., Chen, T., Boulaksil, Y., Xiao, L., Allaoui, H.: Collaborative planning of multi-tier sustainable supply chains: A reinforcement learning enhanced heuristic approach. Computers and Industrial Engineering **185** (2023) https://doi.org/10.1016/j.cie.2023.109669

[56] Gumte, K.M., Devi Pantula, P., Miriyala, S.S., Mitra, K.: Data driven robust optimization for handling uncertainty in supply chain planning models. Chemical Engineering Science **246** (2021) https://doi.org/10.1016/j.ces.2021.116889

[57] Golbabai, A., Ezazipour, S.: A projection-based recurrent neural network and its application in solving convex quadratic bilevel optimization problems. Neural Computing and Applications **32**(8), 3887–3900 (2020) https://doi.org/10.1007/s00521-019-04391-7

[58] Giuffrida, N., Fajardo-Calderin, J., Masegosa, A.D., Werner, F., Steudter, M., Pilla, F.: Optimization and Machine Learning Applied to Last-Mile Logistics: A Review. Sustainability (Switzerland) **14**(9) (2022) https://doi.org/10.3390/su14095329

[59] Gumus, A.T., Guneri, A.F., Keles, S.: Supply chain network design using an integrated neuro-fuzzy and MILP approach: A comparative design study. Expert Systems with Applications **36**(10), 12570–12577 (2009) https://doi.org/10.1016/j.eswa.2009.05.034

[60] Galli, L., Levato, T., Schoen, F., Tigli, L.: Prescriptive analytics for inventory management in health care. Journal of the Operational Research Society **72**(10), 2211–2224 (2021) https://doi.org/10.1080/01605682.2020.1776167

[61] Gutierrez, J.C., Polo Triana, S.I., León Becerra, J.S.: Benefits, challenges, and limitations of inventory control using machine learning algorithms: literature review. OPSEARCH (2024) https://doi.org/10.1007/s12597-024-00839-0

[62] Geraeli, M., Roghanian, E.: A novel adjusted real-time decision-making for dynamic distribution in the grocery supply chain. Journal of Modelling in Management (2024) https://doi.org/10.1108/JM2-02-2024-0048

[63] Garcia, D.J., You, F.: Supply chain design and optimization: Challenges and opportunities. Computers & Chemical Engineering **81**, 153–170 (2015) https://doi.org/10.1016/j.compchemeng.2015.03.015

[64] Han, X.: Path Planning Algorithm for the Multiple Depot Vehicle Routing Problem Based on Parallel Clustering. Scientific Programming **2023** (2023) https://doi.org/10.1155/2023/7588595

[65] Henao-Baena, C.A., Zuluaga-Zuluaga, B., Galeano-Castro, J., Marín-García, E.J., Calvo-Salcedo, A.F.: Methodology for Inventory Management in Neighborhood Stores Using Machine Learning and Integer Linear Programming. Ingenieria (Colombia) **29**(1) (2024) https://doi.org/10.14483/23448393.19423

[66] Hu, S., Dong, Z.S., Dai, R.: A machine learning based sample average approximation for supplier selection with option contract in humanitarian relief. Transportation Research Part E: Logistics and Transportation Review **186** (2024) https://doi.org/10.1016/j.tre.2024.103531

[67] Hou, J., Fan, X., Liu, R.: Optimal spatial allocation of irrigation water under uncertainty using the bilayer nested optimisation algorithm and geospatial technology. International Journal of Geographical Information Science **30**(12), 2462–2485 (2016) https://doi.org/10.1080/13658816.2016.1181264

[68] Han, Y., Huang, Y., Jia, S., Liu, J.: An interval-parameter fuzzy linear programming with stochastic vertices model for water resources management under uncertainty. Mathematical Problems in Engineering **2013** (2013) https://doi.org/10.1155/2013/942343

[69] Han, Y., Huang, Y., Wang, G.: Interval-parameter linear optimization model with stochastic vertices for land and water resources allocation under dual uncertainty. Environmental Engineering Science **28**(3), 197–205 (2011) https://doi.org/10.1089/ees.2010.0160

[70] He, X., Li, C., Huang, T., Huang, J.: A recurrent neural network for solving bilevel linear programming problem. IEEE Transactions on Neural Networks and Learning Systems **25**(4), 824–830 (2014) https://doi.org/10.1109/TNNLS.2013.2280905

[71] He, B., Meng, Y., Tang, L.: An Off-Policy Reinforcement Learning-Based Adaptive Optimization Method for Dynamic Resource Allocation Problem. IEEE Transactions on Neural Networks and Learning Systems, 1–15 (2023) https://doi.org/10.1109/TNNLS.2023.3338237

[72] Hammler, P., Riesterer, N., Braun, T.: Fully dynamic reorder policies with deep reinforcement learning for multi-echelon inventory management. Informatik-Spektrum **46**(5-6), 240–251 (2023) https://doi.org/10.1007/s00287-023-01556-6

[73] Heydari, F., Saghafian, B., Delavar, M.: Coupled Quantity-Quality Simulation-Optimization Model for Conjunctive Surface-Groundwater Use. Water Resources Management **30**(12), 4381–4397 (2016) https://doi.org/10.1007/s11269-016-1426-3

[74] Hottung, A., Tierney, K.: Neural large neighborhood search for routing problems. Artificial Intelligence **313** (2022) https://doi.org/10.1016/j.artint.2022.103786

[75] Hansuwa, S., Velayudhan Kumar, M.R., Chandrasekharan, R.: Analysis of box and ellipsoidal robust optimization, and attention model based reinforcement learning for a robust vehicle routing problem. Sadhana - Academy Proceedings in Engineering Sciences **47**(2) (2022) https://doi.org/10.1007/s12046-022-01833-2

[76] Hou, Y., Wang, H., Huang, X.: A Q-learning-based multi-objective evolutionary algorithm for integrated green production and distribution scheduling problems. Engineering Applications of Artificial Intelligence **127** (2024) https://doi.org/10.1016/j.engappai.2023.107434

[77] Islam, S., Amin, S.H., Wardley, L.J.: Machine learning and optimization models for supplier selection and order allocation planning. International Journal of Production Economics **242** (2021) https://doi.org/10.1016/j.ijpe.2021.108315

[78] Islam, S., Amin, S.H., Wardley, L.J.: A supplier selection & order allocation planning framework by integrating deep learning, principal component analysis, and optimization techniques. Expert Systems with Applications **235** (2024) https://doi.org/10.1016/j.eswa.2023.121121

[79] Ivanov, D., Dolgui, A.: Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. a position paper motivated by covid-19 outbreak. International Journal of Production Research **58**(10), 2904–2915 (2020) https://doi.org/10.1080/00207543.2020.1750727

[80] Ivanov, D., Gusikhin, O.: Supply chain digital twin design and implementation at scale: A case study at the ford motor company and generalizations. Omega **139**, 103447 (2026) https://doi.org/10.1016/j.omega.2025.103447

[81] Irannezhad, E., Prato, C.G., Hickman, M.: An intelligent decision support system prototype for hinterland port logistics. Decision Support Systems **130** (2020) https://doi.org/10.1016/j.dss.2019.113227

[82] Iklassov, Z., Sobirov, I., Solozabal, R., Takac, M.: Reinforcement Learning Approach to Stochastic Vehicle Routing Problem With Correlated Demands. IEEE Access **11**, 87958–87969 (2023) https://doi.org/10.1109/ACCESS.2023.3306076

[83] Ivanov, D.: Conceptual and formal models for design, adaptation, and control of digital twins in supply chain ecosystems. Omega **137**, 103356 (2025) https://doi.org/10.1016/j.omega.2025.103356

[84] Ivanov, D.: A survey of system-cybernetic principles in supply chain resilience: viability, artificial intelligence, digital twins, and ecosystems. International Journal of Systems Science, 1–16 (2025) https://doi.org/10.1080/00207721.2025.2519204

[85] Jin, J., Cui, T., Bai, R., Qu, R.: Container port truck dispatching optimization

using Real2Sim based deep reinforcement learning. European Journal of Operational Research **315**(1), 161–175 (2024) https://doi.org/10.1016/j.ejor.2023.11.038

[86] Jackson, I., Ivanov, D., Dolgui, A., Namdar, J.: Generative artificial intelligence in supply chain and operations management: a capability-based framework for analysis and implementation. International Journal of Production Research **62**(17), 6120–6145 (2024) https://doi.org/10.1080/00207543.2024.2309309

[87] Jordan, M.I., Mitchell, T.M.: Machine learning: Trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015)

[88] Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Computing Surveys **31**(3), 264–323 (1999)

[89] Jovanović, A.D., Pamučar, D.S., Pejčić-Tarle, S.: Green vehicle routing in urban zones - A neuro-fuzzy approach. Expert Systems with Applications **41**(7), 3189–3203 (2014) https://doi.org/10.1016/j.eswa.2013.11.015

[90] Jozani, K., Sageer, N., Eldardiry, H., Tunc, S., Büyüktahtakın, İ.: A multi-echelon demand-driven supply chain model for proactive optimal control of epidemics: Insights from a COVID-19 study. arXiv preprint arXiv:2510.16969 (2025) https://doi.org/10.48550/arXiv.2510.16969

[91] Kıbış, E.Y., Büyüktahtakın, İ.E.: Optimizing multi-modal cancer treatment under 3d spatio-temporal tumor growth. Mathematical Biosciences **307**, 53–69 (2019) https://doi.org/10.1016/j.mbs.2018.10.010

[92] Kıbış, E.Y., Büyüktahtakın, İ.E., Haight, R.G., Akhundov, N., Knight, K., Flower, C.E.: A multistage stochastic programming approach to the optimal surveillance and control of the emerald ash borer in cities. INFORMS Journal on Computing **33**(2), 808–834 (2021)

[93] Kıbış, E., Büyüktahtakın, İ., Haight, R., Akhundov, N., Knight, K., Flower, C.: A multistage stochastic programming approach to the optimal surveillance and control of the emerald ash borer in cities. INFORMS Journal on Computing **33**(2), 808–834 (2021) https://doi.org/10.1287/ijoc.2020.0963

[94] Kazemi, M., Bozorg-Haddad, O., Fallah-Mehdipour, E., Loáiciga, H.A.: Inter-basin hydropolitics for optimal water resources allocation. Environmental Monitoring and Assessment **192**(7) (2020) https://doi.org/10.1007/s10661-020-08439-3

[95] Küp, E.T., Cebeci, S., Bayram, B., Aydın, G., Bozkaya, B., Akhavan-Tabatabaei, R.: An Integrated Framework for Dynamic Vehicle Routing Problems with Pick-up and Delivery Time Windows and Shared Fleet Capacity Planning. Symmetry **16**(4) (2024) https://doi.org/10.3390/sym16040505

[96] Kou, A.Q., Cheng, Y., Guo, L., Xia, X.Q.: A Data-Driven Newsvendor Model with Unknown Demand and Supply Distribution. JOURNAL OF INDUSTRIAL AND MANAGEMENT OPTIMIZATION (2024) https://doi.org/10.3934/jimo.2024080

[97] Keith, K., Castillo-Villar, K.K., Alaeddini, A.: Machine Learning-Based Problem Space Reduction in Stochastic Programming Models: An Application in Biofuel Supply Chain Network Design. IEEE Access **12**, 33852–33866 (2024) https://doi.org/10.1109/ACCESS.2024.3372516

[98] Khalili-Damghani, K., Dideh-Khani, H., Sadi-Nezhad, S.: A two-stage approach based on anfis and fuzzy goal programming for supplier selection. International Journal of Applied Decision Sciences **6**(1), 1–14 (2013) https://doi.org/10.1504/IJADS.2013.052629

[99] Khalili-Fard, A., Hashemi, M., Bakhshi, A., Yazdani, M., Jolai, F., Aghsami, A.: Integrated relief pre-positioning and procurement planning considering non-governmental organizations support and perishable relief items in a humanitarian supply chain network. Omega (United Kingdom) **127** (2024) https://doi.org/10.1016/j.omega.2024.103111

[100] Kong, F., Jiang, B., Wang, J., Wang, H., Song, H.: Collaborative Delivery Optimization with Multiple Drones via Constrained Hybrid Pointer Network. IEEE Internet of Things Journal **11**(5), 7739–7755 (2024) https://doi.org/10.1109/JIOT.2023.3318524

[101] Kochakkashani, F., Kayvanfar, V., Baldacci, R.: Innovative Applications of Unsupervised Learning in Uncertainty-Aware Pharmaceutical Supply Chain Planning. IEEE Access **12**, 107984–107999 (2024) https://doi.org/10.1109/ACCESS.2024.3435439

[102] Kayhomayoon, Z., Milan, S.G., Azar, N.A., Bettinger, P., Babaian, F., Jaafari, A.: A Simulation-Optimization Modeling Approach for Conjunctive Water Use Management in a Semi-Arid Region of Iran. Sustainability (Switzerland) **14**(5) (2022) https://doi.org/10.3390/su14052691

[103] Kosasih, E.E., Papadakis, E., Baryannis, G., Brintrup, A.: A review of explainable artificial intelligence in supply chain management: A systematic review of neurosymbolic approaches. International Journal of Production Research **62**(4), 1510–1540 (2024) https://doi.org/10.1080/00207543.2023.2281663

[104] Kim, H., Park, J., Kwon, C.: A Neural Separation Algorithm for the Rounded Capacity Inequalities. INFORMS Journal on Computing **36**(4), 987–1005 (2024) https://doi.org/10.1287/ijoc.2022.0310

[105] Khedr, A.M., Sheeja Rani, S.: Enhancing supply chain management with deep learning and machine learning techniques: A review. Journal of Open Innovation:

Technology, Market, and Complexity **10**(4), 100379 (2024) https://doi.org/10.1016/j.joitmc.2024.100379

[106] Kalatzantonakis, P., Sifaleras, A., Samaras, N.: A reinforcement learning-Variable neighborhood search method for the capacitated Vehicle Routing Problem. Expert Systems with Applications **213** (2023) https://doi.org/10.1016/j.eswa.2022.118812

[107] Kalantari, S., Safarzadeh Ramhormozi, R., Wang, Y., Sun, S., Wang, X.: Trailer allocation and truck routing using bipartite graph assignment and deep reinforcement learning. Transactions in GIS **27**(4), 996–1020 (2023) https://doi.org/10.1111/tgis.13057

[108] Lee, D.H., Ahn, J.: Multi-start team orienteering problem for UAS mission re-planning with data-efficient deep reinforcement learning. Applied Intelligence **54**(6), 4467–4489 (2024) https://doi.org/10.1007/s10489-024-05367-4

[109] Li, N., Arnold, D.M., Down, D.G., Barty, R., Blake, J., Chiang, F., Courtney, T., Waito, M., Trifunov, R., Heddle, N.M.: From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization. Transfusion **62**(1), 87–99 (2022) https://doi.org/10.1111/trf.16739

[110] Lee, K.H., Abdollahian, M., Schreider, S., Taheri, S.: Supply Chain Demand Forecasting and Price Optimisation Models with Substitution Effect. Mathematics **11**(11) (2023) https://doi.org/10.3390/math11112502

[111] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015) https://doi.org/10.1038/nature14539

[112] Liu, J., Chen, W., Yang, J., Xiong, H., Chen, C.: Iterative Prediction-and-Optimization for E-Logistics Distribution Network Design. INFORMS Journal on Computing **34**(2), 769–789 (2022) https://doi.org/10.1287/ijoc.2021.1107

[113] Li, Y., Chen, K., Zhang, J.: In: Ivanov, D., Dolgui, A., Sokolov, B. (eds.) AI-Powered Supply Chain and Operations Management (SCOM): Capabilities and Challenges, pp. 139–165. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-85508-5_7

[114] Li, B., Dai, T., Chen, W., Song, X., Zang, Y., Huang, Z., Lin, Q., Cai, K.: T-PORP: A Trusted Parallel Route Planning Model on Dynamic Road Networks. IEEE Transactions on Intelligent Transportation Systems **24**(1), 1238–1250 (2023) https://doi.org/10.1109/TITS.2022.3216310

[115] Li, J., Dai, B.T., Niu, Y., Xiao, J., Wu, Y.: Multi-Type Attention for Solving Multi-Depot Vehicle Routing Problems. IEEE Transactions on Intelligent Transportation Systems, 1–10 (2024) https://doi.org/10.1109/TITS.2024.3413077

[116] Li, M., Hao, J.K., Wu, Q.: A flow based formulation and a reinforcement learning based strategic oscillation for cross-dock door assignment. European Journal of Operational Research **312**(2), 473–492 (2024) https://doi.org/10.1016/j.ejor.2023.07.014

[117] Laldin Ismaeil, M.K., Lalla, A.F.: The Role and Impact of Artificial Intelligence on Supply Chain Management: Efficiency, Challenges, and Strategic Implementation. Journal of Ecohumanism **3**(4), 89–106 (2024) https://doi.org/10.62754/joe.v3i4.3461

[118] Liu, C., Lyu, J., Fang, K.: Integrated packing and routing: A model and its solutions. Computers and Operations Research **172** (2024) https://doi.org/10.1016/j.cor.2024.106790

[119] Liu, Z., Li, X., Khojandi, A.: The flying sidekick traveling salesman problem with stochastic travel time: A reinforcement learning approach. Transportation Research Part E: Logistics and Transportation Review **164** (2022) https://doi.org/10.1016/j.tre.2022.102816

[120] Li, J., Ma, Y., Cao, Z., Wu, Y., Song, W., Zhang, J., Chee, Y.M.: Learning Feature Embedding Refiner for Solving Vehicle Routing Problems. IEEE Transactions on Neural Networks and Learning Systems, 1–13 (2023) https://doi.org/10.1109/TNNLS.2023.3285077

[121] Li, J., Ma, Y., Gao, R., Cao, Z., Lim, A., Song, W., Zhang, J.: Deep Reinforcement Learning for Solving the Heterogeneous Capacitated Vehicle Routing Problem. IEEE Transactions on Cybernetics **52**(12), 13572–13585 (2022) https://doi.org/10.1109/TCYB.2021.3111082

[122] Lopes Silva, M.A., Souza, S.R., Freitas Souza, M.J., Bazzan, A.L.C.: A reinforcement learning-based multi-agent framework applied for solving routing and scheduling problems. Expert Systems with Applications **131**, 148–171 (2019) https://doi.org/10.1016/j.eswa.2019.04.056

[123] Liu, S., Wang, J., Wang, R., Zhang, Y., Song, Y., Xing, L.: Data-driven dynamic pricing and inventory management of an omni-channel retailer in an uncertain demand environment. Expert Systems with Applications **244** (2024) https://doi.org/10.1016/j.eswa.2023.122948

[124] Li, Q., Wang, Y., Xiong, Y., Zhang, S., Zhou, Y.: Machine learning-based optimisation in a two-echelon logistics network for the dry port operation in China. International Journal of Systems Science: Operations and Logistics **10**(1) (2023) https://doi.org/10.1080/23302674.2023.2252321

[125] Li, Y., Xu, M., Wang, H., Zhang, X., Chen, Y., Liu, Y., Liang, Y., Jiang, Y.: Modeling of Cross-Network Collaboration Supply Chain and Solution via Reinforcement Learning-Based Evolutionary Algorithm. International Journal

of Industrial Engineering : Theory Applications and Practice **31**(2), 204–233 (2024) https://doi.org/10.23055/ijietap.2024.31.2.9523

[126] Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J.: Self-supervised learning: Generative or contrastive. IEEE transactions on knowledge and data engineering **35**(1), 857–876 (2021) https://doi.org/10.1109/TKDE.2021.3090866

[127] Liu, H., Zong, Z., Li, Y., Jin, D.: NeuroCrossover: An intelligent genetic locus selection scheme for genetic algorithm using reinforcement learning. Applied Soft Computing **146** (2023) https://doi.org/10.1016/j.asoc.2023.110680

[128] Liu, F., Zhang, Q., Zhu, Q., Tong, X., Yuan, M.: Machine Learning Assisted Multiobjective Evolutionary Algorithm for Routing and Packing. IEEE Transactions on Evolutionary Computation, 1–1 (2024) https://doi.org/10.1109/TEVC.2024.3357819

[129] Mahmud, S., Abbasi, A., Chakrabortty, R.K., Ryan, M.J.: A self-adaptive hyper-heuristic based multi-objective optimisation approach for integrated supply chain scheduling problems. Knowledge-Based Systems **251** (2022) https://doi.org/10.1016/j.knosys.2022.109190

[130] Mokhtarinejad, M., Ahmadi, A., Karimi, B., Rahmati, S.H.A.: A novel learning based approach for a new integrated location-routing and scheduling problem within cross-docking considering direct shipment. Applied Soft Computing Journal **34**, 274–285 (2015) https://doi.org/10.1016/j.asoc.2015.04.062

[131] Mohanty, N., Behera, B.K., Ferrie, C.: Analysis of the Vehicle Routing Problem Solved via Hybrid Quantum Algorithms in the Presence of Noisy Channels. IEEE Transactions on Quantum Engineering **4** (2023) https://doi.org/10.1109/TQE.2023.3303989

[132] Mohanty, N., Behera, B.K., Ferrie, C.: Solving the vehicle routing problem via quantum support vector machines. Quantum Machine Intelligence **6**(1) (2024) https://doi.org/10.1007/s42484-024-00161-4

[133] Momenitabar, M., Dehdari Ebrahimi, Z., Abdollahi, A., Helmi, W., Bengtson, K., Ghasemi, P.: An integrated machine learning and quantitative optimization method for designing sustainable bioethanol supply chain networks. Decision Analytics Journal **7** (2023) https://doi.org/10.1016/j.dajour.2023.100236

[134] Momenitabar, M., Dehdari Ebrahimi, Z., Ghasemi, P.: Designing a sustainable bioethanol supply chain network: A combination of machine learning and meta-heuristic algorithms. Industrial Crops and Products **189** (2022) https://doi.org/10.1016/j.indcrop.2022.115848

[135] Morabit, M., Desaulniers, G., Lodi, A.: Machine-learning-based column selection

for column generation. Transportation Science **55**(4), 815–831 (2021) https://doi.org/10.1287/trsc.2021.1045

[136] Morabit, M., Desaulniers, G., Lodi, A.: Learning to repeatedly solve routing problems. Networks **83**(3), 503–526 (2024) https://doi.org/10.1002/net.22200

[137] Marandi, F., Fatemi Ghomi, S.M.T.: Network configuration multi-factory scheduling with batch delivery: A learning-oriented simulated annealing approach. Computers and Industrial Engineering **132**, 293–310 (2019) https://doi.org/10.1016/j.cie.2019.04.032

[138] Ma, L., Hao, X., Zhou, W., He, Q., Zhang, R., Chen, L.: A hybrid neural combinatorial optimization framework assisted by automated algorithm design. Complex and Intelligent Systems (2024) https://doi.org/10.1007/s40747-024-01600-2

[139] Mohamed-Iliasse, M., Loubna, B., Abdelaziz, B.: Machine Learning in Supply Chain Management: A Systematic Literature Review. International Journal of Supply and Operations Management **9**(4), 398–416 (2022) https://doi.org/10.22034/ijsom.2021.109189.2279

[140] Min, H.: Artificial intelligence in supply chain management: theory and applications. International Journal of Logistics: Research and Applications **13**(1), 13–39 (2010) https://doi.org/10.1080/13675560902736537

[141] Maheshwari, S., Jain, P.K., Kotecha, K.: Route Optimization of Mobile Medical Unit with Reinforcement Learning. Sustainability (Switzerland) **15**(5) (2023) https://doi.org/10.3390/su15053937

[142] Mamoudan, M.M., Jafari, A., Mohammadnazari, Z., Nasiri, M.M., Yazdani, M.: Hybrid machine learning-metaheuristic model for sustainable agri-food production and supply chain planning under water scarcity. Resources, Environment and Sustainability **14** (2023) https://doi.org/10.1016/j.resenv.2023.100133

[143] Meidute-Kavaliauskiene, I., Taşkın, K., Ghorbani, S., Činčikaitė, R., Kačenauskaitė, R.: Reviewing the Applications of Neural Networks in Supply Chain: Exploring Research Propositions for Future Directions. Information (Switzerland) **13**(5) (2022) https://doi.org/10.3390/info13050261

[144] Moradi, B.: The new optimization algorithm for the vehicle routing problem with time windows using multi-objective discrete learnable evolution model. Soft Computing **24**(9), 6741–6769 (2020) https://doi.org/10.1007/s00500-019-04312-9

[145] Meisheri, H., Sultana, N.N., Baranwal, M., Baniwal, V., Nath, S., Verma, S., Ravindran, B., Khadilkar, H.: Scalable multi-product inventory control with lead

time constraints using reinforcement learning. Neural Computing and Applications **34**(3), 1735–1757 (2022) https://doi.org/10.1007/s00521-021-06129-w

[146] Moghaddas, M., Tohidi, G.: A neurodynamic scheme to bi-level revenue-based centralized resource allocation models. Journal of Intelligent and Fuzzy Systems **37**(2), 2741–2756 (2019) https://doi.org/10.3233/JIFS-182953

[147] Mousa, M., Berg, D., Kotecha, N., Rio Chanona, E.A., Mowbray, M.: An analysis of multi-agent reinforcement learning for decentralized inventory control systems. Computers and Chemical Engineering **188** (2024) https://doi.org/10.1016/j.compchemeng.2024.108783

[148] Ma, N., Wang, Z., Ba, Z., Li, X., Yang, N., Yang, X., Zhang, H.: Hierarchical Reinforcement Learning for Crude Oil Supply Chain Scheduling. Algorithms **16**(7) (2023) https://doi.org/10.3390/a16070354

[149] Najjartabar Bisheh, M., Nasiri, G.R., Esmaeili, E., Davoudpour, H., Chang, S.I.: A new supply chain distribution network design for two classes of customers using transfer recurrent neural network. International Journal of System Assurance Engineering and Management **13**(5), 2604–2618 (2022) https://doi.org/10.1007/s13198-022-01670-w

[150] Novelan, M.S., Efendi, S., Sihombing, P., Mawengkang, H.: Vehicle Routing Problem Optimization with Machine Learning in Imbalanced Classification Vehicle Route Data. Eastern-European Journal of Enterprise Technologies **5**(3(125)), 49–56 (2023) https://doi.org/10.15587/1729-4061.2023.288280

[151] Nezamoddini, N., Gholami, A., Aqlan, F.: A risk-based optimization framework for integrated supply chains using genetic algorithm and artificial neural networks. International Journal of Production Economics **225** (2020) https://doi.org/10.1016/j.ijpe.2019.107569

[152] Nouicer, D., Msadaa, I.C., Grayaa, K.: A Novel Routing Solution for EV Fleets: A Real-World Case Study Leveraging Double DQNs and Graph-Structured Data to Solve the EVRPTW Problem. IEEE Access **11**, 122116–122127 (2023) https://doi.org/10.1109/ACCESS.2023.3327324

[153] Noh, J.: Reinforcement Learning for Optimizing Can-Order Policy with the Rolling Horizon Method. Systems **11**(7) (2023) https://doi.org/10.3390/systems11070350

[154] Neria, G., Tzur, M.: The Dynamic Pickup and Allocation with Fairness Problem. Transportation Science **58**(4), 821–840 (2024) https://doi.org/10.1287/trsc.2023.0228

[155] Ochoa-Barragán, R., Munguía-López, A.D.C., Ponce-Ortega, J.M.: A hybrid

machine learning-mathematical programming optimization approach for municipal solid waste management during the pandemic. Environment, Development and Sustainability **26**(7), 17653–17672 (2024) https://doi.org/10.1007/s10668-023-03354-2

[156] Ochoa-Barragán, R., Serrano-Arévalo, T.I., Pulido-Ocegueda, J.C., Cerda-Flores, S.C., Ramírez-Márquez, C., Nápoles-Rivera, F., Ponce-Ortega, J.M.: Sustainable lime production in Michoacan Mexico: An optimal and equitable approach with machine learning. Journal of Cleaner Production **442** (2024) https://doi.org/10.1016/j.jclepro.2024.141017

[157] Oguntola, I.O., Ülkü, M.A., Saif, A., Engau, A.: On the value of shipment consolidation and machine learning techniques for the optimal design of a multimodal logistics network. INFOR **62**(1), 1–52 (2024) https://doi.org/10.1080/03155986.2023.2202079

[158] Pettit, T.J., Fiksel, J., Croxton, K.L.: Ensuring supply chain resilience: A conceptual framework. Journal of Business Logistics (2010)

[159] Phiboonbanakit, T., Horanont, T., Huynh, V.N., Supnithi, T.: A Hybrid Reinforcement Learning-Based Model for the Vehicle Routing Problem in Transportation Logistics. IEEE Access **9**, 163325–163347 (2021) https://doi.org/10.1109/ACCESS.2021.3131799

[160] Perez, H.D., Hubbs, C.D., Li, C., Grossmann, I.E.: Algorithmic approaches to inventory management optimization. Processes **9**(1), 1–17 (2021) https://doi.org/10.3390/pr9010102

[161] Pan, W., Liu, S.Q.: Deep reinforcement learning for the dynamic and uncertain vehicle routing problem. Applied Intelligence **53**(1), 405–422 (2023) https://doi.org/10.1007/s10489-022-03456-w

[162] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., *et al.*: The prisma 2020 statement: an updated guideline for reporting systematic reviews. BMJ **372**, 71 (2021) https://doi.org/10.1136/bmj.n71

[163] Pirayesh Neghab, D., Khayyati, S., Karaesmen, F.: An integrated data-driven method using deep learning for a newsvendor problem with unobservable features. European Journal of Operational Research **302**(2), 482–496 (2022) https://doi.org/10.1016/j.ejor.2021.12.047

[164] Punia, S., Singh, S.P., Madaan, J.K.: From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. Decision Support Systems **136** (2020) https://doi.org/10.1016/j.dss.2020.113340

[165] Qin, W., Zhuang, Z., Huang, Z., Huang, H.: A novel reinforcement learning-based hyper-heuristic for heterogeneous vehicle routing problem. Computers and Industrial Engineering **156** (2021) https://doi.org/10.1016/j.cie.2021.107252

[166] RamachandranPillai, R., Arock, M.: An Adaptive Spiking Neural P System for Solving Vehicle Routing Problems. Arabian Journal for Science and Engineering **45**(4), 2513–2529 (2020) https://doi.org/10.1007/s13369-019-04153-6

[167] Rolf, B., Beier, A., Jackson, I., Müller, M., Reggelin, T., Stuckenschmidt, H., Lang, S.: A review on unsupervised learning algorithms and applications in supply chain management. International Journal of Production Research (2024) https://doi.org/10.1080/00207543.2024.2390968

[168] Rizqi, Z.U., Chou, S.Y.: Neuroevolution reinforcement learning for multi-echelon inventory optimization with delivery options and uncertain discount. Engineering Applications of Artificial Intelligence **134** (2024) https://doi.org/10.1016/j.engappai.2024.108670

[169] Rodríguez-Esparza, E., Masegosa, A.D., Oliva, D., Onieva, E.: A new Hyper-heuristic based on Adaptive Simulated Annealing and Reinforcement Learning for the Capacitated Electric Vehicle Routing Problem. Expert Systems with Applications **252** (2024) https://doi.org/10.1016/j.eswa.2024.124197

[170] Ren, L., Fan, X., Cui, J., Shen, Z., Lv, Y., Xiong, G.: A Multi-Agent Reinforcement Learning Method With Route Recorders for Vehicle Routing in Supply Chain Management. IEEE Transactions on Intelligent Transportation Systems **23**(9), 16410–16420 (2022) https://doi.org/10.1109/TITS.2022.3150151

[171] Robinson La Rocca, C., Cordeau, J.F., Frejinger, E.: One-Shot Learning for MIPs with SOS1 Constraints. Operations Research Forum **5**(3) (2024) https://doi.org/10.1007/s43069-024-00336-6

[172] Ramakurthi, V.B., Manupati, V.K., Machado, J., Varela, L.: A hybrid multi-objective evolutionary algorithm-based semantic foundation for sustainable distributed manufacturing systems. Applied Sciences (Switzerland) **11**(14) (2021) https://doi.org/10.3390/app11146314

[173] Raza, S.M., Sajid, M., Singh, J.: Vehicle routing problem using reinforcement learning: Recent advancements. In: Advanced Machine Intelligence and Signal Processing, pp. 269–280. Springer, New York (2022). https://doi.org/10.1007/978-981-19-0840-8_20

[174] Ramírez-Villamil, A., Montoya-Torres, J.R., Jaegler, A., Cuevas-Torres, J.M.: Reconfiguration of last-mile supply chain for parcel delivery using machine learning and routing optimization. Computers and Industrial Engineering **184** (2023) https://doi.org/10.1016/j.cie.2023.109604

[175] Sharifnia, S.M.E., Amrollahi Biyouki, S., Sawhney, R., Hwangbo, H.: Robust simulation optimization for supply chain problem under uncertainty via neural network metamodeling. Computers and Industrial Engineering **162** (2021) https://doi.org/10.1016/j.cie.2021.107693

[176] Sarker, I.H.: Machine learning: Algorithms, real-world applications and research directions. SN computer science **2**(3), 160 (2021) https://doi.org/10.1007/s42979-021-00592-x

[177] Sutton, R.S., Barto, A.G., *et al.*: Reinforcement Learning: An Introduction vol. 1. MIT Press, Cambridge (1998)

[178] Schmidt, M.: Sankey diagrams: Visual analysis for resource flows. Journal of Industrial Ecology **12**(1), 82–94 (2008) https://doi.org/10.1111/j.1530-9290.2008.00004.x

[179] Shiri, M., Fattahi, P., Sogandi, F.: Two-stage approach for COVID-19 vaccine supply chain network under uncertainty using the machine learning algorithms: A case study. Engineering Applications of Artificial Intelligence **135** (2024) https://doi.org/10.1016/j.engappai.2024.108837

[180] Sangiorgio, M., Guariso, G.: NN-based implicit stochastic optimization of multi-reservoir systems management. Water (Switzerland) **10**(3) (2018) https://doi.org/10.3390/w10030303

[181] Sui, Z., Gosavi, A., Lin, L.: A reinforcement learning approach for inventory replenishment in vendor-managed inventory systems with consignment inventory. EMJ - Engineering Management Journal **22**(4), 44–53 (2010) https://doi.org/10.1080/10429247.2010.11431878

[182] Sheffi, Y., Jr., J.B.R.: A supply chain view of the resilient enterprise. MIT Sloan Management Review (2005)

[183] Shang, Y., Li, S.: Hybrid combinatorial remanufacturing strategy for medical equipment in the pandemic. Computers and Industrial Engineering **174** (2022) https://doi.org/10.1016/j.cie.2022.108811

[184] Seyedan, M., Mafakheri, F.: Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. Journal of Big Data **7**(1), 53 (2020) https://doi.org/10.1186/s40537-020-00329-2

[185] Soroka, A.G., Mikhelson, G.V., Mescheryakov, A.V., Gerasimov, S.V.: Smart Routes: A System for Development and Comparison of Algorithms for Solving Vehicle Routing Problems with Realistic Constraints. Automation and Remote Control **85**(3), 309–319 (2024) https://doi.org/10.1134/S0005117924030081

[186] Shahbazian, R., Pugliese, L.D.P., Guerriero, F., Macrina, G.: Integrating

Machine Learning Into Vehicle Routing Problem: Methods and Applications. IEEE Access **12**, 93087–93115 (2024) https://doi.org/10.1109/ACCESS.2024.3422479

[187] Soeanu, A., Ray, S., Debbabi, M., Berger, J., Boukhtouta, A.: A learning based evolutionary algorithm for distributed multi-depot VRP. Frontiers in Artificial Intelligence and Applications **243**, 49–58 (2012) https://doi.org/10.3233/978-1-61499-105-2-49

[188] Stranieri, F., Stella, F., Kouki, C.: Performance of deep reinforcement learning algorithms in two-echelon inventory control systems. International Journal of Production Research **62**(17), 6211–6226 (2024) https://doi.org/10.1080/00207543.2024.2311180

[189] Tang, C.S.: Perspectives in supply chain risk management. International Journal of Production Economics (2006)

[190] Tan, Y., Gu, L., Xu, S., Li, M.: Supply Chain Inventory Management from the Perspective of "Cloud Supply Chain"—A Data Driven Approach. Mathematics **12**(4) (2024) https://doi.org/10.3390/math12040573

[191] Tarapata, Z., Kulas, W., Antkiewicz, R.: Machine Learning Algorithms for the Problem of Optimizing the Distribution of Parcels in Time-Dependent Networks: The Case Study. Archives of Transport **61**(1), 133–147 (2022) https://doi.org/10.5604/01.3001.0015.8269

[192] Tunga, H., Kar, S., Giri, D.: Multi-objective brainstorm optimization algorithm based on Bayesian inference learning automaton for solving CVRP with time windows. Applied Soft Computing **162** (2024) https://doi.org/10.1016/j.asoc.2024.111755

[193] Triantafyllou, N., Papathanasiou, M.M.: Deep learning enhanced mixed integer optimization: Learning to reduce model dimensionality. Computers and Chemical Engineering **187** (2024) https://doi.org/10.1016/j.compchemeng.2024.108725

[194] Teerasoponpong, S., Sopadang, A.: Decision support system for adaptive sourcing and inventory management in small- and medium-sized enterprises. Robotics and Computer-Integrated Manufacturing **73** (2022) https://doi.org/10.1016/j.rcim.2021.102226

[195] Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., Fischl, M.: Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research **122**, 502–517 (2021) https://doi.org/10.1016/j.jbusres.2020.09.009

[196] Tsolaki, K., Vafeiadis, T., Nizamis, A., Ioannidis, D., Tzovaras, D.: Utilizing

machine learning on freight transportation and logistics applications: A review. ICT Express **9**(3), 284–295 (2023) https://doi.org/10.1016/j.icte.2022.02.001

[197] Vanvuchelen, N., De Boeck, K., Boute, R.N.: Cluster-based lateral transshipments for the Zambian health supply chain. European Journal of Operational Research **313**(1), 373–386 (2024) https://doi.org/10.1016/j.ejor.2023.08.005

[198] Van Dyk, D.A., Meng, X.-L.: The art of data augmentation. Journal of Computational and Graphical Statistics **10**(1), 1–50 (2001) https://doi.org/10.1198/10618600152418584

[199] Hezewijk, L., Dellaert, N., Van Woensel, T., Gademann, N.: Using the proximal policy optimisation algorithm for solving the stochastic capacitated lot sizing problem. International Journal of Production Research **61**(6), 1955–1978 (2023) https://doi.org/10.1080/00207543.2022.2056540

[200] Vlachos, I., Reddy, P.G.: Machine learning in supply chain management: systematic literature review and future research agenda. International Journal of Production Research, 1–30 (2025) https://doi.org/10.1080/00207543.2025.2466062

[201] Wang, Q.: VARL: a variational autoencoder-based reinforcement learning Framework for vehicle routing problems. Applied Intelligence **52**(8), 8910–8923 (2022) https://doi.org/10.1007/s10489-021-02920-3

[202] Wu, C., Barnes, D.: A dynamic feedback model for partner selection in agile supply chains. International Journal of Operations and Production Management **32**(1), 79–103 (2012) https://doi.org/10.1108/01443571211195745

[203] Wang, Z., Bai, R., Khan, F., Özcan, E., Zhang, T.: Gase: graph attention sampling with edges fusion for solving vehicle routing problems. Memetic Computing **16**(3), 337–353 (2024) https://doi.org/10.1007/s12293-024-00428-0

[204] Wu, G., Carvalho Servia, M.Á., Mowbray, M.: Distributional reinforcement learning for inventory management in multi-echelon supply chains. Digital Chemical Engineering **6** (2023) https://doi.org/10.1016/j.dche.2022.100073

[205] Wang, Q., Hao, Y.: Routing optimization with Monte Carlo Tree Search-based multi-agent reinforcement learning. Applied Intelligence **53**(21), 25881–25896 (2023) https://doi.org/10.1007/s10489-023-04881-1

[206] Wang, Q., Hao, Y., Zhang, J.: Generative inverse reinforcement learning for learning 2-opt heuristics without extrinsic rewards in routing problems. Journal of King Saud University - Computer and Information Sciences **35**(9) (2023) https://doi.org/10.1016/j.jksuci.2023.101787

[207] Wang, T., Jeff Hong, L.: Large-Scale Inventory Optimization: A Recurrent Neural Networks–Inspired Simulation Approach. INFORMS Journal on Computing **35**(1), 196–215 (2023) https://doi.org/10.1287/IJOC.2022.1253

[208] Wang, F., Lin, L.: Spare parts supply chain network modeling based on a novel scale-free network and replenishment path optimization with Q learning. Computers and Industrial Engineering **157** (2021) https://doi.org/10.1016/j.cie.2021.107312

[209] Wang, X., Li, X., Leung, V.C.: Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges. IEEE Access **3**, 1379–1391 (2015) https://doi.org/10.1109/ACCESS.2015.2467174

[210] Wang, K., Long, C., Ong, D.J., Zhang, J., Yuan, X.M.: Single-Site Perishable Inventory Management Under Uncertainties: A Deep Reinforcement Learning Approach. IEEE Transactions on Knowledge and Data Engineering **35**(10), 10807–10813 (2023) https://doi.org/10.1109/TKDE.2023.3250249

[211] Wang, Q., Lai, K.H., Tang, C.: Solving combinatorial optimization problems over graphs with BERT-Based Deep Reinforcement Learning. Information Sciences **619**, 930–946 (2023) https://doi.org/10.1016/j.ins.2022.11.073

[212] Wu, Y., Song, W., Cao, Z., Zhang, J., Lim, A.: Learning Improvement Heuristics for Solving Routing Problems. IEEE Transactions on Neural Networks and Learning Systems **33**(9), 5057–5069 (2022) https://doi.org/10.1109/TNNLS.2021.3068828

[213] Wang, L., Song, C., Sun, Y., Lu, C., Chen, Q.: A Neural Multi-Objective Capacitated Vehicle Routing Optimization Algorithm Based on Preference Adjustment. Electronics (Switzerland) **12**(19) (2023) https://doi.org/10.3390/electronics12194167

[214] Wang, Q., Tang, C.: Deep reinforcement learning for transportation network combinatorial optimization: A survey. Knowledge-Based Systems **233** (2021) https://doi.org/10.1016/j.knosys.2021.107526

[215] Wang, F., Wang, X., Sun, S.: A reinforcement learning level-based particle swarm optimization algorithm for large-scale optimization. Information Sciences **602**, 298–312 (2022) https://doi.org/10.1016/j.ins.2022.04.053

[216] Wang, C., Yu, Z., McAleer, S., Yu, T., Yang, Y.: ASP: Learn a Universal Neural Solver! IEEE Transactions on Pattern Analysis and Machine Intelligence **46**(6), 4102–4114 (2024) https://doi.org/10.1109/TPAMI.2024.3352096

[217] Wang, Q., Zhang, C., Tang, C.: Discovering Lin-Kernighan-Helsgaun heuristic for routing optimization using self-supervised reinforcement learning. Journal of King Saud University - Computer and Information Sciences **35**(8) (2023)

https://doi.org/10.1016/j.jksuci.2023.101723

[218] Wang, J., Zheng, R., Wang, Z.: Supply Chain Optimization Strategy Research Based on Deep Learning Algorithm. Mobile Information Systems **2022** (2022) https://doi.org/10.1155/2022/9058490

[219] Xu, Y., Fang, M., Chen, L., Xu, G., Du, Y., Zhang, C.: Reinforcement Learning With Multiple Relational Attention for Solving Vehicle Routing Problems. IEEE Transactions on Cybernetics **52**(10), 11107–11120 (2022) https://doi.org/10.1109/TCYB.2021.3089179

[220] Xue, S.: An adaptive ant colony algorithm for crowdsourcing multi-depot vehicle routing problem with time windows. Sustainable Operations and Computers **4**, 62–75 (2023) https://doi.org/10.1016/j.susoc.2023.02.002

[221] Yin, X., Büyüktahtakın, I.E.: A multi-stage stochastic programming approach to epidemic resource allocation with equity considerations. Health Care Management Science **24**(3), 597–622 (2021) https://doi.org/10.1007/s10729-021-09559-z

[222] Yilmaz, D., Büyüktahtakın, İ.E.: Learning optimal solutions via an lstm-optimization framework. Operations Research Forum **4**(2), 48 (2023) https://doi.org/10.1007/s43069-023-00224-5

[223] Yilmaz, D., Büyüktahtakın, İ.E.: A non-anticipative learning-optimization framework for solving multi-stage stochastic programs. Annals of Operations Research, 1–41 (2024) https://doi.org/10.1007/s10479-024-06100-7

[224] Yilmaz, D., Büyüktahtakın, İ.E.: A deep reinforcement learning framework for solving two-stage stochastic programs. Optimization Letters **18**(9), 1993–2020 (2024) https://doi.org/10.1007/s11590-023-02009-5

[225] Yilmaz, D., Büyüktahtakın, İ.E.: An expandable machine learning-optimization framework to sequential decision-making. European Journal of Operational Research **314**(1), 280–296 (2024) https://doi.org/10.1016/j.ejor.2023.10.045

[226] Yin, X., Büyüktahtakın, İ.E., Patel, B.P.: COVID-19: Data-driven optimal allocation of ventilator supply under uncertainty and risk. European journal of operational research **304**(1), 255–275 (2023) https://doi.org/10.1016/j.ejor.2021.11.052

[227] Yin, X., Bushaj, S., Yuan, Y., Büyüktahtakın, İ.E.: COVID-19: Agent-based simulation-optimization to vaccine center location vaccine allocation problem. IISE Transactions **56**(7), 699–714 (2024)

[228] Yang, J., Lau, H.C., Wang, H.: Optimization of customer service and driver dispatch areas for on-demand food delivery. Transportation Research Part C:

Emerging Technologies **165** (2024) https://doi.org/10.1016/j.trc.2024.104653

[229] Yue, B., Ma, J., Shi, J., Yang, J.: A Deep Reinforcement Learning-Based Adaptive Search for Solving Time-Dependent Green Vehicle Routing Problem. IEEE Access **12**, 33400–33419 (2024) https://doi.org/10.1109/ACCESS.2024.3369474

[230] Yunusoglu, P., Ozsoydan, F.B., Bilgen, B.: A machine learning-based two-stage approach for the location of undesirable facilities in the biomass-to-bioenergy supply chain. Applied Energy **362** (2024) https://doi.org/10.1016/j.apenergy.2024.122961

[231] Yi, W., Qu, R., Jiao, L., Niu, B.: Automated Design of Metaheuristics Using Reinforcement Learning Within a Novel General Search Framework. IEEE Transactions on Evolutionary Computation **27**(4), 1072–1084 (2023) https://doi.org/10.1109/TEVC.2022.3197298

[232] Yang, B., Xu, X., Gong, Y., Rekik, Y.: Data-driven optimization models for inventory and financing decisions in online retailing platforms. Annals of Operations Research **339**(1-2), 741–764 (2024) https://doi.org/10.1007/s10479-023-05234-4

[233] Yu, J.J.Q., Yu, W., Gu, J.: Online Vehicle Routing with Neural Combinatorial Optimization and Deep Reinforcement Learning. IEEE Transactions on Intelligent Transportation Systems **20**(10), 3806–3817 (2019) https://doi.org/10.1109/TITS.2019.2909109

[234] Yu, L., Zhang, C., Jiang, J., Yang, H., Shang, H.: Reinforcement learning approach for resource allocation in humanitarian logistics. Expert Systems with Applications **173** (2021) https://doi.org/10.1016/j.eswa.2021.114663

[235] Zhang, X., Chen, L., Gendreau, M., Langevin, A.: Learning-Based Branch-and-Price Algorithms for the Vehicle Routing Problem with Time Windows and Two-Dimensional Loading Constraints. INFORMS Journal on Computing **34**(3), 1419–1436 (2022) https://doi.org/10.1287/ijoc.2021.1110

[236] Zheng, Y.J., Chen, X., Song, Q., Yang, J., Wang, L.: Evolutionary Optimization of COVID-19 Vaccine Distribution With Evolutionary Demands. IEEE Transactions on Evolutionary Computation **27**(1), 141–154 (2023) https://doi.org/10.1109/TEVC.2022.3164260

[237] Zou, Y., Hao, J.K., Wu, Q.: A reinforcement learning guided hybrid evolutionary algorithm for the latency location routing problem. Computers and Operations Research **170** (2024) https://doi.org/10.1016/j.cor.2024.106758

[238] Zhou, Y., Li, W., Wang, X., Qiu, Y., Shen, W.: Adaptive gradient descent enabled ant colony optimization for routing problems. Swarm and Evolutionary Computation **70** (2022) https://doi.org/10.1016/j.swevo.2022.101046

[239] Zhao, J., Mao, M., Zhao, X., Zou, J.: A Hybrid of Deep Reinforcement Learning and Local Search for the Vehicle Routing Problems. IEEE Transactions on Intelligent Transportation Systems **22**(11), 7208–7218 (2021) https://doi.org/10.1109/TITS.2020.3003163

[240] Zwaida, T.A., Pham, C., Beauregard, Y.: Optimization of inventory management to prevent drug shortages in the hospital supply chain. Applied Sciences (Switzerland) **11**(6) (2021) https://doi.org/10.3390/app11062726

[241] Zhang, Z., Qi, G., Guan, W.: Coordinated multi-agent hierarchical deep reinforcement learning to solve multi-trip vehicle routing problems with soft time windows. IET Intelligent Transport Systems **17**(10), 2034–2051 (2023) https://doi.org/10.1049/itr2.12394

[242] Zhang, J., Sun, Y., Feng, Q., Zhao, Y., Wang, Z.: Data-Driven Robust Optimization of the Vehicle Routing Problem with Uncertain Customers. Complexity **2022** (2022) https://doi.org/10.1155/2022/9064669

[243] Zavvar Sabegh, M.H., Mohammadi, M., Naderi, B.: Multi-objective optimization considering quality concepts in a green healthcare supply chain for natural disaster response: neural network approaches. International Journal of System Assurance Engineering and Management **8**, 1689–1703 (2017) https://doi.org/10.1007/s13198-017-0645-1

[244] Zhang, J., Tutun, S., Anvaryazdi, S.F., Amini, M., Sundaramoorthi, D., Sundaramoorthi, H.: Management of resource sharing in emergency response using data-driven analytics. Annals of Operations Research **339**(1-2), 663–692 (2024) https://doi.org/10.1007/s10479-023-05702-x

[245] Zhou, J., Wu, Y., Cao, Z., Song, W., Zhang, J., Chen, Z.: Learning Large Neighborhood Search for Vehicle Routing in Airport Ground Handling. IEEE Transactions on Knowledge and Data Engineering **35**(9), 9769–9782 (2023) https://doi.org/10.1109/TKDE.2023.3249799

[246] Zhang, C., Wu, Y., Ma, Y., Song, W., Le, Z., Cao, Z., Zhang, J.: A review on learning to solve combinatorial optimisation problems in manufacturing. IET Collaborative Intelligent Manufacturing **5**(1) (2023) https://doi.org/10.1049/cim2.12072

[247] Zhang, D., Wang, M., Mango, J., Li, X., Xu, X.: A survey on applications of reinforcement learning in spatial resource allocation. Computational Urban Science **4**(1) (2024) https://doi.org/10.1007/s43762-024-00127-z

[248] Zhang, Z., Wu, Z., Zhang, H., Wang, J.: Meta-Learning-Based Deep Reinforcement Learning for Multiobjective Optimization Problems. IEEE Transactions on Neural Networks and Learning Systems **34**(10), 7978–7991 (2023) https://doi.org/10.1109/TNNLS.2022.3148435

[249] Zhang, S., Xu, J., Qiao, Y.: Multi-Objective Q-Learning-Based Brain Storm Optimization for Integrated Distributed Flow Shop and Distribution Scheduling Problems. Mathematics **11**(20) (2023) https://doi.org/10.3390/math11204306

[250] Zhao, S., You, F.: Distributionally robust chance constrained programming with generative adversarial networks (GANs). AIChE Journal **66**(6) (2020) https://doi.org/10.1002/aic.16963

[251] Zhang, J., Yu, M., Feng, Q., Leng, L., Zhao, Y.: Data-Driven Robust Optimization for Solving the Heterogeneous Vehicle Routing Problem with Customer Demand Uncertainty. Complexity **2021** (2021) https://doi.org/10.1155/2021/6634132

[252] Zhang, R., Yu, R., Xia, W.: Constraint-aware Policy Optimization to Solve the Vehicle Routing Problem with Time Windows. Information Technology and Control **51**(1), 126–138 (2022) https://doi.org/10.5755/j01.itc.51.1.29924

# Appendix A Glossary of Acronyms

Table A1: List of Acronyms Used in This Paper.

| Acronym | Full Form |
|---|---|
| **AC** | Actor Critique |
| **AI** | Artificial Intelligence |
| **ANN** | Artificial Neural Network |
| **BIP** | Binary Integer Programming |
| **BPTT** | Backpropagation Through Time |
| **CART** | Classification and Regression Tree |
| **CNN** | Convolutional Neural Network |
| **CVRP** | Capacitated Vehicle Routing Problem |
| **DDPG** | Deep Deterministic Policy Gradient |
| **DP** | Dynamic Programming |
| **DQN** | Deep Q-Network |
| **DRL** | Deep Reinforcement Learning |
| **GAN** | Generative Adversarial Network |
| **GNN** | Graph Neural Network |
| **GRU** | Gated Recurrent Unit |
| **HRL** | Hierarchical Reinforcement Learning |
| **ILP** | Integer Linear Programming |
| **IP** | Integer Programming |
| **IoT** | Internet of Things |
| **LP** | Linear Programming |
| **LSTM** | Long Short-Term Memory |
| **MAE** | Mean Absolute Error |
| **MARL** | Multi-Agent Reinforcement Learning |
| **MDP** | Markov Decision Process |
| **MILP** | Mixed-Integer Linear Programming |
| **MINLP** | Mixed-Integer Non-Linear Programming |
| **MIP** | Mixed-Integer Programming |
| **ML** | Machine Learning |

Continued on next page

Table A1: List of Acronyms Used in This Paper. (Continued)

| Acronym | Full Form |
|---|---|
| **MSE** | Mean Squared Error |
| **NERL** | Neuroevolutionary Methods |
| **NLP** | Non-Linear Programming |
| **PCA** | Principal Component Analysis |
| **PPO** | Proximal Policy Optimization |
| **PRISMA** | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| **RL** | Reinforcement Learning |
| **RMSE** | Root Mean Squared Error |
| **RNN** | Recurrent Neural Network |
| **SCM** | Supply Chain Management |
| **SLR** | Systematic Literature Review |
| **SRQ** | Sub-Research Question |
| **SSRL** | Self-Supervised Reinforcement Learning |
| **SVM** | Support Vector Machine |
| **SC** | Supply Chain |
| **TSP** | Traveling Salesman Problem |
| **VAE** | Variational Autoencoder |
| **VRP** | Vehicle Routing Problem |
| **XAI** | Explainable Artificial Intelligence |

# Appendix B Types of Papers

The literature at the intersection of machine learning and supply chain optimization can be categorized by its primary contribution. This classification helps distinguish between papers that introduce new theoretical frameworks, those that develop novel solution methodologies, and those that synthesize the state of the art. Our analysis reveals that the vast majority of papers present a combination of new models and methods, reflecting a field that is actively developing new problem-solving tools.

The most substantial body of work, with 148 papers classified as Model and 36 as Method, consists of original research contributions that form the creative engine of the field. In practice, these two categories are deeply intertwined, as most papers that propose a new mathematical model for a supply chain problem also develop a novel method to solve it. Model-focused papers are primarily concerned with proposing a new mathematical or conceptual framework that captures the complexities of a specific supply chain problem. For example, many papers formulate novel routing models, such as for electric vehicle fleets [150], drone-assisted delivery [107, 117], and trusted routing in dynamic road networks [112]. Others model entire supply chain networks, such as a two-stage approach to vaccine distribution [181], a hybrid model for sustainable agri-food production [139], and a framework for managing blood supply chains [9]. These papers define the "what", the specific problem being solved. Method-focused papers, on the other hand, concentrate on the "how", the algorithm or solution approach used to solve a new or an existing problem. This includes the development of broad frameworks, such as a hybrid neural combinatorial optimization approach [136], as well as more specific techniques, such as a feature embedding refiner for *VRP*s [118], a neural large neighborhood search [82], a rollout algorithm for multi-depot *VRP*s [7], and a Newton-like neural network for resource allocation [8]. Many articles were explicitly designated as "Model/Methodology," such as for integrated packing and routing [116], for emergency medical resource allocation [34], and as a decision support tool in a bio-ethanol supply chain [10]. This strong overlap indicates that the most common contribution in this research area is the formulation of new problems and the development of customized solution approaches.

A critical component of any academic field, review articles synthesize the existing body of knowledge, and we identified 15 such articles in our study. These articles provide a structured overview of a topic, identify dominant trends, highlight research gaps, and propose agendas for future work. They are invaluable resources for both newcomers seeking to understand the field and established researchers looking for new directions. Our surveyed literature includes several high-quality reviews covering broad topics, such as a systematic review of machine learning in logistics and supply chain management [2] and a review of neural network applications in supply chains [141]. More focused reviews tackle specific problems, such as the integration of machine learning into the *Vehicle Routing Problem* [188], optimization techniques in last-mile logistics [64], and learning-based approaches for combinatorial problems in manufacturing [246]. Other reviews concentrate on specific ML techniques, such as deep reinforcement learning for transportation networks [216], unsupervised learning algorithms in supply chain management [170], and reinforcement learning for spatial resource allocation [247]. The scope of these reviews also includes emerging topics such as the impact of ChatGPT on supply chains [54] and the use of ML for the management of supply chain risk [25].

Table B2: Type of reviewed papers. Entries are sorted by frequency of use.

| Type of Paper | Count | References |
|---|---|---|
| Model | 148 | [1, 4, 9–12, 20, 27, 28, 31–34, 36, 38, 39, 44, 46, 47, 49–53, 57, 58, 60, 62, 66, 69–71, 73–76, 78, 79, 81, 84–88, 90, 92–98, 100, 101, 105, 107–109, 112–114, 116, 117, 119, 121–135, 138, 139, 142–152, 155, 156, 158, 160, 162, 167–169, 171–173, 175, 176, 181–184, 189, 190, 192–194, 196, 199, 202, 206, 208, 211–214, 218–220, 222, 227–234, 236, 239–245, 248–252] |
| Method | 36 | [6–8, 16, 17, 26, 37, 43, 45, 48, 56, 63, 65, 77, 82, 89, 99, 102, 104, 111, 118, 136, 164, 165, 174, 178, 195, 200, 201, 203, 207, 210, 215, 221, 235, 238] |
| Review | 15 | [2, 25, 54, 64, 68, 115, 137, 141, 170, 186, 188, 198, 216, 246, 247] |

# Appendix C A Glossary of Types of Optimization Problems

The definitions and the importance of the types of optimization problems are defined below:

At the foundation of mathematical optimization lies **Linear Programming (LP)**, a technique used to find the best outcome in a model where all relationships are linear. The method was pioneered during World War II, most notably by George Dantzig [40], who developed the simplex algorithm, a highly efficient method for solving LP problems. In an LP model, the goal is to maximize or minimize a linear objective function, such as total profit or cost, subject to a set of linear constraints that represent resource limitations. The core idea, as formalized in foundational texts such as Winston's Operations Research [209], is that the optimal solution will always lie at a corner point of the feasible region defined by the constraints. This geometric property is what allows the simplex algorithm to systematically move from one corner to an adjacent, better one until an optimum is found. LP's power lies in its ability to model a vast array of planning problems and solve them with remarkable computational efficiency, though its primary limitation is the assumption that decision variables can take on any fractional value.

Many real-world supply chain decisions, however, are not fractional; one cannot build 0.7 of a warehouse or dispatch half a truck. This limitation of LP led to the development of **Integer Programming (IP)**, which extends LP to handle discrete choices by requiring some or all decision variables to be integers. The theory and methods for solving these problems, such as branch-and-bound and cutting plane algorithms, were largely developed in the mid-20th century, with foundational work synthesized in texts such as Nemhauser and Wolsey's Integer and Combinatorial Optimization [153]. The most versatile form is **Mixed-Integer Linear Programming (MILP)**, which combines both continuous and integer variables. This allows for modeling complex "yes/no" decisions (e.g., whether to open a facility, represented by a binary 0-1 variable) alongside continuous decisions (e.g., how much to ship from that facility). The inclusion of integer variables makes these problems NP-hard, meaning they are exponentially more difficult to solve than pure LP problems, but their ability to capture the discrete nature of reality has made them an indispensable tool for strategic and tactical decision-making in virtually every industry.

While linear models are powerful, many real-world phenomena exhibit non-linear behavior. **Non-Linear Programming (NLP)** is the framework for optimization problems where the objective function or at least one of the constraints is non-linear. This is essential for modeling complex systems where relationships are not proportional, such as in chemical processes, financial portfolio optimization, where risk is a quadratic function, or supply chains with economies of scale. Unlike LP, where the optimum is always at a corner, in NLP the optimum can be anywhere within the feasible region, and there may be multiple local optima. This makes solving NLPs significantly more challenging, often requiring sophisticated iterative methods such as gradient descent or Newton's method to find a solution, as detailed in seminal texts such as Nocedal and Wright's Numerical Optimization [154]. When combined with integer variables, the problem becomes a **Mixed-Integer Non-Linear Programming (MINLP)** formulation, representing one of the most difficult classes of optimization problems to solve.

**Combinatorial Optimization** is a field of optimization focused on finding an optimal object from a finite, but often astronomically large, set of possible objects. Rather than just deciding "how much," these problems are concerned with finding the best arrangement, sequence, or selection of discrete items. Foundational problems such as the *Traveling Salesman Problem (TSP)* and the *Vehicle Routing Problem (VRP)* epitomize this class. The theoretical underpinnings of this field are deeply rooted in computer science and graph theory, with much of the complexity theory established in the 1970s and captured in classic texts such as Papadimitriou and Steiglitz's Combinatorial Optimization: Algorithms and Complexity [166]. Because the number of possible solutions can grow factorially (e.g., a 20-city *TSP* has more possible tours than atoms in the known universe), exact solutions are often intractable. This has driven the development of both sophisticated exact algorithms (such as branch-and-cut) and powerful heuristics that can find high-quality, near-optimal solutions in a reasonable amount of time.

Traditional optimization models often assume that all parameters are known with certainty, a simplification that can lead to fragile solutions in the real world. To address this, **Stochastic Optimization** and **Robust Optimization** were developed. Stochastic Optimization, pioneered in works such as those of Dantzig and later synthesized by authors such as Birge and Louveaux [18], is applied when uncertainty can be described by a probability distribution. It typically aims to find a solution that optimizes the expected performance over all possible future scenarios. In contrast, Robust Optimization, developed more recently by researchers such as Ben-Tal and Nemirovski [22], offers a more risk-averse approach. It does not require precise probability distributions and instead seeks a solution that remains feasible and performs well under the worst-case realization of uncertainty within a given set. This makes robust solutions "immunized" against uncertainty, providing a guarantee of performance that is critical in high-stakes applications such as energy grid management or critical infrastructure design.

For problems that involve making a sequence of decisions over time, **Dynamic Programming (DP)** provides a powerful solution methodology. Introduced by Richard Bellman in the 1950s [14], DP is based on the "Principle of Optimality," which states that an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. In practice, this means breaking down a complex, multi-stage problem into a series of simpler, nested subproblems. By solving each subproblem optimally, typically by working backward from the end of the time horizon, DP can determine the optimal policy for the entire problem. This fundamental concept of value functions and state transitions laid the groundwork for modern control theory and is the direct theoretical ancestor of reinforcement learning.

# Appendix D A Glossary of Categories of Optimization Problems

We discovered the following categories in the reviewed papers:

***Assignment* Problem:** In an *Assignment* problem, a set of objects must be matched one-to-one with a set of tasks such that the overall cost is minimized (or total profit maximized) under the constraint that each object is assigned to at most one task and vice versa [159]. *Assignment* problems have numerous applications in scheduling and resource allocation (e.g., assigning workers to jobs or machines to tasks in a cost-optimal way).

***Bi-level* Optimization:** *Bi-level* optimization problems feature a hierarchical structure with two decision levels (an upper-level "leader" and a lower-level "follower"), where the leader's decision influences the follower's feasible region and objective [30]. Such models often represent Stackelberg game scenarios; for example, a government (leader) sets policies such as tolls or prices, anticipating the optimal reaction of consumers or network users (followers).

***Bin Packing*:** The *Bin Packing* problem involves packing a collection of items of varying sizes into a minimum number of fixed-capacity bins without exceeding each bin's capacity [29]. This NP-hard problem arises in logistics and computing (for instance, loading trucks or allocating tasks to servers), where the aim is to use resources as efficiently as possible.

***Classical Supply Chain*:** *Classical Supply Chain* problems involve sending flow (of goods, service, etc.) through a network in order to optimize an objective, subject to capacities on the arcs. Classically, there is a flow of goods from the manufacturer to the customer. This flow can be transferred through different types of networks. Water resource allocation, healthcare supply chain, and fuel distribution are some of the examples.

***Collaborative Supply Chain*:** A *Collaborative Supply Chain* is a supply network in which independent entities, such as suppliers, manufacturers, and customers, share information, jointly plan operations, and coordinate decisions to improve overall system performance rather than optimizing individually. This collaboration typically includes synchronized forecasting, shared datasets, integrated production or replenishment planning, and cooperative problem solving. The objective of these problems can be reducing uncertainty, cutting total system cost, improving responsiveness, and enhancing service levels; however, in order to achieve these goals, trust, data-sharing capabilities, and aligned incentives are needed, which can be difficult to establish and maintain [185].

***Facility Location*****:** *Facility Location* problems involve determining the optimal locations for one or more facilities to optimize a given objective (such as minimizing transportation costs or maximizing service coverage) given the distribution of customers or demand points [157]. Real-world applications include locating warehouses in a supply chain to minimize logistics costs and siting emergency services (hospitals, fire stations) to maximize coverage and accessibility.

***Inventory Control*****:** *Inventory control* optimization focuses on deciding when and how much stock to order or produce in order to meet demand while minimizing total costs (holding costs, ordering costs, and shortage costs) [237]. Typical use cases include retail or manufacturing settings, where models (e.g., economic order quantity and $(s, S)$ policies) guide replenishment decisions to balance the trade-off between inventory holding versus stockout risk.

**Knapsack Problem:** The knapsack problem asks for the selection of a subset of items, each with a given value and weight, to maximize the total value without exceeding a weight (or capacity) limit [103]. It is a fundamental NP-hard optimization problem modeling resource allocation under a single constraint, with practical examples such as budget allocation (selecting projects or investments to maximize benefit given a budget constraint).

***Lot-sizing*****:** *Lot-sizing* problems involve determining the optimal production or order quantities over a planning horizon to meet time-phased demand at minimum cost, balancing setup/order costs and holding costs [217]. This is central in production planning and inventory management (e.g., deciding how much of a product to produce or order in each period) to avoid excessive inventory while preventing stockouts.

***News-vendor Problem*****:** The *News-vendor* (newsboy) problem is a single-period stochastic inventory model in which a decision-maker must choose an order quantity for a perishable or short-life product before the uncertain demand realizes, aiming to balance the overstock versus understock cost trade-offs [5]. It provides the optimal stock level that maximizes expected profit or service level and is widely applied to inventory decisions for fashion goods, perishables, or any product with one selling period.

***Price Optimization*****:** *Price optimization* involves determining the optimal prices for products or services to maximize a business objective (such as profit or revenue), given how demand responds to price changes [161]. Techniques in this category are central to revenue management and are used in industries such as retail, airlines, and hotels (e.g., dynamically adjusting prices or discounts based on demand elasticity, seasonality, and inventory levels to achieve the best financial outcome).

**Scheduling:** *Scheduling* problems entail allocating jobs or tasks to time slots and/or resources (machines, employees, etc.) in an optimal way, typically to minimize some objective such as total completion time, lateness, or cost [163]. This broad class includes CPU task scheduling, production scheduling on machines, timetabling, and project scheduling (e.g., the classic job-shop scheduling problem seeks an order of jobs on machines that minimizes the overall makespan).

***Supplier Selection*** The *Supplier Selection* problem involves choosing the best supplier(s) (and often determining order quantities to allocate to each) from a set of candidates, based on multiple criteria such as price, quality, reliability, and capacity [204]. This is a key optimization task in procurement and supply chain management, commonly formulated as a multi-criteria or cost-minimization problem (for example, selecting vendors to minimize total purchasing cost while meeting quality standards and demand requirements).

***Traveling Salesman Problem (TSP)*:** The *TSP* asks for the shortest possible route that visits each city (customer node) exactly once and returns to the starting city, thereby finding an optimal tour through all given locations [41]. It is a famous NP-hard combinatorial optimization problem with numerous applications in routing and logistics (such as planning optimal delivery routes or drill bit paths in manufacturing), and it serves as a benchmark for many optimization methods.

***Vehicle Routing Problem (VRP)*:** The *VRP* generalizes the *TSP* to multiple vehicles and additional constraints. *VRP* involves designing optimal routes for a fleet of vehicles to service a set of customers (often with demands), typically starting and ending at a depot, while respecting vehicle capacity and route length constraints [197]. *VRP* is fundamental in distribution logistics (for example, finding minimum-distance delivery routes for trucks to drop off goods to customers) and has many variants (capacitated *VRP*, *VRP* with time windows, etc.) to model real-world routing challenges.

# Appendix E Problem Categories in Reviewed Studies

The overwhelming majority of the articles focus on the *vehicle routing problem (VRP)*. These problems are notoriously difficult to solve at large scales and that can justify the significant attention towards the *VRP* in the literature. How ML techniques were used to help solve *VRP* varies from unsupervised techniques, such as heuristics [49], to adaptive calculation techniques that do not rely on heuristics [169]; however, the majority of articles use reinforcement learning and neural networks to benefit from their similar structure and powerful computational capacity [82, 104, 245].

The next category of problems is *classical supply chain* problems. This category is a general category that includes a wide range of applications from blood resource allocation [108] to allocation of water resources [43]. We name this category *classical supply chain*, since the structure of the problems followed a network flow optimization, where the flow of goods from suppliers to customers is controlled in order to achieve certain goals. For instance, Abbasi et al. [1] and Li et al. [108] focus on the blood distribution in healthcare settings. The scarcity of water resources has made the water supply chain a well-studied topic among different types of supply chain problems [36, 58, 74–76, 101, 139]. Other supply chains, such as emergency resources [34], automobile manufacturing [47], vaccines [181], agricultural products [56], and biofuel [10, 229] are also included in the category of *classical supply chain* problems.

*Inventory Control* problems come next in frequency. These problems usually deal with balancing stock levels and minimizing costs in uncertain demand environments. Approaches range from classical stochastic programming models to machine learning–based prediction methods [1, 108], with reinforcement learning being especially common when adaptive policies are needed in dynamic environments [151, 183]. Neural networks are also employed for demand forecasting and scalable decision rules [122, 143]. The focus is often on bridging the gap between predictive analytics and prescriptive optimization in *inventory control*.

*Traveling Salesperson Problem (TSP)* papers typically pursue "learning to construct tours," hybridizing neural combinatorial models with local search or using meta-learning to generalize across instance distributions. Reinforcement- and evolution-inspired operators (e.g., neurocrossover) appear frequently, often paired with problem-specific decoders and post-improvement [107, 117, 125, 208, 219, 238, 248].

Job-shop/flow-shop *Scheduling* with ML centers on (deep) RL to map shop states to dispatching rules, or hybrids that couple predictive models with classical metaheuristics. Hierarchical [146] and Q-learning [85] schemes are common, and several studies report competitive performance when learned policies are combined with domain heuristics or constraint-aware search [246].

Papers that focus on *lot-sizing* problems [24] mainly use hybrid predictive-prescriptive pipelines and learning-augmented heuristics for capacitated *lot-sizing* variants under uncertainty. Articles use ML to forecast demands/costs, learn *lot-sizing* decisions, or accelerate decomposition/heuristics; several combine evolutionary search with neural surrogates or controllers [147, 171, 184, 190, 200, 212, 224–226].

*Supplier Selection* problems have typically been framed as a multi-criteria classification/ranking task enhanced by ML predictors (quality, lead time, risk), then integrated the predictions into optimization for sourcing or contracts [4, 98]. Recent work explores end-to-end designs that pair predictive models with cost-to-serve or resilience objectives [87].

*Collaborative Supply Chains* form another category of problems studied in the reviewed articles. Collaboration introduces information sharing, joint planning, and risk-awareness mechanisms [60]. ML contributes to demand/price forecasts and learned coordination rules; several papers study learning-enabled contracts or coordination policies under uncertainty and disruption risk [65, 124, 149, 182, 244].

When it comes to *Facility Location* problems, ML is used to learn demand fields, cost surfaces, or feasibility filters that then feed mixed-integer location models, and in integrated location-routing settings to warm-start or steer search [26]. Recent work also explores deep architectures as surrogates for expensive assignment subproblems [195].

*Assignment* problems appear to be mainly handled by neural approximators and one-shot solvers, as well as distributed formulations where learned subproblem policies improve throughput or latency. The emphasis is on amortized inference and robust generalization to new instance sizes [8, 114, 174].

*Bi-level* studies explore neurodynamic solvers and recurrent schemes for problems where upper-level strategic choices face learned or algorithmic lower-level responses. This line points to the integration of differentiable surrogates or learned lower-level policies to speed the bilevel search [63, 77, 144].

*Bin Packing* papers are fewer, but representative approaches combine learned patterns (e.g., clustering/NN features of items) with bin packing heuristics, or embed RL to decide placement/sequencing; integrated logistics settings sometimes fold packing as a submodule [116, 246].

The work focusing on *News-vendor* problems uses ML to produce calibrated demand (or demand distribution) estimates that plug into single-period cost trade-offs; several papers stress distributional robustness or integrate learned features in the cost parameters. When the loss surface is well captured, ML-augmented *news-vendors* can materially reduce mismatch costs [95, 165, 167].

This niche group uses tree-style learners /boosting-style learners for demand/price response, then optimizes markups under inventory or service constraints; the ML piece primarily improves elasticity estimation and heterogeneity capture [109, 122]. *Knapsack* problems [23] Knapsack problems have received little attention in this area and appear in only one reviewed study [240].

Other *Review Papers* included in our study synthesize how ML is used in OR across forecasting and optimization workflows, hybrid heuristics, and learning-to-optimize paradigms. They also catalog common pitfalls (data leakage, objective mismatch) and report evidence on when learning actually beats carefully tuned classical baselines [2, 25, 64, 141, 170].

# Appendix F A Glossary of Machine Learning Models

## F.1 Supervised Learning

**Boosting** Boosting is an ensemble technique that aims to create a strong classifier from an ensemble of weak classifiers. In a boosting approach, models are trained sequentially: each new model focuses on the training examples that were misclassified by the previous models, thereby progressively improving accuracy. The quintessential example is AdaBoost (Adaptive Boosting), introduced by Freund and Schapire, which assigns higher weights to misclassified instances on each round and combines many "weak" decision stumps into a powerful committee [55]. Boosting algorithms, including AdaBoost and later variants such as Gradient Boosting and XGBoost, have proven effective at reducing bias and variance, often achieving high predictive performance on classification tasks. By emphasizing the harder-to-predict cases, boosting can fit complex decision boundaries, though it must be regularized to avoid overfitting.

**Classification** Classification is a fundamental task in supervised machine learning where the objective is to assign input instances to one of several discrete categories or classes. In a classification problem, the learning algorithm is trained on a labeled dataset (examples with known class labels) and must learn a decision rule that can predict the correct class for new, unseen examples. Classic classification algorithms include logistic regression, which models the probability of class membership via a logistic function, decision trees, support vector machines, $k$-nearest neighbors, and neural network classifiers. The performance of a classifier is typically evaluated using metrics such as accuracy, precision, recall, and the F1-score. A good classifier will generalize well, meaning it achieves high accuracy not only on training data but also on new data, by capturing the true patterns that distinguish the classes rather than noise or incidental correlations [15].

**Decision Tree** A decision tree is a tree-structured predictive model used for both classification and regression tasks. In a decision tree, each internal node represents a test on a feature (attribute), each outgoing branch corresponds to one of the possible outcomes of that test, and each leaf node represents a predicted value or class label. The tree is constructed by recursively splitting the training data based on the feature that yields the highest information gain or impurity reduction at each node. This greedy splitting results in a flowchart-like model where following the path from the root to a leaf provides a series of decisions leading to a prediction. For classification trees, a leaf may assign a class, whereas for regression trees, each leaf typically contains a constant value. Decision trees form the foundation of more advanced ensemble methods such as random forests and boosting. The simplicity and interpretability of decision trees make them popular in applications where model transparency is important, despite their sometimes lower accuracy compared to other modern methods [15].

**Ensemble Learning** Ensemble learning is a strategy in machine learning where multiple models (often referred to as "base learners" or "weak learners") are combined to produce a more powerful model. The underlying principle is that a group of diverse models, when aggregated, can outperform any single model if their errors are at least somewhat uncorrelated. Ensembles can dramatically improve predictive performance and robustness, as the ensemble can "average out" the individual errors of its members. There are several main paradigms of ensemble learning: bagging and random forests, boosting, and stacking [21]. A necessary condition for ensemble performance gain is model diversity: the individual learners should make different errors. Techniques such as bagging and random feature selection ensure diversity among trees in a forest, while boosting introduces diversity by reweighting data. Empirically, ensemble models such as random forests and gradient boosting machines are among the top performers in many machine learning tasks and competitions [42]. The price of higher accuracy of ensemble models is that ensembles lose some interpretability and are computationally heavier since multiple models must be trained and stored. Boosting and Random Forest are explained separately.

**Linear Machine Learning** Linear models in machine learning are models that make predictions as a linear combination of input features. These models assume that the target outcome is (approximately) a linear function of the features. Linear models are among the simplest and most interpretable machine learning models. Examples include linear regression for predicting a continuous outcome and linear classifiers such as the perceptron or linear support vector machine, which find a linear decision boundary (hyperplane) that separates classes. Despite their simplicity, linear models can be very effective, especially for problems where the relationship between the features and the outcome is roughly linear or when the number of features is very large compared to the number of samples. Linear models assume the data is linearly separable or linearly related to the target; while this is a strong and over-simplifying assumption, it leads to simple, well-understood, and computationally efficient models [83].

**Logistic Regression** Logistic regression is a widely used statistical model for binary classification problems. Despite its name, logistic regression is actually a classification algorithm. It models the probability that a given input belongs to the positive class using the logistic function. In its simplest form for a binary outcome, logistic regression assumes a linear combination of the input features is transformed using the sigmoid function ($\sigma(t) = \frac{1}{1+e^{-t}}$) that outputs any real value into the range $[0, 1]$. The model outputs a probability between 0 and 1, and a decision threshold is used to assign the instance to one of the two classes. Logistic regression is trained by maximizing the likelihood of the observed labels using methods such as gradient descent. Probabilistic output and interpretability of logistic regression make it a popular method in many applied domains for predicting a yes/no outcome[83].

**Neural Networks** Neural networks, specifically artificial neural networks (ANNs), are computational models inspired by the networks of neurons in the brain. An ANN consists of simple processing units (neurons) organized in layers and connected by weighted links. Each neuron receives inputs from some other neurons, computes a weighted sum, applies a nonlinear activation function, and passes the result to neurons in the next layer. In a typical feedforward neural network, information flows from an input layer (representing the features of the data) through one or more hidden layers to an output layer (producing a prediction). Neural networks are a form of parametric function approximator that can approximate extremely complex functions. They learn by adjusting their weights based on training data. The most common training algorithm is some variant of stochastic gradient descent with backpropagation. One advantage of neural networks is their ability to automatically learn features from raw data, reducing the need for manual feature engineering. However, they typically require a lot of training data and computation, and the resulting models are not interpretable and can be prone to overfitting if not regularized. Despite these challenges, neural networks form the basis of many state-of-the-art systems in AI. As hardware and algorithms have improved, neural networks have transitioned from a relatively obscure technique to a dominant paradigm in machine learning [59].

**SVM (Support Vector Machine)** A Support Vector Machine (SVM) is a supervised learning algorithm traditionally used for classification. The core idea of SVM is to find the optimal separating hyperplane that maximizes the margin between two classes of data. Given labeled training data, a linear SVM will determine a hyperplane such that the margin (the distance between the hyperplane and the support vector) is as large as possible. This maximal margin criterion leads to a quadratic optimization problem with convex constraints, which can be solved efficiently. The result is a classifier with high generalization properties, particularly when classes are well-separated. For cases where the data are not linearly separable in the original feature space, SVMs employ the kernel trick. A kernel function implicitly maps inputs into a higher-dimensional feature space where a linear separator might exist, without explicitly computing the coordinates in that space. By using a kernel, SVM can form nonlinear decision boundaries in the original input space while still relying on the concept of a maximum-margin separation. Another strength of SVMs is the use of regularization through an optimization formulation that trades off margin size with classification error on the training set. SVM techniques are memory-efficient and have well-understood generalization bounds. However, SVMs do not scale as well to very large datasets and can be less effective on very high-dimensional sparse data compared to linear models with appropriate regularization [35].

**Quantum Supervised ML** Quantum supervised learning applies quantum computing to classical tasks such as classification and regression, aiming for computational or representational advantages. Instead of classical models, it uses parameterized quantum circuits, where data is encoded into quantum states (e.g., via amplitude or angle encoding). Model predictions are obtained from quantum measurements, and the circuit parameters are optimized in a hybrid quantum–classical loop to minimize a cost function, analogous to training in classical supervised learning. [187, 191]. The main promise of quantum supervised learning lies in its ability to exploit the exponentially large Hilbert space of quantum states to represent complex data distributions. In particular Quantum Supervised ML can implicitly compute inner products in high-dimensional spaces that may be classically intractable [72].

**Random Forest** A random forest is an ensemble learning method for classification and regression that operates by constructing a multitude of decision trees and outputting their aggregated result (majority vote for classification or average for regression). Random forests were introduced by Breiman [21] as an improvement over plain bagging of trees by adding an extra layer of randomness. In a random forest, each decision tree is trained on a bootstrapped sample of the training data, and additionally, at each candidate split in the tree, a random subset of features is considered as split candidates. This random feature selection ensures that the trees in the forest are decorrelated from each other, which is important because simply bagging many trees can still result in correlated predictions if one or a few strong predictors dominate the splits of all trees. By contrast, the randomness in feature selection forces the forest to explore different predictors and more diverse split structures. The combination of many high-variance, low-bias trees yields a model that often achieves both low bias and low variance. Random forests are renowned for their strong performance on many tasks, with relatively few hyperparameters to tune. They handle high-dimensional data well, can model nonlinear decision boundaries, and provide an internal estimate of generalization error without needing a separate validation set. However, they can be computationally intensive for large datasets and may not extrapolate well beyond the range of training data in regression tasks.

**Regression** In machine learning, regression refers to a family of supervised learning tasks where the goal is to predict a continuous-valued outcome given an input. This is in contrast to classification, which predicts discrete class labels. A regression model tries to learn the mapping that is a good estimate of the true response associated with the input. The simplest and most widely known regression method is linear regression, which assumes a linear combination of the input features as the mapping function. Linear regression can be solved analytically or with gradient-based optimization, and yields coefficients that indicate how each feature influences the output. When linear assumptions are too restrictive, one can use polynomial regression or other basis expansions, but these can lead to overfitting if the degree is high. The quality of a regression model is often evaluated using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), or $R^2$ (coefficient of determination). A key challenge in regression is avoiding overfitting, especially when the model is very flexible relative to the amount of training data. Techniques such as regularization used in Ridge or Lasso for linear models [83], or cross-validation to select model complexity, are commonly employed.

## F.2 Reinforcement Learning

**Reinforcement Learning** Reinforcement learning (RL) is a paradigm of learning where an agent learns to make decisions by performing actions in an environment and receiving feedback in the form of rewards or punishments. The goal of the agent is to learn a policy that maximizes the cumulative reward it expects to receive over time. Unlike supervised learning, the agent is never directly told the correct action for a given state; instead, it discovers good or bad actions through trial-and-error interactions. A formal framework for RL is typically given by a Markov Decision Process (MDP), defined by state space, action space, transition dynamics, reward function, and a discount factor. At each time step, the agent observes the current state, chooses an action, and then the environment transitions to a new state according to its dynamics and provides the agent a reward. Over an episode or over an infinite horizon, the agent accumulates rewards. The return is often defined as the discounted sum of future rewards. RL algorithms seek to maximize the expected return. RL methods include value-based methods (such as Q-learning), which learn value functions that estimate how good it is to take an action given a certain state. Another important RL technique is policy-based method, which directly adjusts the policy to maximize reward. Actor-critic methods combine value-based and policy-based methods by learning a value function (critic) to guide the improvement of the policy (actor). A hallmark of reinforcement learning is the exploration-exploitation trade-off: the agent must exploit known rewarding actions to gather reward, but it also must explore new actions to discover potentially better strategies. Techniques such as $\epsilon$-greedy or more sophisticated methods are used to manage this trade-off. Reinforcement learning is a very general framework, but it can be challenging because feedback is often sparse, delayed, or noisy, and the state space can be enormous. Sutton et al. [179] provides a thorough introduction to the theory and algorithms of reinforcement learning. Different types of RL algorithms are presented here:

- **Q-Learning (QL).** Q-learning is a model-free reinforcement learning algorithm that learns an action-value function, $Q(s, a)$, estimating the expected cumulative reward of taking action $a$ in state $s$ and following the optimal policy thereafter. It uses temporal-difference updates and is proven to converge to the optimal value function under certain conditions, even with exploratory actions [205].

- **Deep Q-Network (DQN).** DQN extends Q-learning to environments with large state spaces by approximating the $Q$-function with deep neural networks. Stabilization methods such as experience replay and target networks make training feasible, enabling human-level performance on Atari 2600 games [140].
- **Proximal Policy Optimization (PPO).** PPO is a policy-gradient algorithm that stabilizes training by using a clipped surrogate objective, preventing large updates that destabilize learning. Its balance between simplicity and robustness has made it a widely used baseline in continuous control tasks [177].
- **Deep Reinforcement Learning (DRL).** DRL combines reinforcement learning with deep neural networks, enabling agents to learn directly from high-dimensional inputs such as images or text. It underlies many breakthroughs such as AlphaGo, but faces challenges of instability and high sample complexity [3].
- **Actor-Critic Methods.** Actor-critic algorithms consist of an actor that learns the policy and a critic that estimates value functions to guide the actor. This reduces the variance of policy gradient estimates and improves learning stability, with popular variants including A2C and A3C [106].
- **Multi-Agent Reinforcement Learning (MARL).** In MARL, multiple agents learn and interact within a shared environment. Agents may cooperate, compete, or both, which makes the problem non-stationary and introduces challenges such as credit assignment. Applications span robotics, resource allocation, and autonomous driving [13].
- **Self-Supervised Reinforcement Learning .** Self-supervised learning introduces auxiliary tasks (e.g., prediction of future states or representations) that help RL agents learn robust features from interaction data, improving sample efficiency and generalization in sparse reward settings [61].
- **Hierarchical Reinforcement Learning (HRL).** HRL structures decision-making across multiple levels: high-level policies set goals or sub-tasks, and low-level controllers execute them. This improves exploration and efficiency in long-horizon tasks with delayed rewards [19].
- **Neuroevolutionary Methods.** Neuroevolution uses evolutionary algorithms to optimize neural networks for reinforcement learning, avoiding reliance on gradients. It fosters diversity and robustness, with notable methods such as NEAT evolving both architectures and weights [180].
- **Multi-Armed Bandits (MAB).** The bandit problem models the trade-off between exploration and exploitation: an agent repeatedly chooses among uncertain options to maximize cumulative reward. Solutions such as epsilon-greedy, UCB, and Thompson Sampling provide principled strategies with applications from online advertising to clinical trials [120].

## F.3 Deep Learning and Generative AI

**Deep Learning** Deep learning refers to a class of machine learning techniques based on artificial neural networks with multiple layers. Deep learning methods automatically learn hierarchical representations of data by composing nonlinear transformations across many layers of neurons. Each layer in a deep neural network transforms its input to a higher-level, more abstract representation: for example, in image analysis, early layers may learn to detect edges, mid-layers assemble edges into motifs or shapes, and deeper layers recognize complex objects. The hallmark of deep learning is that these features are learned from data, rather than manually engineered. Deep learning has dramatically improved the state-of-the-art in numerous domains such as computer vision, speech recognition, and natural language processing. This breakthrough came from advances in neural network architectures, large-scale datasets, increased computational power, and techniques to train very deep networks reliably. Training a deep network typically involves minimizing a loss function using stochastic gradient descent and backpropagation to adjust weights. Deep models often require a lot of data and computation, but when those requirements are met, they can excel at tasks such as image classification [110], machine translation, and even strategic games. The depth of the model allows it to capture highly complex functions and interactions in the data. While powerful, deep learning models can be seen as black boxes, raising issues in interpretability and requiring careful tuning of hyperparameters.

**Generative AI** Generative AI refers to models that learn the underlying distribution of training data in order to generate new, similar data samples. Unlike discriminative models (which predict a label or outcome given input features), generative models aim to capture data structure rules and create content that resembles the examples on which they were trained. Classic examples of generative models include probabilistic graphical models such as Bayesian networks, Hidden Markov Models, and topic models. Generative Adversarial Networks (GANs), introduced by Goodfellow et al. [67], put two neural networks (a generator and a discriminator) against each other in a minimax game, leading to remarkably realistic generated samples. Another major approach is variational autoencoders (VAEs), which pair an encoder and decoder network trained to reconstruct data through a latent distribution, thus learning to generate new data by sampling the latent space. More recently, transformer-based models such as GPT have been used in generative settings. Generative AI represents a powerful paradigm wherein machines exhibit creativity by learning to model complex data distributions and sampling from them [59].

**LSTM (Long Short-Term Memory)** LSTM is a special type of recurrent neural network (RNN) architecture designed to model sequential data and capture long-term dependencies. Introduced by Hochreiter and Schmidhuber [80], LSTM networks address the vanishing gradient problem that plagued earlier RNNs by regulating the information flow in their memory cell that contains three main gates: an input gate, a forget gate, and an output gate. This gating architecture allows LSTMs to maintain and update a cell state over long sequences, effectively remembering values for long durations (hence the name "long short-term" memory, because it learns long-term dependencies using short-term computational steps). In practice, LSTMs can learn to carry relevant information forward while discarding irrelevant information at each time step. This makes them extremely effective for tasks such as language modeling, machine translation, speech recognition, and time-series prediction, where context from far earlier in the sequence can be crucial. The development of LSTMs was a key step in enabling RNNs to learn long-term patterns in sequential data.

### F.4 Unsupervised Learning

**Clustering** Clustering is an unsupervised learning task that involves grouping a set of data points into clusters such that points in the same cluster are more similar to each other than to points in other clusters. Here, "similarity" is defined by a distance or similarity measure over the feature space of the data. The aim of clustering is to reveal the inherent structure or natural groupings in the data without any pre-provided labels. Common clustering algorithms include K-means, Hierarchical clustering, (Density-Based Spatial Clustering of Applications with Noise) DBSCAN, and Gaussian mixture models. Clustering is widely used in exploratory data analysis, where discovering latent group structure is useful [223].

**PCA (Principal Component Analysis)** Principal Component Analysis (PCA) is a dimensionality reduction technique and unsupervised learning method used to summarize and visualize data by reducing the number of its dimensions while retaining most of the variance present in the dataset. PCA works by finding new orthogonal axes (called principal components) that successively account for the largest possible variance in the data. The first principal component is the direction in the feature space along which the data have the highest variance. The second principal component is the orthogonal direction with the next highest variance, and so on. By projecting the original data onto the first few principal components, we obtain a lower-dimensional representation of the data that captures its most important structure [91]. PCA is especially useful for high-dimensional datasets where many features are correlated or redundant.

## Appendix G Geographical Scale of Tangible Supply Chain Problems

The scale of a supply chain problem profoundly influences its nature, complexity, and the types of decisions involved. In this review, we analyze the geographical scale that defines the physical scope of the optimization problem. This dimension is critical because decisions made at a local or city level are typically operational and tactical (e.g., last-mile delivery), whereas those at a national or global scale are strategic (e.g., *facility location*, network design). It is important to note that other dimensions of scale, such as the temporal scale (the length of the planning horizon, from real-time to multi-year) and the product scale (the number of SKUs), also significantly impact problem complexity, though our classification focuses on the geographical dimension. Our geographical scale analysis, as summarized in Table G3, reveals that a vast number of studies develop scale-agnostic methods, while significant research also targets specific challenges at both small and large geographical scales.

The largest group of 134 papers is classified as "Any", indicating that the developed models are not tied to a specific geographical context. These studies typically focus on advancing the methodology for solving a class of problems, such as *VRP* or inventory management, and are validated on benchmark instances. Much of the research in neural combinatorial optimization falls into this category, aiming to produce universal solution frameworks. For example, [136] develops a hybrid neural framework, [118] proposes a feature embedding refiner, and [230] uses reinforcement learning to automate metaheuristic design. Many studies tackle specific *VRP* variants without a fixed scale, including capacitated *VRP* [102, 134], *VRP* with time windows [133], multi-depot *VRP* [6, 7, 113, 189], and heterogeneous fleet *VRP* [119, 168, 251]. Other general *VRP* solution methods include the use of multi-agent RL [121], quantum algorithms [129, 130], and the development of novel learning-based heuristics [82, 208, 214]. Inventory management studies are also often scale-agnostic, presenting generalizable models for multi-echelon systems [171, 190], can-order policies [151], *lot-sizing* [147, 200], and vendor-managed inventory [183]. The category also includes papers on network design [10, 65], *bi-level* optimization [144], and *supplier selection* [86, 97]. Naturally, the 15 review articles fall into this category, as they cover methods applicable on multiple scales [2, 25, 54, 64, 68, 115, 137, 141, 170, 186, 188, 198, 216, 246, 247].

Research at the Small-Scale level, which includes 38 papers, focuses on operational and tactical problems within a local, city, or regional context. These studies often address challenges such as traffic congestion, tight delivery windows, and dynamic events. A major theme is urban logistics, with articles on reliable routing in dynamic road networks [112], routing for electric vehicle fleets [150], optimizing on-demand food delivery [227], and managing last-mile delivery [38, 64]. Several case-specific studies solve problems in particular regions, such as optimizing mobile medical unit routes [138], managing municipal solid waste [155], designing bioethanol supply chains [132], and optimizing dry port operations [123]. Drone delivery is another key application on this scale, as seen in the flying sidekick *TSP* model [117]. The category also includes articles on regional resource allocation for water [36, 74, 75, 81, 93], inventory management for red blood cells [108], and collaborative manufacturing planning [60]. Further work includes green *VRP* [92], adaptive routing for parcels [193], and solving stochastic *VRP* [37].

Finally, 27 papers focus on Large-Scale problems, which involve strategic or major tactical decisions at a national or multi-regional level. These studies tackle the complexity of large networks, long distances, and multiple supply chain echelons. A significant portion addresses national public service and healthcare logistics, such as the distribution of COVID-19 vaccines [181, 235], the management of national blood supply chains [1, 9], and the exchange of emergency resources [244]. Other national-level studies focus on sustainable network design for agri-food production [139], bioethanol [131], and biomass [31], as well as managing water resources [43, 44, 76, 101, 182] and pharmaceutical supply chains [199]. Another theme within this category is the development of methodologies specifically designed for scalability. This includes a review on the solution of large combinatorial problems in manufacturing [246], a model for large-scale location-production-routing [26], an approach for large-scale inventory optimization [210], a framework for dynamic *VRP*s with time windows [17], and models for large multi-echelon inventory systems [206].

Table G3: Scale of optimization problems in the reviewed supply chain optimization studies, sorted by frequency of use.

| Purpose of Use | Count | References |
|---|---|---|
| Any | 134 | [2, 6–8, 10–12, 16, 20, 25, 28, 32, 33, 38, 39, 45–50, 52–54, 62–66, 68, 69, 71, 77–79, 82, 84, 85, 89, 90, 94–96, 99, 100, 102, 104, 107, 113–116, 118, 119, 121, 122, 125, 127–130, 133–137, 141–145, 147, 149, 151, 152, 162, 165, 167–174, 178, 183, 184, 186, 188–190, 192–196, 198, 200–203, 207, 208, 211–216, 218–220, 222, 227, 228, 230–234, 236, 239–243, 245–249, 251, 252] |
| Small-Scale | 38 | [4, 27, 36, 37, 51, 56, 58, 60, 74, 75, 81, 86, 88, 92, 93, 97, 98, 105, 108, 111, 112, 117, 123, 124, 126, 132, 138, 146, 148, 150, 155, 156, 158, 160, 164, 176, 238, 250] |
| Large-Scale | 27 | [1, 9, 17, 26, 31, 34, 43, 44, 57, 70, 73, 76, 87, 101, 109, 131, 139, 175, 181, 182, 199, 206, 210, 221, 229, 235, 244] |

# References


[1] Abbasi, B., Babaei, T., Hosseinifard, Z., Smith-Miles, K., Dehghani, M.: Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management. Computers and Operations Research **119** (2020) https://doi.org/10.1016/j.cor.2020.104941

[2] Akbari, M., Do, T.N.A.: A systematic review of machine learning in logistics and supply chain management: current trends and future directions. Benchmarking **28**(10), 2977–3005 (2021) https://doi.org/10.1108/BIJ-10-2020-0514

[3] Arulkumaran, K., Deisenroth, M.P., Brundage, M., Bharath, A.A.: Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine **34**(6), 26–38 (2017) https://doi.org/10.1109/MSP.2017.2743240

[4] Aghaie, A., Heidary, M.H.: Simulation-based optimization of a stochastic supply chain considering supplier disruption: Agent-based modeling and reinforcement learning. Scientia Iranica **26**(6), 3780–3795 (2019) https://doi.org/10.24200/sci.2018.20789

[5] Arrow, K.J., Harris, T., Marschak, J.: Optimal inventory policy. Econometrica **19**(3), 250–272 (1951) https://doi.org/10.2307/1906813

[6] Arishi, A., Krishnan, K.: A multi-agent deep reinforcement learning approach for solving the multi-depot vehicle routing problem. Journal of Management Analytics **10**(3), 493–515 (2023) https://doi.org/10.1080/23270012.2023.2229842

[7] Anuar, W.K., Lee, L.S., Seow, H.V., Pickl, S.: A multi-depot vehicle routing

problem with stochastic road capacity and reduced two-stage stochastic integer linear programming models for rollout algorithm. Mathematics **9**(13) (2021) https://doi.org/10.3390/math9131572

[8] Anderson, T., Martínez, S.: Distributed resource allocation with binary decisions via Newton-like Neural Network dynamics. Automatica **128** (2021) https://doi.org/10.1016/j.automatica.2021.109564

[9] Altaf, M.M., Roshdy, A.S., AlSagri, H.S.: Deep reinforcement learning model for blood bank vehicle routing multi-objective optimization. Computers, Materials and Continua **70**(2), 3955–3967 (2022) https://doi.org/10.32604/cmc.2022.019448

[10] Akbarian-Saravi, N., Mobini, M., Rabbani, M.: Development of a comprehensive decision support tool for strategic and tactical planning of a sustainable bioethanol supply chain: Real case study, discussions and policy implications. Journal of Cleaner Production **244** (2020) https://doi.org/10.1016/j.jclepro.2019.118871

[11] Ashtab, S., Tosarkani, B.M.: Scenario-based multi-objective optimisation model based on supervised machine learning to configure a plastic closed-loop supply chain network. International Journal of Business Performance and Supply Chain Modelling **14**(1), 106–128 (2023) https://doi.org/10.1504/IJBPSCM.2023.130469

[12] Bandaru, S., Aslam, T., Ng, A.H.C., Deb, K.: Generalized higher-level automated innovization with application to inventory management. European Journal of Operational Research **243**(2), 480–496 (2015) https://doi.org/10.1016/j.ejor.2014.11.015

[13] Busoniu, L., Babuska, R., De Schutter, B.: A comprehensive survey of multiagent reinforcement learning. IEEE Transactions on Systems, Man, and Cybernetics, Part C **38**(2), 156–172 (2008) https://doi.org/10.1109/TSMCC.2007.913919

[14] Bellman, R.: The Theory of Dynamic Programming. RAND Corporation, Santa Monica (1954). https://www.ams.org/journals/bull/1954-60-06/S0002-9904-1954-09848-8/

[15] Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. Wadsworth International Group, Belmont, CA (1984). https://doi.org/10.1201/9781315139470

[16] Bhosekar, A., Ierapetritou, M.: A framework for supply chain optimization for modular manufacturing with production feasibility analysis. Computers and Chemical Engineering **145** (2021) https://doi.org/10.1016/j.compchemeng.2020.107175

[17] Baty, L., Jungel, K., Klein, P.S., Parmentier, A., Schiffer, M.: Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows. Transportation Science **58**(4), 708–725 (2024) https://doi.org/10.1287/trsc.2023.0107

[18] Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer, New York (2011)

[19] Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning. Discrete Event Dynamic Systems **13**(1-2), 41–77 (2003) https://doi.org/10.1023/A:1025696116075

[20] Boccia, M., Mancuso, A., Masone, A., Murino, T., Sterle, C.: New features for customer classification in the Flying Sidekick Traveling Salesman Problem. Expert Systems with Applications **247** (2024) https://doi.org/10.1016/j.eswa.2023.123106

[21] Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001) https://doi.org/10.1023/A:1010933404324

[22] Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)

[23] Büyüktahtakın, I.E.: Stage-t scenario dominance for risk-averse multi-stage stochastic mixed-integer programs. Annals of Operations Research **309**(1), 1–35 (2022) https://doi.org/10.1007/s10479-021-04388-3

[24] Büyüktahtakın, İ.E.: Scenario-dominance to multi-stage stochastic lot-sizing and knapsack problems. Computers & Operations Research **153**, 106149 (2023) https://doi.org/10.1016/j.cor.2023.106149

[25] Baryannis, G., Validi, S., Dani, S., Antoniou, G.: Supply chain risk management and artificial intelligence: state of the art and future research directions. International Journal of Production Research **57**(7), 2179–2202 (2019) https://doi.org/10.1080/00207543.2018.1530476

[26] Bootaki, B., Zhang, G.: A location-production-routing problem for distributed manufacturing platforms: A neural genetic algorithm solution methodology. International Journal of Production Economics **275** (2024) https://doi.org/10.1016/j.ijpe.2024.109325

[27] Chobar, A.P., Adibi, M.A., Kazemi, A.: Multi-objective hub-spoke network design of perishable tourism products using combination machine learning and meta-heuristic algorithms. Environment, Development and Sustainability (2022) https://doi.org/10.1007/s10668-022-02350-2

[28] Chi, C.Y., Chen, D.F., Doan, H.P., Kuo, C.H.: UAV parcel Delivery System with

Deep Reinforcement Learning Based Collision Avoidance and Route Optimization. International Journal of Automation and Smart Technology **14**(1), 8–18 (2024) https://doi.org/10.5875/P25FHT10

[29] Coffman Jr, E.G., Garey, M.R., Johnson, D.S.: Approximation algorithms for bin-packing—an updated survey. In: Algorithm Design for Computer System Design, pp. 49–106. Springer, New York (1984). https://doi.org/10.1007/978-3-7091-4338-4_3

[30] Colson, B., Marcotte, P., Savard, G.: An overview of bilevel optimization. Annals of Operations Research **153**(1), 235–256 (2007) https://doi.org/10.1007/s10479-007-0176-2

[31] Christou, I.T., Psathas, F., Rentizelas, A., Papadakis, A., Georgiou, P.N., Anastasopoulos, D., Lappas, P.: A heuristic for improving clustering in biomass supply chains. International Journal of Systems Science: Operations and Logistics **11**(1) (2024) https://doi.org/10.1080/23302674.2024.2378859

[32] Chen, B., Qu, R., Bai, R., Laesanklang, W.: A variable neighborhood search algorithm with reinforcement learning for a real-life periodic vehicle routing problem with time windows and open routes. RAIRO - Operations Research **54**(5), 1467–1494 (2020) https://doi.org/10.1051/ro/2019080

[33] Chamurally, S., Rieck, J.: A practical and robust approach for solving the multi-compartment vehicle routing problem under demand uncertainty using machine learning. Networks **84**(3), 300–325 (2024) https://doi.org/10.1002/net.22231

[34] Chen, Z.Y., Sun, M., Han, X.X.: Prediction-driven collaborative emergency medical resource allocation with deep learning and optimization. Journal of the Operational Research Society **74**(2), 590–603 (2023) https://doi.org/10.1080/01605682.2022.2101953

[35] Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning **20**(3), 273–297 (1995) https://doi.org/10.1007/BF00994018

[36] Cai, Y., Wang, H., Yue, W., Xie, Y., Liang, Q.: An integrated approach for reducing spatially coupled water-shortage risks of Beijing-Tianjin-Hebei urban agglomeration in China. Journal of Hydrology **603** (2021) https://doi.org/10.1016/j.jhydrol.2021.127123

[37] Cai, H., Xu, P., Tang, X., Lin, G.: Solving the Vehicle Routing Problem with Stochastic Travel Cost Using Deep Reinforcement Learning. Electronics (Switzerland) **13**(16) (2024) https://doi.org/10.3390/electronics13163242

[38] Chu, H., Zhang, W., Bai, P., Chen, Y.: Data-driven optimization for last-mile delivery. Complex and Intelligent Systems **9**(3), 2271–2284 (2023) https://doi.org/10.1007/s40747-021-00293-1

[39] Chen, J., Zong, Z., Zhuang, Y., Yan, H., Jin, D., Li, Y.: Reinforcement Learning for Practical Express Systems with Mixed Deliveries and Pickups. ACM Transactions on Knowledge Discovery from Data **17**(3) (2023) https://doi.org/10.1145/3546952

[40] Dantzig, G.B.: Linear programming. Operations research **50**(1), 42–47 (2002) https://doi.org/10.1287/opre.50.1.42.17798

[41] Dantzig, G.B., Fulkerson, D.R., Johnson, S.M.: Solution of a large-scale traveling-salesman problem. Journal of the Operations Research Society of America **2**(4), 393–410 (1954) https://doi.org/10.1287/opre.2.4.393

[42] Dietterich, T.G.: Ensemble methods in machine learning. In: Multiple Classifier Systems (MCS 2000). Lecture Notes in Computer Science, vol. 1857, pp. 1–15. Springer, Berlin, Heidelberg (2000). https://doi.org/10.1007/3-540-45014-9_1

[43] Dariane, A.B., Moradi, A.M.: Comparative analysis of evolving artificial neural network and reinforcement learning in stochastic optimization of multireservoir systems. Hydrological Sciences Journal **61**(6), 1141–1156 (2016) https://doi.org/10.1080/02626667.2014.986485

[44] Delavar, M., Moghadasi, M., Morid, S.: Real-Time Model for Optimal Water Allocation in Irrigation Systems during Droughts. Journal of Irrigation and Drainage Engineering **138**(6), 517–524 (2012) https://doi.org/10.1061/(ASCE)IR.1943-4774.0000440

[45] Dornemann, J.: Solving the capacitated vehicle routing problem with time windows via graph convolutional network assisted tree search and quantum-inspired computing. Frontiers in Applied Mathematics and Statistics **9** (2023) https://doi.org/10.3389/fams.2023.1155356

[46] Deng, J., Wang, J., Wang, X., Cai, Y., Liu, P.: Multi-Task Multi-Objective Evolutionary Search Based on Deep Reinforcement Learning for Multi-Objective Vehicle Routing Problems with Time Windows. Symmetry **16**(8) (2024) https://doi.org/10.3390/sym16081030

[47] Dong, W., Zhang, X., Guo, H., Qu, J., Dong, L., Zhang, H., Jiang, Y.: Optimization of Sustainable Automotive Manufacturing Supply Chain with Demand Prediction: A Combinatorial Framework. International Journal of Industrial Engineering : Theory Applications and Practice **30**(4), 1042–1051 (2023) https://doi.org/10.23055/ijietap.2023.30.4.9035

[48] Ezazipour, S., Golbabai, A.: A globally convergent neurodynamics optimization model for mathematical programming with equilibrium constraints. Kybernetika **56**(3), 383–409 (2020) https://doi.org/10.14736/kyb-2020-3-0383

[49] Ewbank, H., Wanke, P., Hadi-Vencheh, A.: An unsupervised fuzzy clustering

approach to the capacitated vehicle routing problem. Neural Computing and Applications **27**(4), 857–867 (2016) https://doi.org/10.1007/s00521-015-1901-4

[50] Fitzpatrick, J., Ajwani, D., Carroll, P.: A scalable learning approach for the capacitated vehicle routing problem. Computers and Operations Research **171** (2024) https://doi.org/10.1016/j.cor.2024.106787

[51] Fellek, G., Farid, A., Gebreyesus, G., Fujimura, S., Yoshie, O.: Graph Transformer with Reinforcement Learning for Vehicle Routing Problem. IEEJ Transactions on Electrical and Electronic Engineering **18**(5), 701–713 (2023) https://doi.org/10.1002/tee.23771

[52] Frias, N., Johnson, F., Valle, C.: Hybrid Algorithms for Energy Minimizing Vehicle Routing Problem: Integrating Clusterization and Ant Colony Optimization. IEEE Access **11**, 125800–125821 (2023) https://doi.org/10.1109/ACCESS.2023.3325787

[53] Furian, N., O'Sullivan, M., Walker, C., Cela, E.: A machine learning-based branch and price algorithm for a sampled vehicle routing problem. OR Spectrum **43**(3), 693–732 (2021) https://doi.org/10.1007/s00291-020-00615-8

[54] Frederico, G.F.: ChatGPT in Supply Chains: Initial Evidence of Applications and Potential Research Agenda. Logistics **7**(2) (2023) https://doi.org/10.3390/logistics7020026

[55] Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences **55**(1), 119–139 (1997) https://doi.org/10.1006/jcss.1997.1504

[56] Flores, H., Villalobos, J.R.: A stochastic planning framework for the discovery of complementary, agricultural systems. European Journal of Operational Research **280**(2), 707–729 (2020) https://doi.org/10.1016/j.ejor.2019.07.053

[57] Fu, Y., Wang, Y., Gao, K., Suganthan, P.N., Huang, M.: Integrated scheduling of multi-constraint open shop and vehicle routing: Mathematical model and learning-driven brain storm optimization algorithm. Applied Soft Computing **163** (2024) https://doi.org/10.1016/j.asoc.2024.111943

[58] Fu, J., Yang, Y.: Construction of an intelligent optimization model for water resource allocation under drought conditions. Applied Mathematics and Nonlinear Sciences **9**(1) (2024) https://doi.org/10.2478/amns-2024-1162

[59] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, MA (2016)

[60] Guo, Y., Chen, T., Boulaksil, Y., Xiao, L., Allaoui, H.: Collaborative planning of multi-tier sustainable supply chains: A reinforcement learning enhanced heuristic

approach. Computers and Industrial Engineering **185** (2023) https://doi.org/10.1016/j.cie.2023.109669

[61] Gui, J., Chen, T., Zhang, J., Cao, Q., Sun, Z., Luo, H., Tao, D.: A Survey on Self-supervised Learning: Algorithms, Applications, and Future Trends (2024). https://arxiv.org/abs/2301.05712

[62] Gumte, K.M., Devi Pantula, P., Miriyala, S.S., Mitra, K.: Data driven robust optimization for handling uncertainty in supply chain planning models. Chemical Engineering Science **246** (2021) https://doi.org/10.1016/j.ces.2021.116889

[63] Golbabai, A., Ezazipour, S.: A projection-based recurrent neural network and its application in solving convex quadratic bilevel optimization problems. Neural Computing and Applications **32**(8), 3887–3900 (2020) https://doi.org/10.1007/s00521-019-04391-7

[64] Giuffrida, N., Fajardo-Calderin, J., Masegosa, A.D., Werner, F., Steudter, M., Pilla, F.: Optimization and Machine Learning Applied to Last-Mile Logistics: A Review. Sustainability (Switzerland) **14**(9) (2022) https://doi.org/10.3390/su14095329

[65] Gumus, A.T., Guneri, A.F., Keles, S.: Supply chain network design using an integrated neuro-fuzzy and MILP approach: A comparative design study. Expert Systems with Applications **36**(10), 12570–12577 (2009) https://doi.org/10.1016/j.eswa.2009.05.034

[66] Galli, L., Levato, T., Schoen, F., Tigli, L.: Prescriptive analytics for inventory management in health care. Journal of the Operational Research Society **72**(10), 2211–2224 (2021) https://doi.org/10.1080/01605682.2020.1776167

[67] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems (NIPS 2014), pp. 2672–2680 (2014)

[68] Gutierrez, J.C., Polo Triana, S.I., León Becerra, J.S.: Benefits, challenges, and limitations of inventory control using machine learning algorithms: literature review. OPSEARCH (2024) https://doi.org/10.1007/s12597-024-00839-0

[69] Geraeli, M., Roghanian, E.: A novel adjusted real-time decision-making for dynamic distribution in the grocery supply chain. Journal of Modelling in Management (2024) https://doi.org/10.1108/JM2-02-2024-0048

[70] Han, X.: Path Planning Algorithm for the Multiple Depot Vehicle Routing Problem Based on Parallel Clustering. Scientific Programming **2023** (2023) https://doi.org/10.1155/2023/7588595

[71] Henao-Baena, C.A., Zuluaga-Zuluaga, B., Galeano-Castro, J., Marín-García,

E.J., Calvo-Salcedo, A.F.: Methodology for Inventory Management in Neighborhood Stores Using Machine Learning and Integer Linear Programming. Ingenieria (Colombia) **29**(1) (2024) https://doi.org/10.14483/23448393.19423

[72] Havlíček, V., Córcoles, A.D., Temme, K., Harrow, A.W., Kandala, A., Chow, J.M., Gambetta, J.M.: Supervised learning with quantum-enhanced feature spaces. Nature **567**(7747), 209–212 (2019) https://doi.org/10.1038/s41586-019-0980-2

[73] Hu, S., Dong, Z.S., Dai, R.: A machine learning based sample average approximation for supplier selection with option contract in humanitarian relief. Transportation Research Part E: Logistics and Transportation Review **186** (2024) https://doi.org/10.1016/j.tre.2024.103531

[74] Hou, J., Fan, X., Liu, R.: Optimal spatial allocation of irrigation water under uncertainty using the bilayer nested optimisation algorithm and geospatial technology. International Journal of Geographical Information Science **30**(12), 2462–2485 (2016) https://doi.org/10.1080/13658816.2016.1181264

[75] Han, Y., Huang, Y., Jia, S., Liu, J.: An interval-parameter fuzzy linear programming with stochastic vertices model for water resources management under uncertainty. Mathematical Problems in Engineering **2013** (2013) https://doi.org/10.1155/2013/942343

[76] Han, Y., Huang, Y., Wang, G.: Interval-parameter linear optimization model with stochastic vertices for land and water resources allocation under dual uncertainty. Environmental Engineering Science **28**(3), 197–205 (2011) https://doi.org/10.1089/ees.2010.0160

[77] He, X., Li, C., Huang, T., Huang, J.: A recurrent neural network for solving bilevel linear programming problem. IEEE Transactions on Neural Networks and Learning Systems **25**(4), 824–830 (2014) https://doi.org/10.1109/TNNLS.2013.2280905

[78] He, B., Meng, Y., Tang, L.: An Off-Policy Reinforcement Learning-Based Adaptive Optimization Method for Dynamic Resource Allocation Problem. IEEE Transactions on Neural Networks and Learning Systems, 1–15 (2023) https://doi.org/10.1109/TNNLS.2023.3338237

[79] Hammler, P., Riesterer, N., Braun, T.: Fully dynamic reorder policies with deep reinforcement learning for multi-echelon inventory management. Informatik-Spektrum **46**(5-6), 240–251 (2023) https://doi.org/10.1007/s00287-023-01556-6

[80] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997) https://doi.org/10.1162/neco.1997.9.8.1735

[81] Heydari, F., Saghafian, B., Delavar, M.: Coupled Quantity-Quality Simulation-Optimization Model for Conjunctive Surface-Groundwater Use. Water Resources Management **30**(12), 4381–4397 (2016) https://doi.org/10.1007/s11269-016-1426-3

[82] Hottung, A., Tierney, K.: Neural large neighborhood search for routing problems. Artificial Intelligence **313** (2022) https://doi.org/10.1016/j.artint.2022.103786

[83] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer, New York (2009)

[84] Hansuwa, S., Velayudhan Kumar, M.R., Chandrasekharan, R.: Analysis of box and ellipsoidal robust optimization, and attention model based reinforcement learning for a robust vehicle routing problem. Sadhana - Academy Proceedings in Engineering Sciences **47**(2) (2022) https://doi.org/10.1007/s12046-022-01833-2

[85] Hou, Y., Wang, H., Huang, X.: A Q-learning-based multi-objective evolutionary algorithm for integrated green production and distribution scheduling problems. Engineering Applications of Artificial Intelligence **127** (2024) https://doi.org/10.1016/j.engappai.2023.107434

[86] Islam, S., Amin, S.H., Wardley, L.J.: Machine learning and optimization models for supplier selection and order allocation planning. International Journal of Production Economics **242** (2021) https://doi.org/10.1016/j.ijpe.2021.108315

[87] Islam, S., Amin, S.H., Wardley, L.J.: A supplier selection & order allocation planning framework by integrating deep learning, principal component analysis, and optimization techniques. Expert Systems with Applications **235** (2024) https://doi.org/10.1016/j.eswa.2023.121121

[88] Irannezhad, E., Prato, C.G., Hickman, M.: An intelligent decision support system prototype for hinterland port logistics. Decision Support Systems **130** (2020) https://doi.org/10.1016/j.dss.2019.113227

[89] Iklassov, Z., Sobirov, I., Solozabal, R., Takac, M.: Reinforcement Learning Approach to Stochastic Vehicle Routing Problem With Correlated Demands. IEEE Access **11**, 87958–87969 (2023) https://doi.org/10.1109/ACCESS.2023.3306076

[90] Jin, J., Cui, T., Bai, R., Qu, R.: Container port truck dispatching optimization using Real2Sim based deep reinforcement learning. European Journal of Operational Research **315**(1), 161–175 (2024) https://doi.org/10.1016/j.ejor.2023.11.038

[91] Jolliffe, I.T.: Principal Component Analysis, 2nd edn. Springer, New York (2002)

[92] Jovanović, A.D., Pamučar, D.S., Pejčić-Tarle, S.: Green vehicle routing in urban zones - A neuro-fuzzy approach. Expert Systems with Applications **41**(7), 3189–3203 (2014) https://doi.org/10.1016/j.eswa.2013.11.015

[93] Kazemi, M., Bozorg-Haddad, O., Fallah-Mehdipour, E., Loáiciga, H.A.: Inter-basin hydropolitics for optimal water resources allocation. Environmental Monitoring and Assessment **192**(7) (2020) https://doi.org/10.1007/s10661-020-08439-3

[94] Küp, E.T., Cebeci, S., Bayram, B., Aydın, G., Bozkaya, B., Akhavan-Tabatabaei, R.: An Integrated Framework for Dynamic Vehicle Routing Problems with Pick-up and Delivery Time Windows and Shared Fleet Capacity Planning. Symmetry **16**(4) (2024) https://doi.org/10.3390/sym16040505

[95] Kou, A.Q., Cheng, Y., Guo, L., Xia, X.Q.: A Data-Driven Newsvendor Model with Unknown Demand and Supply Distribution. JOURNAL OF INDUSTRIAL AND MANAGEMENT OPTIMIZATION (2024) https://doi.org/10.3934/jimo.2024080

[96] Keith, K., Castillo-Villar, K.K., Alaeddini, A.: Machine Learning-Based Problem Space Reduction in Stochastic Programming Models: An Application in Biofuel Supply Chain Network Design. IEEE Access **12**, 33852–33866 (2024) https://doi.org/10.1109/ACCESS.2024.3372516

[97] Khalili-Damghani, K., Dideh-Khani, H., Sadi-Nezhad, S.: A two-stage approach based on anfis and fuzzy goal programming for supplier selection. International Journal of Applied Decision Sciences **6**(1), 1–14 (2013) https://doi.org/10.1504/IJADS.2013.052629

[98] Khalili-Fard, A., Hashemi, M., Bakhshi, A., Yazdani, M., Jolai, F., Aghsami, A.: Integrated relief pre-positioning and procurement planning considering non-governmental organizations support and perishable relief items in a humanitarian supply chain network. Omega (United Kingdom) **127** (2024) https://doi.org/10.1016/j.omega.2024.103111

[99] Kong, F., Jiang, B., Wang, J., Wang, H., Song, H.: Collaborative Delivery Optimization with Multiple Drones via Constrained Hybrid Pointer Network. IEEE Internet of Things Journal **11**(5), 7739–7755 (2024) https://doi.org/10.1109/JIOT.2023.3318524

[100] Kochakkashani, F., Kayvanfar, V., Baldacci, R.: Innovative Applications of Unsupervised Learning in Uncertainty-Aware Pharmaceutical Supply Chain Planning. IEEE Access **12**, 107984–107999 (2024) https://doi.org/10.1109/ACCESS.2024.3435439

[101] Kayhomayoon, Z., Milan, S.G., Azar, N.A., Bettinger, P., Babaian, F., Jaafari, A.: A Simulation-Optimization Modeling Approach for Conjunctive Water Use

Management in a Semi-Arid Region of Iran. Sustainability (Switzerland) **14**(5) (2022) https://doi.org/10.3390/su14052691

[102] Kim, H., Park, J., Kwon, C.: A Neural Separation Algorithm for the Rounded Capacity Inequalities. INFORMS Journal on Computing **36**(4), 987–1005 (2024) https://doi.org/10.1287/ijoc.2022.0310

[103] Kellerer, H., Pferschy, U., Pisinger, D.: Knapsack Problems. Springer, New York (2004). https://doi.org/10.1007/978-3-540-24777-7

[104] Kalatzantonakis, P., Sifaleras, A., Samaras, N.: A reinforcement learning-Variable neighborhood search method for the capacitated Vehicle Routing Problem. Expert Systems with Applications **213** (2023) https://doi.org/10.1016/j.eswa.2022.118812

[105] Kalantari, S., Safarzadeh Ramhormozi, R., Wang, Y., Sun, S., Wang, X.: Trailer allocation and truck routing using bipartite graph assignment and deep reinforcement learning. Transactions in GIS **27**(4), 996–1020 (2023) https://doi.org/10.1111/tgis.13057

[106] Konda, V.R., Tsitsiklis, J.N.: Actor-critic algorithms. Advances in Neural Information Processing Systems **12** (2000)

[107] Lee, D.H., Ahn, J.: Multi-start team orienteering problem for UAS mission re-planning with data-efficient deep reinforcement learning. Applied Intelligence **54**(6), 4467–4489 (2024) https://doi.org/10.1007/s10489-024-05367-4

[108] Li, N., Arnold, D.M., Down, D.G., Barty, R., Blake, J., Chiang, F., Courtney, T., Waito, M., Trifunov, R., Heddle, N.M.: From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization. Transfusion **62**(1), 87–99 (2022) https://doi.org/10.1111/trf.16739

[109] Lee, K.H., Abdollahian, M., Schreider, S., Taheri, S.: Supply Chain Demand Forecasting and Price Optimisation Models with Substitution Effect. Mathematics **11**(11) (2023) https://doi.org/10.3390/math11112502

[110] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015) https://doi.org/10.1038/nature14539

[111] Liu, J., Chen, W., Yang, J., Xiong, H., Chen, C.: Iterative Prediction-and-Optimization for E-Logistics Distribution Network Design. INFORMS Journal on Computing **34**(2), 769–789 (2022) https://doi.org/10.1287/ijoc.2021.1107

[112] Li, B., Dai, T., Chen, W., Song, X., Zang, Y., Huang, Z., Lin, Q., Cai, K.: T-PORP: A Trusted Parallel Route Planning Model on Dynamic Road Networks. IEEE Transactions on Intelligent Transportation Systems **24**(1), 1238–1250

(2023) https://doi.org/10.1109/TITS.2022.3216310

[113] Li, J., Dai, B.T., Niu, Y., Xiao, J., Wu, Y.: Multi-Type Attention for Solving Multi-Depot Vehicle Routing Problems. IEEE Transactions on Intelligent Transportation Systems, 1–10 (2024) https://doi.org/10.1109/TITS.2024.3413077

[114] Li, M., Hao, J.K., Wu, Q.: A flow based formulation and a reinforcement learning based strategic oscillation for cross-dock door assignment. European Journal of Operational Research **312**(2), 473–492 (2024) https://doi.org/10.1016/j.ejor.2023.07.014

[115] Laldin Ismaeil, M.K., Lalla, A.F.: The Role and Impact of Artificial Intelligence on Supply Chain Management: Efficiency, Challenges, and Strategic Implementation. Journal of Ecohumanism **3**(4), 89–106 (2024) https://doi.org/10.62754/joe.v3i4.3461

[116] Liu, C., Lyu, J., Fang, K.: Integrated packing and routing: A model and its solutions. Computers and Operations Research **172** (2024) https://doi.org/10.1016/j.cor.2024.106790

[117] Liu, Z., Li, X., Khojandi, A.: The flying sidekick traveling salesman problem with stochastic travel time: A reinforcement learning approach. Transportation Research Part E: Logistics and Transportation Review **164** (2022) https://doi.org/10.1016/j.tre.2022.102816

[118] Li, J., Ma, Y., Cao, Z., Wu, Y., Song, W., Zhang, J., Chee, Y.M.: Learning Feature Embedding Refiner for Solving Vehicle Routing Problems. IEEE Transactions on Neural Networks and Learning Systems, 1–13 (2023) https://doi.org/10.1109/TNNLS.2023.3285077

[119] Li, J., Ma, Y., Gao, R., Cao, Z., Lim, A., Song, W., Zhang, J.: Deep Reinforcement Learning for Solving the Heterogeneous Capacitated Vehicle Routing Problem. IEEE Transactions on Cybernetics **52**(12), 13572–13585 (2022) https://doi.org/10.1109/TCYB.2021.3111082

[120] Lattimore, T., Szepesvári, C.: Bandit Algorithms. Cambridge University Press, Cambridge, UK (2020)

[121] Lopes Silva, M.A., Souza, S.R., Freitas Souza, M.J., Bazzan, A.L.C.: A reinforcement learning-based multi-agent framework applied for solving routing and scheduling problems. Expert Systems with Applications **131**, 148–171 (2019) https://doi.org/10.1016/j.eswa.2019.04.056

[122] Liu, S., Wang, J., Wang, R., Zhang, Y., Song, Y., Xing, L.: Data-driven dynamic pricing and inventory management of an omni-channel retailer in an uncertain demand environment. Expert Systems with Applications **244** (2024) https://doi.org/10.1016/j.eswa.2023.122948

[123] Li, Q., Wang, Y., Xiong, Y., Zhang, S., Zhou, Y.: Machine learning-based optimisation in a two-echelon logistics network for the dry port operation in China. International Journal of Systems Science: Operations and Logistics **10**(1) (2023) https://doi.org/10.1080/23302674.2023.2252321

[124] Li, Y., Xu, M., Wang, H., Zhang, X., Chen, Y., Liu, Y., Liang, Y., Jiang, Y.: Modeling of Cross-Network Collaboration Supply Chain and Solution via Reinforcement Learning-Based Evolutionary Algorithm. International Journal of Industrial Engineering : Theory Applications and Practice **31**(2), 204–233 (2024) https://doi.org/10.23055/ijietap.2024.31.2.9523

[125] Liu, H., Zong, Z., Li, Y., Jin, D.: NeuroCrossover: An intelligent genetic locus selection scheme for genetic algorithm using reinforcement learning. Applied Soft Computing **146** (2023) https://doi.org/10.1016/j.asoc.2023.110680

[126] Liu, F., Zhang, Q., Zhu, Q., Tong, X., Yuan, M.: Machine Learning Assisted Multiobjective Evolutionary Algorithm for Routing and Packing. IEEE Transactions on Evolutionary Computation, 1–1 (2024) https://doi.org/10.1109/TEVC.2024.3357819

[127] Mahmud, S., Abbasi, A., Chakrabortty, R.K., Ryan, M.J.: A self-adaptive hyper-heuristic based multi-objective optimisation approach for integrated supply chain scheduling problems. Knowledge-Based Systems **251** (2022) https://doi.org/10.1016/j.knosys.2022.109190

[128] Mokhtarinejad, M., Ahmadi, A., Karimi, B., Rahmati, S.H.A.: A novel learning based approach for a new integrated location-routing and scheduling problem within cross-docking considering direct shipment. Applied Soft Computing Journal **34**, 274–285 (2015) https://doi.org/10.1016/j.asoc.2015.04.062

[129] Mohanty, N., Behera, B.K., Ferrie, C.: Analysis of the Vehicle Routing Problem Solved via Hybrid Quantum Algorithms in the Presence of Noisy Channels. IEEE Transactions on Quantum Engineering **4** (2023) https://doi.org/10.1109/TQE.2023.3303989

[130] Mohanty, N., Behera, B.K., Ferrie, C.: Solving the vehicle routing problem via quantum support vector machines. Quantum Machine Intelligence **6**(1) (2024) https://doi.org/10.1007/s42484-024-00161-4

[131] Momenitabar, M., Dehdari Ebrahimi, Z., Abdollahi, A., Helmi, W., Bengtson, K., Ghasemi, P.: An integrated machine learning and quantitative optimization method for designing sustainable bioethanol supply chain networks. Decision Analytics Journal **7** (2023) https://doi.org/10.1016/j.dajour.2023.100236

[132] Momenitabar, M., Dehdari Ebrahimi, Z., Ghasemi, P.: Designing a sustainable bioethanol supply chain network: A combination of machine learning and meta-heuristic algorithms. Industrial Crops and Products **189** (2022) https://doi.org/

10.1016/j.indcrop.2022.115848

[133] Morabit, M., Desaulniers, G., Lodi, A.: Machine-learning-based column selection for column generation. Transportation Science **55**(4), 815–831 (2021) https://doi.org/10.1287/trsc.2021.1045

[134] Morabit, M., Desaulniers, G., Lodi, A.: Learning to repeatedly solve routing problems. Networks **83**(3), 503–526 (2024) https://doi.org/10.1002/net.22200

[135] Marandi, F., Fatemi Ghomi, S.M.T.: Network configuration multi-factory scheduling with batch delivery: A learning-oriented simulated annealing approach. Computers and Industrial Engineering **132**, 293–310 (2019) https://doi.org/10.1016/j.cie.2019.04.032

[136] Ma, L., Hao, X., Zhou, W., He, Q., Zhang, R., Chen, L.: A hybrid neural combinatorial optimization framework assisted by automated algorithm design. Complex and Intelligent Systems (2024) https://doi.org/10.1007/s40747-024-01600-2

[137] Mohamed-Iliasse, M., Loubna, B., Abdelaziz, B.: Machine Learning in Supply Chain Management: A Systematic Literature Review. International Journal of Supply and Operations Management **9**(4), 398–416 (2022) https://doi.org/10.22034/ijsom.2021.109189.2279

[138] Maheshwari, S., Jain, P.K., Kotecha, K.: Route Optimization of Mobile Medical Unit with Reinforcement Learning. Sustainability (Switzerland) **15**(5) (2023) https://doi.org/10.3390/su15053937

[139] Mamoudan, M.M., Jafari, A., Mohammadnazari, Z., Nasiri, M.M., Yazdani, M.: Hybrid machine learning-metaheuristic model for sustainable agri-food production and supply chain planning under water scarcity. Resources, Environment and Sustainability **14** (2023) https://doi.org/10.1016/j.resenv.2023.100133

[140] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., *et al.*: Human-level control through deep reinforcement learning. nature **518**(7540), 529–533 (2015) https://doi.org/10.1038/nature14236

[141] Meidute-Kavaliauskiene, I., Taşkın, K., Ghorbani, S., Činčikaitė, R., Kačenauskaitė, R.: Reviewing the Applications of Neural Networks in Supply Chain: Exploring Research Propositions for Future Directions. Information (Switzerland) **13**(5) (2022) https://doi.org/10.3390/info13050261

[142] Moradi, B.: The new optimization algorithm for the vehicle routing problem with time windows using multi-objective discrete learnable evolution model. Soft Computing **24**(9), 6741–6769 (2020) https://doi.org/10.1007/s00500-019-04312-9

[143] Meisheri, H., Sultana, N.N., Baranwal, M., Baniwal, V., Nath, S., Verma, S., Ravindran, B., Khadilkar, H.: Scalable multi-product inventory control with lead time constraints using reinforcement learning. Neural Computing and Applications **34**(3), 1735–1757 (2022) https://doi.org/10.1007/s00521-021-06129-w

[144] Moghaddas, M., Tohidi, G.: A neurodynamic scheme to bi-level revenue-based centralized resource allocation models. Journal of Intelligent and Fuzzy Systems **37**(2), 2741–2756 (2019) https://doi.org/10.3233/JIFS-182953

[145] Mousa, M., Berg, D., Kotecha, N., Rio Chanona, E.A., Mowbray, M.: An analysis of multi-agent reinforcement learning for decentralized inventory control systems. Computers and Chemical Engineering **188** (2024) https://doi.org/10.1016/j.compchemeng.2024.108783

[146] Ma, N., Wang, Z., Ba, Z., Li, X., Yang, N., Yang, X., Zhang, H.: Hierarchical Reinforcement Learning for Crude Oil Supply Chain Scheduling. Algorithms **16**(7) (2023) https://doi.org/10.3390/a16070354

[147] Najjartabar Bisheh, M., Nasiri, G.R., Esmaeili, E., Davoudpour, H., Chang, S.I.: A new supply chain distribution network design for two classes of customers using transfer recurrent neural network. International Journal of System Assurance Engineering and Management **13**(5), 2604–2618 (2022) https://doi.org/10.1007/s13198-022-01670-w

[148] Novelan, M.S., Efendi, S., Sihombing, P., Mawengkang, H.: Vehicle Routing Problem Optimization with Machine Learning in Imbalanced Classification Vehicle Route Data. Eastern-European Journal of Enterprise Technologies **5**(3(125)), 49–56 (2023) https://doi.org/10.15587/1729-4061.2023.288280

[149] Nezamoddini, N., Gholami, A., Aqlan, F.: A risk-based optimization framework for integrated supply chains using genetic algorithm and artificial neural networks. International Journal of Production Economics **225** (2020) https://doi.org/10.1016/j.ijpe.2019.107569

[150] Nouicer, D., Msadaa, I.C., Grayaa, K.: A Novel Routing Solution for EV Fleets: A Real-World Case Study Leveraging Double DQNs and Graph-Structured Data to Solve the EVRPTW Problem. IEEE Access **11**, 122116–122127 (2023) https://doi.org/10.1109/ACCESS.2023.3327324

[151] Noh, J.: Reinforcement Learning for Optimizing Can-Order Policy with the Rolling Horizon Method. Systems **11**(7) (2023) https://doi.org/10.3390/systems11070350

[152] Neria, G., Tzur, M.: The Dynamic Pickup and Allocation with Fairness Problem. Transportation Science **58**(4), 821–840 (2024) https://doi.org/10.1287/trsc.2023.0228

[153] Nemhauser, G.L., Wolsey, L.A.: Integer and Combinatorial Optimization. John Wiley & Sons, New York (1988)

[154] Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (2006)

[155] Ochoa-Barragán, R., Munguía-López, A.D.C., Ponce-Ortega, J.M.: A hybrid machine learning-mathematical programming optimization approach for municipal solid waste management during the pandemic. Environment, Development and Sustainability **26**(7), 17653–17672 (2024) https://doi.org/10.1007/s10668-023-03354-2

[156] Ochoa-Barragán, R., Serrano-Arévalo, T.I., Pulido-Ocegueda, J.C., Cerda-Flores, S.C., Ramírez-Márquez, C., Nápoles-Rivera, F., Ponce-Ortega, J.M.: Sustainable lime production in Michoacan Mexico: An optimal and equitable approach with machine learning. Journal of Cleaner Production **442** (2024) https://doi.org/10.1016/j.jclepro.2024.141017

[157] Owen, S.H., Daskin, M.S.: Strategic facility location: A review. European Journal of Operational Research **111**(3), 423–447 (1998) https://doi.org/10.1016/S0377-2217(98)00186-6

[158] Oguntola, I.O., Ülkü, M.A., Saif, A., Engau, A.: On the value of shipment consolidation and machine learning techniques for the optimal design of a multimodal logistics network. INFOR **62**(1), 1–52 (2024) https://doi.org/10.1080/03155986.2023.2202079

[159] Pentico, D.W.: Assignment problems: A golden anniversary survey. European Journal of Operational Research **176**(2), 774–793 (2007) https://doi.org/10.1016/j.ejor.2005.09.014

[160] Phiboonbanakit, T., Horanont, T., Huynh, V.N., Supnithi, T.: A Hybrid Reinforcement Learning-Based Model for the Vehicle Routing Problem in Transportation Logistics. IEEE Access **9**, 163325–163347 (2021) https://doi.org/10.1109/ACCESS.2021.3131799

[161] Phillips, R.: Pricing and Revenue Optimization. Stanford University Press, Stanford, Calif (2005). https://books.google.com/books?id=InuQPrC6GtQC

[162] Perez, H.D., Hubbs, C.D., Li, C., Grossmann, I.E.: Algorithmic approaches to inventory management optimization. Processes **9**(1), 1–17 (2021) https://doi.org/10.3390/pr9010102

[163] Pinedo, M.L.: Scheduling: Theory, Algorithms, and Systems, 4th edn. Springer, New York (2012). https://doi.org/10.1007/978-1-4614-2361-4

[164] Pan, W., Liu, S.Q.: Deep reinforcement learning for the dynamic and uncertain vehicle routing problem. Applied Intelligence **53**(1), 405–422 (2023) https://doi.

org/10.1007/s10489-022-03456-w

[165] Pirayesh Neghab, D., Khayyati, S., Karaesmen, F.: An integrated data-driven method using deep learning for a newsvendor problem with unobservable features. European Journal of Operational Research **302**(2), 482–496 (2022) https://doi.org/10.1016/j.ejor.2021.12.047

[166] Papadimitriou, C.H., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover Publications, Englewood Cliffs (1998). https://doi.org/10.1109/TASSP.1984.1164450

[167] Punia, S., Singh, S.P., Madaan, J.K.: From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. Decision Support Systems **136** (2020) https://doi.org/10.1016/j.dss.2020.113340

[168] Qin, W., Zhuang, Z., Huang, Z., Huang, H.: A novel reinforcement learning-based hyper-heuristic for heterogeneous vehicle routing problem. Computers and Industrial Engineering **156** (2021) https://doi.org/10.1016/j.cie.2021.107252

[169] RamachandranPillai, R., Arock, M.: An Adaptive Spiking Neural P System for Solving Vehicle Routing Problems. Arabian Journal for Science and Engineering **45**(4), 2513–2529 (2020) https://doi.org/10.1007/s13369-019-04153-6

[170] Rolf, B., Beier, A., Jackson, I., Müller, M., Reggelin, T., Stuckenschmidt, H., Lang, S.: A review on unsupervised learning algorithms and applications in supply chain management. International Journal of Production Research (2024) https://doi.org/10.1080/00207543.2024.2390968

[171] Rizqi, Z.U., Chou, S.Y.: Neuroevolution reinforcement learning for multi-echelon inventory optimization with delivery options and uncertain discount. Engineering Applications of Artificial Intelligence **134** (2024) https://doi.org/10.1016/j.engappai.2024.108670

[172] Rodríguez-Esparza, E., Masegosa, A.D., Oliva, D., Onieva, E.: A new Hyper-heuristic based on Adaptive Simulated Annealing and Reinforcement Learning for the Capacitated Electric Vehicle Routing Problem. Expert Systems with Applications **252** (2024) https://doi.org/10.1016/j.eswa.2024.124197

[173] Ren, L., Fan, X., Cui, J., Shen, Z., Lv, Y., Xiong, G.: A Multi-Agent Reinforcement Learning Method With Route Recorders for Vehicle Routing in Supply Chain Management. IEEE Transactions on Intelligent Transportation Systems **23**(9), 16410–16420 (2022) https://doi.org/10.1109/TITS.2022.3150151

[174] Robinson La Rocca, C., Cordeau, J.F., Frejinger, E.: One-Shot Learning for MIPs with SOS1 Constraints. Operations Research Forum **5**(3) (2024) https://doi.org/10.1007/s43069-024-00336-6

[175] Ramakurthi, V.B., Manupati, V.K., Machado, J., Varela, L.: A hybrid multi-objective evolutionary algorithm-based semantic foundation for sustainable distributed manufacturing systems. Applied Sciences (Switzerland) **11**(14) (2021) https://doi.org/10.3390/app11146314

[176] Ramírez-Villamil, A., Montoya-Torres, J.R., Jaegler, A., Cuevas-Torres, J.M.: Reconfiguration of last-mile supply chain for parcel delivery using machine learning and routing optimization. Computers and Industrial Engineering **184** (2023) https://doi.org/10.1016/j.cie.2023.109604

[177] Schulman, J., *et al.*: Proximal policy optimization algorithms. In: arXiv Preprint arXiv:1707.06347 (2017). https://doi.org/10.48550/arXiv.1707.06347

[178] Sharifnia, S.M.E., Amrollahi Biyouki, S., Sawhney, R., Hwangbo, H.: Robust simulation optimization for supply chain problem under uncertainty via neural network metamodeling. Computers and Industrial Engineering **162** (2021) https://doi.org/10.1016/j.cie.2021.107693

[179] Sutton, R.S., Barto, A.G., *et al.*: Reinforcement Learning: An Introduction vol. 1. MIT Press, Cambridge (1998)

[180] Stanley, K.O., Clune, J., Lehman, J., Miikkulainen, R.: Designing Neural Networks Through Neuroevolution. Nature Machine Intelligence, London, UK (2019). https://doi.org/10.1038/s42256-018-0006-z

[181] Shiri, M., Fattahi, P., Sogandi, F.: Two-stage approach for COVID-19 vaccine supply chain network under uncertainty using the machine learning algorithms: A case study. Engineering Applications of Artificial Intelligence **135** (2024) https://doi.org/10.1016/j.engappai.2024.108837

[182] Sangiorgio, M., Guariso, G.: NN-based implicit stochastic optimization of multi-reservoir systems management. Water (Switzerland) **10**(3) (2018) https://doi.org/10.3390/w10030303

[183] Sui, Z., Gosavi, A., Lin, L.: A reinforcement learning approach for inventory replenishment in vendor-managed inventory systems with consignment inventory. EMJ - Engineering Management Journal **22**(4), 44–53 (2010) https://doi.org/10.1080/10429247.2010.11431878

[184] Shang, Y., Li, S.: Hybrid combinatorial remanufacturing strategy for medical equipment in the pandemic. Computers and Industrial Engineering **174** (2022) https://doi.org/10.1016/j.cie.2022.108811

[185] Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E.: Designing and Managing the Supply Chain: Concepts, Strategies, and Cases. McGraw-hill New York, ??? (1999)

[186] Soroka, A.G., Mikhelson, G.V., Mescheryakov, A.V., Gerasimov, S.V.: Smart Routes: A System for Development and Comparison of Algorithms for Solving Vehicle Routing Problems with Realistic Constraints. Automation and Remote Control **85**(3), 309–319 (2024) https://doi.org/10.1134/S0005117924030081

[187] Schuld, M., Petruccione, F.: Supervised learning with quantum computers. Springer (2018) https://doi.org/10.1007/978-3-319-96424-9

[188] Shahbazian, R., Pugliese, L.D.P., Guerriero, F., Macrina, G.: Integrating Machine Learning Into Vehicle Routing Problem: Methods and Applications. IEEE Access **12**, 93087–93115 (2024) https://doi.org/10.1109/ACCESS.2024.3422479

[189] Soeanu, A., Ray, S., Debbabi, M., Berger, J., Boukhtouta, A.: A learning based evolutionary algorithm for distributed multi-depot VRP. Frontiers in Artificial Intelligence and Applications **243**, 49–58 (2012) https://doi.org/10.3233/978-1-61499-105-2-49

[190] Stranieri, F., Stella, F., Kouki, C.: Performance of deep reinforcement learning algorithms in two-echelon inventory control systems. International Journal of Production Research **62**(17), 6211–6226 (2024) https://doi.org/10.1080/00207543.2024.2311180

[191] Schuld, M., Sinayskiy, I., Petruccione, F.: An introduction to quantum machine learning. Contemporary Physics **56**(2), 172–185 (2015) https://doi.org/10.1080/00107514.2014.964942

[192] Tan, Y., Gu, L., Xu, S., Li, M.: Supply Chain Inventory Management from the Perspective of "Cloud Supply Chain"—A Data Driven Approach. Mathematics **12**(4) (2024) https://doi.org/10.3390/math12040573

[193] Tarapata, Z., Kulas, W., Antkiewicz, R.: Machine Learning Algorithms for the Problem of Optimizing the Distribution of Parcels in Time-Dependent Networks: The Case Study. Archives of Transport **61**(1), 133–147 (2022) https://doi.org/10.5604/01.3001.0015.8269

[194] Tunga, H., Kar, S., Giri, D.: Multi-objective brainstorm optimization algorithm based on Bayesian inference learning automaton for solving CVRP with time windows. Applied Soft Computing **162** (2024) https://doi.org/10.1016/j.asoc.2024.111755

[195] Triantafyllou, N., Papathanasiou, M.M.: Deep learning enhanced mixed integer optimization: Learning to reduce model dimensionality. Computers and Chemical Engineering **187** (2024) https://doi.org/10.1016/j.compchemeng.2024.108725

[196] Teerasoponpong, S., Sopadang, A.: Decision support system for adaptive sourcing and inventory management in small- and medium-sized enterprises. Robotics and Computer-Integrated Manufacturing **73** (2022) https://doi.org/10.1016/j.rcim.2021.102226

[197] Toth, P., Vigo, D.: The Vehicle Routing Problem. SIAM, Philadelphia (2002)

[198] Tsolaki, K., Vafeiadis, T., Nizamis, A., Ioannidis, D., Tzovaras, D.: Utilizing machine learning on freight transportation and logistics applications: A review. ICT Express **9**(3), 284–295 (2023) https://doi.org/10.1016/j.icte.2022.02.001

[199] Vanvuchelen, N., De Boeck, K., Boute, R.N.: Cluster-based lateral transshipments for the Zambian health supply chain. European Journal of Operational Research **313**(1), 373–386 (2024) https://doi.org/10.1016/j.ejor.2023.08.005

[200] Hezewijk, L., Dellaert, N., Van Woensel, T., Gademann, N.: Using the proximal policy optimisation algorithm for solving the stochastic capacitated lot sizing problem. International Journal of Production Research **61**(6), 1955–1978 (2023) https://doi.org/10.1080/00207543.2022.2056540

[201] Wang, Q.: VARL: a variational autoencoder-based reinforcement learning Framework for vehicle routing problems. Applied Intelligence **52**(8), 8910–8923 (2022) https://doi.org/10.1007/s10489-021-02920-3

[202] Wu, C., Barnes, D.: A dynamic feedback model for partner selection in agile supply chains. International Journal of Operations and Production Management **32**(1), 79–103 (2012) https://doi.org/10.1108/01443571211195745

[203] Wang, Z., Bai, R., Khan, F., Özcan, E., Zhang, T.: Gase: graph attention sampling with edges fusion for solving vehicle routing problems. Memetic Computing **16**(3), 337–353 (2024) https://doi.org/10.1007/s12293-024-00428-0

[204] Weber, C.A., Current, J., Benton, W.C.: Vendor selection criteria and methods. European Journal of Operational Research **50**(1), 2–18 (1991) https://doi.org/10.1016/0377-2217(91)90033-R

[205] Watkins, C.J., Dayan, P.: Q-learning. Machine Learning **8**(3-4), 279–292 (1992) https://doi.org/10.1007/BF00992698

[206] Wu, G., Carvalho Servia, M.Á., Mowbray, M.: Distributional reinforcement learning for inventory management in multi-echelon supply chains. Digital Chemical Engineering **6** (2023) https://doi.org/10.1016/j.dche.2022.100073

[207] Wang, Q., Hao, Y.: Routing optimization with Monte Carlo Tree Search-based multi-agent reinforcement learning. Applied Intelligence **53**(21), 25881–25896 (2023) https://doi.org/10.1007/s10489-023-04881-1

[208] Wang, Q., Hao, Y., Zhang, J.: Generative inverse reinforcement learning for learning 2-opt heuristics without extrinsic rewards in routing problems. Journal of King Saud University - Computer and Information Sciences **35**(9) (2023) https://doi.org/10.1016/j.jksuci.2023.101787

[209] Winston, W.L.: Operations Research: Applications and Algorithms. Thomson/Brooks/Cole, Belmont (2004)

[210] Wang, T., Jeff Hong, L.: Large-Scale Inventory Optimization: A Recurrent Neural Networks–Inspired Simulation Approach. INFORMS Journal on Computing **35**(1), 196–215 (2023) https://doi.org/10.1287/IJOC.2022.1253

[211] Wang, F., Lin, L.: Spare parts supply chain network modeling based on a novel scale-free network and replenishment path optimization with Q learning. Computers and Industrial Engineering **157** (2021) https://doi.org/10.1016/j.cie.2021.107312

[212] Wang, K., Long, C., Ong, D.J., Zhang, J., Yuan, X.M.: Single-Site Perishable Inventory Management Under Uncertainties: A Deep Reinforcement Learning Approach. IEEE Transactions on Knowledge and Data Engineering **35**(10), 10807–10813 (2023) https://doi.org/10.1109/TKDE.2023.3250249

[213] Wang, Q., Lai, K.H., Tang, C.: Solving combinatorial optimization problems over graphs with BERT-Based Deep Reinforcement Learning. Information Sciences **619**, 930–946 (2023) https://doi.org/10.1016/j.ins.2022.11.073

[214] Wu, Y., Song, W., Cao, Z., Zhang, J., Lim, A.: Learning Improvement Heuristics for Solving Routing Problems. IEEE Transactions on Neural Networks and Learning Systems **33**(9), 5057–5069 (2022) https://doi.org/10.1109/TNNLS.2021.3068828

[215] Wang, L., Song, C., Sun, Y., Lu, C., Chen, Q.: A Neural Multi-Objective Capacitated Vehicle Routing Optimization Algorithm Based on Preference Adjustment. Electronics (Switzerland) **12**(19) (2023) https://doi.org/10.3390/electronics12194167

[216] Wang, Q., Tang, C.: Deep reinforcement learning for transportation network combinatorial optimization: A survey. Knowledge-Based Systems **233** (2021) https://doi.org/10.1016/j.knosys.2021.107526

[217] Wagner, H.M., Whitin, T.M.: Dynamic version of the economic lot size model. Management Science **5**(1), 89–96 (1958) https://doi.org/10.1287/mnsc.5.1.89

[218] Wang, C., Yu, Z., McAleer, S., Yu, T., Yang, Y.: ASP: Learn a Universal Neural Solver! IEEE Transactions on Pattern Analysis and Machine Intelligence **46**(6), 4102–4114 (2024) https://doi.org/10.1109/TPAMI.2024.3352096

[219] Wang, Q., Zhang, C., Tang, C.: Discovering Lin-Kernighan-Helsgaun heuristic for routing optimization using self-supervised reinforcement learning. Journal of King Saud University - Computer and Information Sciences **35**(8) (2023) https://doi.org/10.1016/j.jksuci.2023.101723

[220] Wang, J., Zheng, R., Wang, Z.: Supply Chain Optimization Strategy Research Based on Deep Learning Algorithm. Mobile Information Systems **2022** (2022) https://doi.org/10.1155/2022/9058490

[221] Xu, Y., Fang, M., Chen, L., Xu, G., Du, Y., Zhang, C.: Reinforcement Learning With Multiple Relational Attention for Solving Vehicle Routing Problems. IEEE Transactions on Cybernetics **52**(10), 11107–11120 (2022) https://doi.org/10.1109/TCYB.2021.3089179

[222] Xue, S.: An adaptive ant colony algorithm for crowdsourcing multi-depot vehicle routing problem with time windows. Sustainable Operations and Computers **4**, 62–75 (2023) https://doi.org/10.1016/j.susoc.2023.02.002

[223] Xu, R., Wunsch, D.: Survey of clustering algorithms. IEEE Transactions on Neural Networks **16**(3), 645–678 (2005) https://doi.org/10.1109/TNN.2005.845141

[224] Yilmaz, D., Büyüktahtakın, İ.E.: Learning optimal solutions via an lstm-optimization framework. Operations Research Forum **4**(2), 48 (2023) https://doi.org/10.1007/s43069-023-00224-5

[225] Yilmaz, D., Büyüktahtakın, İ.E.: A non-anticipative learning-optimization framework for solving multi-stage stochastic programs. Annals of Operations Research, 1–41 (2024) https://doi.org/10.1007/s10479-024-06100-7

[226] Yilmaz, D., Büyüktahtakın, İ.E.: An expandable machine learning-optimization framework to sequential decision-making. European Journal of Operational Research **314**(1), 280–296 (2024) https://doi.org/10.1016/j.ejor.2023.10.045

[227] Yang, J., Lau, H.C., Wang, H.: Optimization of customer service and driver dispatch areas for on-demand food delivery. Transportation Research Part C: Emerging Technologies **165** (2024) https://doi.org/10.1016/j.trc.2024.104653

[228] Yue, B., Ma, J., Shi, J., Yang, J.: A Deep Reinforcement Learning-Based Adaptive Search for Solving Time-Dependent Green Vehicle Routing Problem. IEEE Access **12**, 33400–33419 (2024) https://doi.org/10.1109/ACCESS.2024.3369474

[229] Yunusoglu, P., Ozsoydan, F.B., Bilgen, B.: A machine learning-based two-stage approach for the location of undesirable facilities in the biomass-to-bioenergy supply chain. Applied Energy **362** (2024) https://doi.org/10.1016/j.apenergy.2024.122961

[230] Yi, W., Qu, R., Jiao, L., Niu, B.: Automated Design of Metaheuristics Using Reinforcement Learning Within a Novel General Search Framework. IEEE Transactions on Evolutionary Computation **27**(4), 1072–1084 (2023) https://doi.org/10.1109/TEVC.2022.3197298

[231] Yang, B., Xu, X., Gong, Y., Rekik, Y.: Data-driven optimization models for inventory and financing decisions in online retailing platforms. Annals of Operations Research **339**(1-2), 741–764 (2024) https://doi.org/10.1007/s10479-023-05234-4

[232] Yu, J.J.Q., Yu, W., Gu, J.: Online Vehicle Routing with Neural Combinatorial Optimization and Deep Reinforcement Learning. IEEE Transactions on Intelligent Transportation Systems **20**(10), 3806–3817 (2019) https://doi.org/10.1109/TITS.2019.2909109

[233] Yu, L., Zhang, C., Jiang, J., Yang, H., Shang, H.: Reinforcement learning approach for resource allocation in humanitarian logistics. Expert Systems with Applications **173** (2021) https://doi.org/10.1016/j.eswa.2021.114663

[234] Zhang, X., Chen, L., Gendreau, M., Langevin, A.: Learning-Based Branch-and-Price Algorithms for the Vehicle Routing Problem with Time Windows and Two-Dimensional Loading Constraints. INFORMS Journal on Computing **34**(3), 1419–1436 (2022) https://doi.org/10.1287/ijoc.2021.1110

[235] Zheng, Y.J., Chen, X., Song, Q., Yang, J., Wang, L.: Evolutionary Optimization of COVID-19 Vaccine Distribution With Evolutionary Demands. IEEE Transactions on Evolutionary Computation **27**(1), 141–154 (2023) https://doi.org/10.1109/TEVC.2022.3164260

[236] Zou, Y., Hao, J.K., Wu, Q.: A reinforcement learning guided hybrid evolutionary algorithm for the latency location routing problem. Computers and Operations Research **170** (2024) https://doi.org/10.1016/j.cor.2024.106758

[237] Zipkin, P.H.: Foundations of Inventory Management. McGraw-Hill, Boston (2000)

[238] Zhou, Y., Li, W., Wang, X., Qiu, Y., Shen, W.: Adaptive gradient descent enabled ant colony optimization for routing problems. Swarm and Evolutionary Computation **70** (2022) https://doi.org/10.1016/j.swevo.2022.101046

[239] Zhao, J., Mao, M., Zhao, X., Zou, J.: A Hybrid of Deep Reinforcement Learning and Local Search for the Vehicle Routing Problems. IEEE Transactions on Intelligent Transportation Systems **22**(11), 7208–7218 (2021) https://doi.org/10.1109/TITS.2020.3003163

[240] Zwaida, T.A., Pham, C., Beauregard, Y.: Optimization of inventory management to prevent drug shortages in the hospital supply chain. Applied Sciences

(Switzerland) **11**(6) (2021) https://doi.org/10.3390/app11062726

[241] Zhang, Z., Qi, G., Guan, W.: Coordinated multi-agent hierarchical deep reinforcement learning to solve multi-trip vehicle routing problems with soft time windows. IET Intelligent Transport Systems **17**(10), 2034–2051 (2023) https://doi.org/10.1049/itr2.12394

[242] Zhang, J., Sun, Y., Feng, Q., Zhao, Y., Wang, Z.: Data-Driven Robust Optimization of the Vehicle Routing Problem with Uncertain Customers. Complexity **2022** (2022) https://doi.org/10.1155/2022/9064669

[243] Zavvar Sabegh, M.H., Mohammadi, M., Naderi, B.: Multi-objective optimization considering quality concepts in a green healthcare supply chain for natural disaster response: neural network approaches. International Journal of System Assurance Engineering and Management **8**, 1689–1703 (2017) https://doi.org/10.1007/s13198-017-0645-1

[244] Zhang, J., Tutun, S., Anvaryazdi, S.F., Amini, M., Sundaramoorthi, D., Sundaramoorthi, H.: Management of resource sharing in emergency response using data-driven analytics. Annals of Operations Research **339**(1-2), 663–692 (2024) https://doi.org/10.1007/s10479-023-05702-x

[245] Zhou, J., Wu, Y., Cao, Z., Song, W., Zhang, J., Chen, Z.: Learning Large Neighborhood Search for Vehicle Routing in Airport Ground Handling. IEEE Transactions on Knowledge and Data Engineering **35**(9), 9769–9782 (2023) https://doi.org/10.1109/TKDE.2023.3249799

[246] Zhang, C., Wu, Y., Ma, Y., Song, W., Le, Z., Cao, Z., Zhang, J.: A review on learning to solve combinatorial optimisation problems in manufacturing. IET Collaborative Intelligent Manufacturing **5**(1) (2023) https://doi.org/10.1049/cim2.12072

[247] Zhang, D., Wang, M., Mango, J., Li, X., Xu, X.: A survey on applications of reinforcement learning in spatial resource allocation. Computational Urban Science **4**(1) (2024) https://doi.org/10.1007/s43762-024-00127-z

[248] Zhang, Z., Wu, Z., Zhang, H., Wang, J.: Meta-Learning-Based Deep Reinforcement Learning for Multiobjective Optimization Problems. IEEE Transactions on Neural Networks and Learning Systems **34**(10), 7978–7991 (2023) https://doi.org/10.1109/TNNLS.2022.3148435

[249] Zhang, S., Xu, J., Qiao, Y.: Multi-Objective Q-Learning-Based Brain Storm Optimization for Integrated Distributed Flow Shop and Distribution Scheduling Problems. Mathematics **11**(20) (2023) https://doi.org/10.3390/math11204306

[250] Zhao, S., You, F.: Distributionally robust chance constrained programming with generative adversarial networks (GANs). AIChE Journal **66**(6) (2020) https:

//doi.org/10.1002/aic.16963

[251] Zhang, J., Yu, M., Feng, Q., Leng, L., Zhao, Y.: Data-Driven Robust Optimization for Solving the Heterogeneous Vehicle Routing Problem with Customer Demand Uncertainty. Complexity **2021** (2021) https://doi.org/10.1155/2021/6634132

[252] Zhang, R., Yu, R., Xia, W.: Constraint-aware Policy Optimization to Solve the Vehicle Routing Problem with Time Windows. Information Technology and Control **51**(1), 126–138 (2022) https://doi.org/10.5755/j01.itc.51.1.29924