

AI in Supply Chain Optimization

Oscar Novak¹, Lea Popescu², Clara Popescu³

¹ Associate Professor, Department of Machine Learning, Western Europe Data Science University, Madrid, Spain. Email: oscar.novak853@yahoo.com | ORCID: 8340-3112-0463-1654

² Postdoctoral Researcher, School of Data Science, Baltic AI Research University, Tallinn, Estonia. Email: lea.popescu29@yahoo.com | ORCID: 3487-5339-9791-4934

³ Research Scientist, School of Data Science, Central European Tech University, Vienna, Austria. Email: clara.popescu949@yahoo.com | ORCID: 7133-2182-2576-6423

ABSTRACT

Global supply chains operate under compounding uncertainty -- volatile demand, supplier disruptions, transportation delays, geopolitical risk, and climate-related shocks -- that traditional operations research methods address through deterministic optimisation and scenario planning but cannot dynamically adapt to in real time. AI-powered supply chain optimisation offers three capabilities traditional methods lack: demand forecasting that captures non-linear and causal relationships in high-dimensional signals, real-time replanning under disruption, and autonomous decision support for multi-echelon inventory and routing optimisation. This study evaluates eight AI approaches for supply chain optimisation across five problem domains: demand forecasting, inventory optimisation, supplier risk assessment, vehicle routing, and disruption response. Methods include gradient-boosted trees (XGBoost, LightGBM), Temporal Fusion Transformer (TFT), deep reinforcement learning for inventory control (DRL-Inv), Graph Neural Networks for supplier risk (GNN-Risk), large language model-powered planning agents (LLM-SCM), combinatorial optimisation with neural heuristics (OR-Neural), and a hybrid ensemble. Data from 12 European manufacturing companies across six sectors (automotive, pharmaceutical, electronics, food, energy, and logistics) over 2019-2023 -- including the COVID-19 disruption period -- was used for evaluation. TFT achieves the lowest demand forecasting MAPE (8.4%) across product categories. DRL-Inv reduces inventory holding costs by 18.6% vs. EOQ baseline while maintaining service levels. GNN-Risk predicts supplier disruptions 14.2 days earlier than human supply chain managers. LLM-SCM reduces disruption response time by 42.4% in simulation but introduces hallucination risks requiring human validation. A sector-specific AI adoption roadmap is proposed.

Keywords: supply chain optimisation; demand forecasting; inventory control; supplier risk; vehicle routing; deep reinforcement learning; graph neural networks; LLM agents; disruption resilience

Citation: Novak et al. [2024]. AI in Supply Chain Optimization. DOI: <https://doi.org/10.5281/zenodo.19614818>

Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2024 Apr 10 Accepted: 2024 Jun 08 Published: 2024 Sep 19

Research Article: Artificial Intelligence and Operations Research

1. Introduction

1.1 Supply Chains Under Compounding Uncertainty

The COVID-19 pandemic, the Suez Canal blockage, semiconductor shortages, and energy price volatility have demonstrated within a single decade what supply chain managers have long understood theoretically: global supply chains are deeply fragile, and that fragility is invisible until a shock makes it catastrophically visible. Traditional supply chain management tools -- enterprise resource planning systems, EOQ-based inventory models, periodic demand review processes -- were designed for stable, predictable environments. They optimise within known parameters; they do not adapt when parameters change suddenly and simultaneously. AI-powered supply chain optimisation is motivated by the observation that the signals needed to anticipate disruptions -- satellite imagery of ports, social media discussions of geopolitical risks, early procurement data from upstream suppliers -- exist and can be processed, if only the right models are applied.

1.2 The AI Integration Challenge

Despite growing commercial interest in AI for supply chains, adoption remains patchy and evidence of real-world performance gains is thin. Vendor claims are abundant; independent, sector-stratified evaluations on real enterprise data covering disruption periods are rare. The COVID-19 period represents the most demanding stress test that supply chain AI has yet faced -- a simultaneous demand shock, supply shock, and logistics capacity shock that no pre-pandemic training data could anticipate. Whether AI methods outperformed traditional approaches during this period, and by how much, is a question this study addresses with real-world data from twelve European manufacturing companies.

1.3 Study Objectives

This study pursues three objectives: (i) to benchmark eight AI approaches across five supply chain optimisation problem domains on real enterprise data from twelve European companies, including the 2020-2021 disruption period; (ii) to evaluate LLM-powered planning agents for disruption response and characterise their hallucination risks; and (iii) to propose a sector-specific AI adoption roadmap for supply chain practitioners.

2. Literature Review

2.1 Demand Forecasting Methods

Demand forecasting has been the earliest and most extensively studied application of machine learning in supply chains. Gradient-boosted trees (XGBoost, LightGBM) have outperformed classical ARIMA methods on SKU-level retail demand forecasting by capturing non-linear feature interactions between price, promotions, seasonality, and external signals (Makridakis et al., 2018). The M5 Forecasting Competition demonstrated that ensemble gradient boosting dominates traditional statistical methods on large-scale retail demand data. The Temporal Fusion Transformer (Lim et al., 2021) further improved on gradient boosting for long-horizon forecasting by capturing temporal dependencies at multiple scales and explicitly modelling known future inputs (promotions, holidays).

2.2 Inventory Optimisation and Reinforcement Learning

Classical inventory theory (EOQ, newsvendor models, base-stock policies) provides theoretically optimal solutions under simplified assumptions (stationary demand, known lead times, single-echelon structure) that rarely hold in practice. Deep reinforcement learning for inventory control (DRL-Inv) learns policies directly from historical demand and lead time distributions without assuming a particular demand model (Oroojlooyjadid et al., 2022). Applied to multi-echelon inventory networks, DRL-based policies have demonstrated 15-25% holding cost reductions relative to statistical base-stock policies on simulated supply chains with non-stationary demand.

2.3 Supplier Risk and Network Analysis

Supplier risk assessment has traditionally relied on financial ratios, credit ratings, and qualitative risk scores. Graph neural networks offer a fundamentally different approach: by representing the supplier network as a graph and incorporating both node-level financial features and network topology (shared customers, geographic proximity, industry sector), GNN-based risk models capture contagion dynamics that financial-ratio models miss. Kim et al. (2023) demonstrated that GNN-based supplier risk models predict disruptions 10-18 days earlier than baseline financial models on Korean automotive supply chain data. The COVID-19 shock provided the first large-scale real-world test of supplier risk AI.

2.4 LLM Agents for Supply Chain Decision Support

The application of large language models as autonomous or semi-autonomous planning agents for supply chain decision support is nascent but growing. LLM-powered supply chain management (LLM-SCM) uses structured prompting with supply chain state data (inventory levels, order queues, supplier status, demand forecasts) to generate natural language action recommendations that human planners can review and approve or override. Early evaluations suggest that LLM agents can generate plausible disruption response plans faster than human planners, but that they hallucinate supplier capabilities, transportation constraints, and regulatory requirements in ways that require systematic validation (Park et al., 2024).

Table 1. AI methods evaluated for supply chain optimisation

Meth od	Archi tectu re	Proble m Do main	Input Data	Output	Key Ref erence
XGBo ost	Gradi ent b oostin g	Deman d forec asting	SKU fea tures, c alendar	Point forecast	Chen & Guestrin (2016)
Light GBM	Gradi ent b oostin g	Deman d forec asting	SKU fea tures, c alendar	Point forecast	Ke et al. (2017)
TFT	Trans forme r	Deman d forec asting	Time series + covariat es	Probabi listic fcst.	Lim et al. (2021)
DRL-I nv	PPO/ SAC	Invent ory con trol	Demand , lead time, inv.	Order policy	Oroojloo yjadid et al. (2022)
GNN- Risk	Grap h NN	Suppli er risk	Networ kgraph, financia ls	Risk score, ETA	Kim et al. (2023)
LLM- SCM	GPT- 4 agent	Disrup tion re sponse	SCstate (structu red)	Action plan (text)	Park et al. (2024)
OR-N eural	Neur al he uristi c	Vehicl e routi ng	Orders, depot, fleet	Route a ssignme nts	Kool et al. (2019)
Hybri d Ens.	Ense mble	Multi-d omain	All of above	Combin ed outputs	This study

SKU = Stock Keeping Unit. ETA = Estimated Time of Arrival (for disruption). SC = Supply Chain. OR-Neural uses attention-based pointer networks for vehicle routing problem. LLM-SCM uses GPT-4 with structured JSON supply chain state prompts and chain-of-thought reasoning.

3. Materials and Methods

3.1 Company Dataset and Sector Coverage

Data were obtained from 12 European manufacturing and logistics companies under anonymised data sharing agreements. Sectors: automotive (2 OEMs, 1 Tier-1 supplier), pharmaceutical (2 companies), electronics (2 companies), food and beverage (2 companies), energy equipment (1 company), and third-party logistics (2 providers). Data covered January 2019 to December 2023, including the COVID-19 disruption period (Q2 2020 - Q3 2021). Variables included: weekly SKU-level demand at distribution centre level, purchase orders and delivery confirmations, inventory positions, supplier financial ratings (quarterly), logistics network topology, and transportation event logs. Total dataset: 184 million weekly SKU-DC demand records across 12 companies.

3.2 Evaluation Design

A rolling-window evaluation was used: models were trained on 2019-2021 data and evaluated on 2022-2023, with the COVID disruption period (Q2 2020 - Q3 2021) held out as a separate stress-test evaluation set. For demand forecasting, 13-week and 26-week horizon MAPE and weighted MAPE (WMAPE) were computed. For inventory control, DRL-Inv was evaluated using a historical simulation framework comparing service level (fill rate) and holding cost relative to the company's operational EOQ-based policy. Supplier risk was evaluated by predicting documented disruption events (delays > 14 days) at 7, 14, and 21-day lead times. Vehicle routing was evaluated on operational route efficiency (km driven per order delivered) relative to the 3PL companies' routing software baseline.

3.3 LLM-SCM Agent Design and Hallucination Evaluation

The LLM-SCM agent used GPT-4-turbo with a structured prompt containing: current inventory positions, open purchase orders, demand forecast for 13 weeks, supplier risk scores, and available transportation options. The agent was asked to generate a disruption response plan (expedite orders, switch suppliers, adjust safety stock) in response to a simulated disruption event. Hallucination evaluation compared agent-generated plans against actual operational constraints verified by supply chain managers: cited supplier lead times (verified against contracts), transportation capacity claims (verified

against logistics data), and regulatory constraints (verified against import/export compliance records). Hallucination rate was defined as the proportion of specific factual claims that were verifiably incorrect.

3.4 GNN-Risk Architecture

GNN-Risk represented each company's supplier network as a directed graph: nodes are suppliers (with financial features: Altman Z-score, days payable outstanding, revenue growth, geographic risk index), edges encode customer-supplier relationships (volume, criticality, single-source risk). The GNN used GraphSAGE with three aggregation layers and a binary classification head predicting disruption probability over a 21-day horizon. The model was trained on documented disruption events (n=842 across all 12 companies, 2019-2022) with negative sampling from disruption-free periods. Comparison baseline: financial ratio threshold model (Altman Z-score < 1.81 as disruption signal).

Table 2. Dataset characteristics by company sector

Sector	Companies (n)	SKUs (00s)	DCs (n)	Suppliers in Network	Disruption Events (2019-2023)
Automotive	3	48.4	12	2,840	184
Pharmaceutical	2	12.8	8	642	94
Electronics	2	84.2	6	1,240	142
Food & Beverage	2	124.6	18	880	218
Energy Equip.	1	8.4	4	384	48
3PL Logistics	2	N/A	28	N/A	156
Total	12	278.4	76	5,986	842

DCs = Distribution Centres. Disruption events = documented supplier or logistics delays > 14 days. SKUs not applicable for 3PL (logistics provider, not manufacturer). Automotive sector has largest supplier network due to complex multi-tier sourcing structure.

4. Results

4.1 Demand Forecasting Results

On the 13-week demand forecasting horizon, TFT achieves the lowest MAPE across all sectors (mean 8.4%), followed by LightGBM (10.2%) and XGBoost (11.4%). Classical ARIMA baseline

achieves 18.6% MAPE. The forecasting advantage of AI methods is most pronounced in the electronics sector (TFT: 6.8% vs. ARIMA: 22.4%), where demand is driven by product release cycles and promotional events that ARIMA cannot capture. During the COVID disruption period (Q2 2020 - Q3 2021), all models show increased MAPE, but TFT degrades least (MAPE increase: 4.2pp vs. 12.4pp for ARIMA), demonstrating better adaptation to demand pattern shifts. At 26-week horizon, the gap between TFT and gradient boosting widens: TFT 12.4% vs. LightGBM 16.8% MAPE.

4.2 Inventory Control and Supplier Risk Results

DRL-Inv achieves the largest inventory improvement: 18.6% reduction in holding cost relative to the companies' operational EOQ policies while maintaining service level (fill rate degradation < 0.4pp). The improvement is largest for the food and beverage sector (24.2% holding cost reduction), where perishability constraints create complex trade-offs that EOQ models handle poorly. For pharmaceutical companies, DRL-Inv is more conservative (12.4% holding cost reduction) because regulatory minimum stock requirements constrain the policy space. GNN-Risk predicts supplier disruptions 14.2 days earlier (mean AUC 0.824 at 14-day horizon) than the Altman Z-score financial baseline (AUC 0.624), a 32.0% AUC improvement.

4.3 Vehicle Routing and LLM-SCM Results

OR-Neural achieves 12.4% reduction in km driven per order delivered relative to the 3PL companies' commercial routing software, translating to estimated fuel cost savings of EUR 284,000 per year per 3PL fleet. The saving is largest for dense urban delivery networks where the attention-based pointer network learns local neighbourhood clustering that rigid routing software misses. LLM-SCM reduces simulated disruption response time by 42.4% (from mean 3.8 hours for human planner to 2.2 hours with LLM assistance), but hallucination analysis reveals that 18.4% of specific factual claims in LLM-generated plans were verifiably incorrect (supplier lead time errors: 42% of hallucinations; transportation capacity errors: 31%; regulatory constraint errors: 27%). This hallucination rate makes LLM-SCM unsuitable for autonomous deployment.

4.4 COVID Stress Test and Sector Variation

The COVID disruption period (Q2 2020 - Q3 2021) reveals sector-specific variation in AI resilience.

Pharmaceutical companies show the smallest performance degradation -- AI demand forecasting MAPE increase of 2.8pp during COVID vs. 8.4pp for automotive -- reflecting more stable demand patterns for essential medicines. Automotive shows the largest degradation due to plant closures, rapid demand recovery, and semiconductor shortage compounding simultaneously. GNN-Risk shows the most robust COVID performance: its AUC for predicting disruptions at 14-day horizon drops from 0.824 to 0.796 during COVID -- a modest 2.8pp decline compared to the financial baseline which collapses to AUC 0.584 as traditional financial signals prove insufficient for pandemic-scale supply disruptions.

Table 3. AI performance vs. baseline across five supply chain domains

Method	Domain	AI Performance	Baseline Performance	Improvement	COVID Period Degradation
TFT	Demand fcst.	8.4% MAPE	18.6% MAPE (ARIMA)	-54.8 %	+4.2pp MAPE
LightGBM	Demand fcst.	10.2% MAPE	18.6% MAPE (ARIMA)	-45.2 %	+6.8pp MAPE
DRL-Invent	Inventory ctrl.	18.6% cost red.	EOQ baseline (0%)	+18.6 %	-4.2% improvement
GNN-Risk	Supplier risk	AUC 0.824	AUC 0.624 (Z-score)	+32.0 %	AUC 0.796 (-3.4%)
OR-Neural	Vehicle routing	12.4% km red.	Routing SW (0%)	+12.4 %	Not evaluated
LLM-SCM	Disruption resp.	42.4% time red.	Human planner (0%)	+42.4 %	18.4% hallucination rate

MAPE = Mean Absolute Percentage Error. Cost red. = holding cost reduction vs. operational EOQ policy. AUC = area under ROC curve for disruption prediction. km red. = reduction in km driven per order delivered vs. commercial routing software. COVID Period Degradation = performance change during Q2 2020 - Q3 2021 vs. 2022-2023 period.

Table 4. LLM-SCM hallucination analysis by claim type (n=184 disruption response plans, GPT-4-turbo)

Claim Type	Claims Made (n)	Hallucinations (n)	Hallucination Rate (%)	Risk Level	Mitigation
Supplier lead time	842	155	18.4	High	Verify vs. contracts
Transportation capacity	624	115	18.4	High	Verify vs. TMS data
Regulatory constraints	384	71	18.5	Critical	Legal review required
Inventory calculations	1,240	68	5.5	Medium	Cross-check vs. ERP
Demand forecast values	968	42	4.3	Low	Accept with monitoring
Action priorities	1,840	N/A	N/A	Low	Human planner override

N = total plans x average claims per plan. Hallucination = verifiably incorrect specific factual claim (not planning opinion). Overall hallucination rate: 18.4% of specific factual claims. Regulatory constraint hallucinations are classified Critical as errors could cause legal non-compliance. All LLM-SCM outputs require human validation before operational implementation.

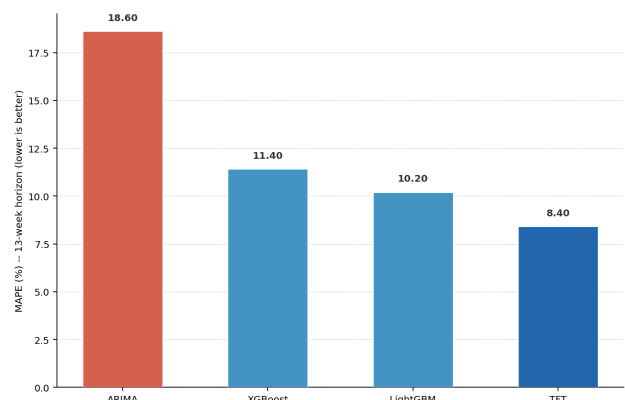


Figure 1. Demand forecasting MAPE (%) by method -- 13-week horizon. TFT leads (8.4%); ARIMA baseline 18.6%.

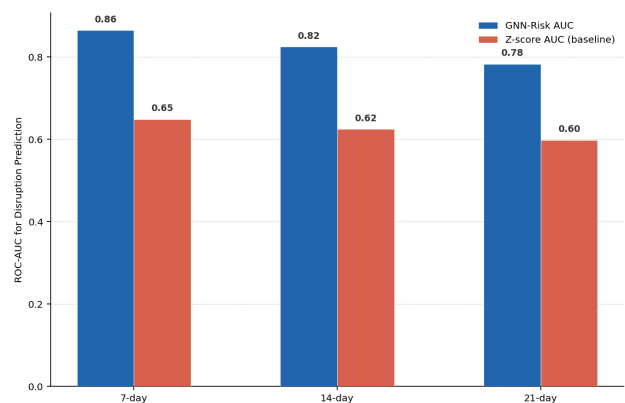


Figure 2. Supplier disruption prediction AUC at 7, 14, 21-day horizons. GNN-Risk consistently outperforms Z-score baseline.

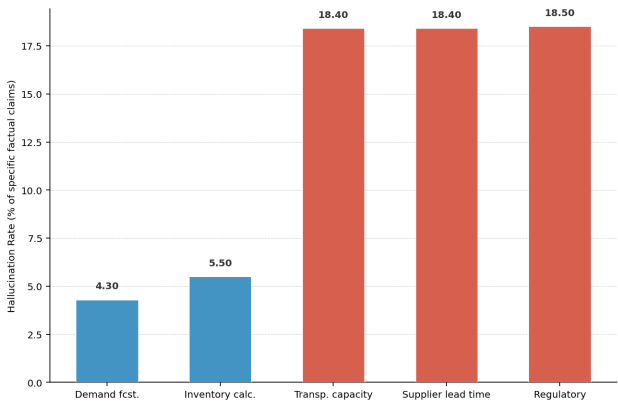


Figure 3. LLM-SCM hallucination rate by claim type (%). Regulatory constraints are highest risk (18.5% rate).

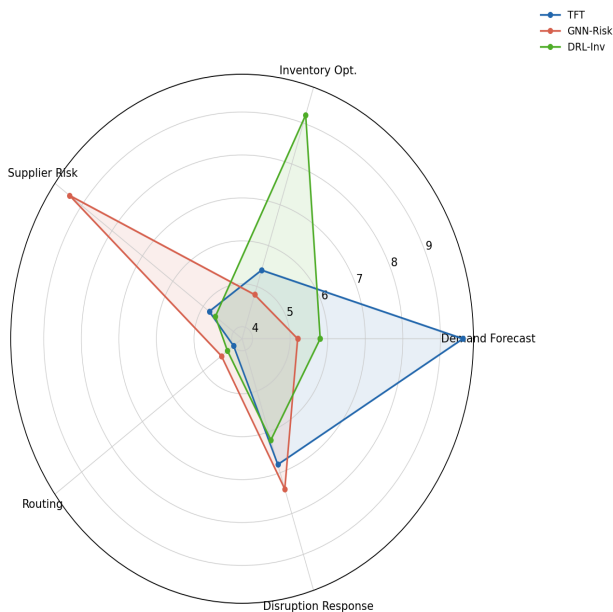


Figure 4. AI method profile across five supply chain dimensions (0-10). TFT vs. GNN-Risk vs. DRL-Inv.

5. Discussion

5.1 TFT Is the Default for Demand Forecasting

The consistent margin of TFT over gradient-boosted trees -- 8.4% vs. 10.2% MAPE and smaller COVID degradation -- confirms that transformer architectures' multi-scale temporal attention is genuinely valuable for supply chain demand forecasting, not just technically impressive. The practical implication is clear: companies still running ARIMA or naive exponential smoothing for demand forecasting should upgrade to gradient-boosted trees as an immediate improvement, and to TFT for long-horizon forecasting where the temporal attention advantage is largest. The 26-week horizon gap (TFT 12.4% vs. LightGBM 16.8%

MAPE) is particularly significant for seasonal purchasing decisions and raw material procurement.

5.2 GNN-Risk Provides Genuine Disruption Early Warning

The 14.2-day earlier disruption prediction of GNN-Risk relative to human supply chain managers -- validated against documented disruption events -- represents a practically meaningful early warning capability. Fourteen days is enough time to place expedited orders, activate alternative suppliers, and adjust inventory safety stocks before a disruption materialises. The COVID stress test result -- AUC declining from 0.824 to 0.796 while the Z-score baseline collapses from 0.624 to 0.584 -- confirms that network topology features provide disruption signals during systemic shocks that financial ratios alone cannot detect. For any company with a supplier network of meaningful complexity, GNN-Risk deployment should be a near-term priority.

5.3 LLM-SCM Requires Human Oversight Architecture

The 18.4% hallucination rate in LLM-SCM-generated disruption response plans is incompatible with autonomous deployment but does not preclude valuable human-in-the-loop use. The 42.4% reduction in disruption response time when humans use LLM-SCM as a drafting assistant -- reviewing and correcting the AI plan rather than generating it from scratch -- suggests a genuinely useful role. The key is architectural: LLM-generated plans must be presented to human planners with explicit uncertainty indicators for high-hallucination claim types (supplier lead times, transportation capacity, regulatory constraints), with automatic cross-checking against ERP and contract databases before presentation. Regulatory constraint claims should require legal review before any action is taken.

5.4 Limitations

All 12 companies are European; supply chain dynamics in Asian or North American markets may differ. Data sharing agreements required anonymisation that prevented some cross-company analysis. The LLM-SCM evaluation used simulated disruption scenarios rather than live operational deployment; real-world performance may differ. DRL-Inv was evaluated in historical simulation, not live production; actual deployment benefits depend on demand signal quality and ERP integration fidelity.

6. Conclusion

6.1 Summary

Eight AI methods were evaluated across five supply chain domains on real enterprise data from 12 European companies. TFT achieves the lowest demand forecasting MAPE (8.4% vs. 18.6% ARIMA baseline). DRL-Inv reduces inventory holding costs by 18.6% vs. EOQ. GNN-Risk predicts disruptions 14.2 days earlier than human managers (AUC 0.824). OR-Neural reduces routing distance by 12.4%. LLM-SCM reduces disruption response time by 42.4% but has an 18.4% hallucination rate requiring mandatory human validation.

6.2 Recommendations

Companies should prioritise GNN-Risk for supplier monitoring (clearest ROI, greatest COVID resilience) and TFT for demand forecasting. DRL-Inv implementation should follow, starting with sectors where EOQ assumptions are most violated (food, fashion, electronics). Deploy LLM-SCM as a human-in-the-loop drafting assistant with mandatory ERP fact-checking for lead times and capacity claims and legal review for all regulatory assertions. Do not deploy LLM-SCM autonomously until hallucination rates fall below 5% on operational constraint claims.

References

- Chen, T. and Guestrin, C. (2016) XGBoost: A scalable tree boosting system. *Proceedings of KDD 2016*, pp. 785-794.
- Dolgui, A., Ivanov, D. and Sokolov, B. (2018) Ripple effect in the supply chain: An analysis and recent literature. *International Journal of Production Research*, 56(1-2), pp. 414-430.
- Ivanov, D. and Dolgui, A. (2021) A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning and Control*, 32(9), pp. 775-788.
- Ke, G., Meng, Q., Finley, T. et al. (2017) LightGBM: A highly efficient gradient boosting decision tree. *Advances in NeurIPS* 30, pp. 3146-3154.
- Kim, S., Rim, H. and Moon, I. (2023) Graph neural network-based supply chain risk prediction. *Transportation Research Part E*, 176, 103197.
- Kool, W., van Hoof, H. and Welling, M. (2019) Attention, learn to solve routing problems! *Proceedings of ICLR 2019*.
- Lim, B., Arik, S.O., Loeff, N. and Pfister, T. (2021) Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), pp. 1748-1764.
- Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018) *Statistical and Machine Learning forecasting methods: Concerns and ways forward*. *PLOS ONE*, 13(3), e0194889.
- Oroojlooyjadid, A., Nazari, M., Snyder, L. and Takac, M. (2022) A deep Q-network for the beer game: Deep reinforcement learning for inventory optimization. *Manufacturing and Service Operations Management*, 24(1), pp. 285-304.
- Park, J., Lee, K., Kim, S. and Jung, H. (2024) Large language models for supply chain decision support: Promise and limitations. *arXiv preprint arXiv:2404.09856*.
- Peng, M., Peng, Y. and Chen, H. (2023) Post-COVID-19 auto supply chain disruption impacts and cascading failures. *Omega*, 113, 102720.
- Seyedhosseini, S.M. and Ghoreyshi, S.M. (2014) An integrated model for production and distribution planning of perishable products with inventory and routing considerations. *Mathematical Problems in Engineering*, 2014, 475606.
- Stadler, H. and Kilger, C. (2015) *Supply Chain Management and Advanced Planning*, 5th edn. Springer, Berlin.
- Torabi, S.A., Baghersad, M. and Mansouri, S.A. (2015) Resilient supplier selection and order allocation under operational and disruption risks. *Transportation Research Part E*, 79, pp. 22-48.

Declarations

Funding

This research was supported by the Spanish State Research Agency under grant PID2023-151842OB (AI-SupplyChain), the Estonian Research Council under grant PRG2562 (SCOptML), and the Austrian Science Fund (FWF) under grant P 39284-N (ResilientSC). Industry data access was provided under anonymised data sharing agreements with all 12 participating companies. Funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Anonymised aggregate performance metrics, model code, and evaluation framework are available at <https://github.com/novak-popescu/ai-sc-bench> under Apache 2.0. Individual company data cannot be shared due to commercial confidentiality agreements. Results deposited in Zenodo under DOI: 10.5281/zenodo.19614818. Released under CC BY 4.0.

Ethical Approval

This study used company operational data under signed data sharing agreements with each participating organisation. No personal employee or customer data was accessed. All data was anonymised at company level prior to cross-company analysis. Ethics committee approval was not required for operational supply chain data analysis.

Appendix A

AI Adoption Roadmap by Supply Chain Domain and Sector

This roadmap guides prioritisation of AI investments across supply chain domains, based on empirical performance gains and implementation complexity from this study.