Asian Business Research Journal
Vol. 11, No. 3, 21-29, 2026
ISSN: 2576-6759
DOI: 10.55220/2576-6759.v11i3.906

Deep Learning Methods for Demand Forecasting and Inventory Optimization in Modern Supply Chains

Pan Li[1] ✉
Jingyi Liu[2]
Lei Qiu[3]

[1]*University of Hull, United Kingdom.*
[2]*Cornell University, United States.*
[3]*Ningbo University of Technology, China.*
(✉ Corresponding Author)

Abstract

Modern supply chain management faces unprecedented challenges in demand forecasting and inventory optimization due to increasing market volatility, consumer behavior complexity, and global disruptions. Deep learning (DL) has emerged as a transformative approach that addresses these challenges by capturing complex nonlinear patterns in demand data and optimizing inventory decisions across multiple echelons. This review examines the current state of DL methods applied to demand forecasting and inventory optimization in supply chains. Recurrent neural networks (RNNs), long short-term memory (LSTM) networks, convolutional neural networks (CNNs), and transformer-based architectures have demonstrated superior performance compared to traditional statistical methods. The integration of DL with reinforcement learning (RL) has enabled adaptive inventory policies that respond dynamically to changing market conditions. Graph neural networks (GNNs) have proven effective in capturing network dependencies across complex supply chain structures. Despite these advances, challenges remain in model interpretability, data quality requirements, computational complexity, and real-time implementation. This paper provides a comprehensive analysis of DL architectures, hybrid approaches, performance metrics, and practical applications while identifying critical research gaps and future directions for advancing intelligent supply chain management systems.

Keywords: Convolutional neural networks, Deep learning, Demand forecasting, Graph neural networks, Inventory optimization, Long short-term memory, Recurrent neural networks, Reinforcement learning, Supply chain management, Transformers.

1. Introduction

Supply chain management has evolved from traditional rule-based systems to sophisticated data-driven approaches that leverage advanced computational methods for decision-making. The complexity of modern supply chains, characterized by multi-echelon networks, uncertain demand patterns, variable lead times, and dynamic market conditions, necessitates forecasting and optimization methods that can capture intricate relationships and adapt to rapid changes [1]. Traditional approaches including autoregressive integrated moving average (ARIMA) models, exponential smoothing, and classical optimization techniques often struggle with the nonlinear patterns, high-dimensional feature spaces, and complex interdependencies inherent in contemporary supply chain data [2]. The emergence of deep learning (DL) as a powerful paradigm in machine learning has opened new possibilities for addressing these limitations through its capacity to automatically learn hierarchical representations from raw data without extensive feature engineering [3].

Demand forecasting serves as the foundation for effective supply chain planning, influencing decisions across procurement, production scheduling, warehouse management, and distribution logistics [4]. Accurate forecasts enable organizations to reduce inventory holding costs, minimize stockouts, improve customer service levels, and optimize working capital allocation [5]. However, demand patterns in modern markets exhibit increasing volatility driven by factors such as seasonal variations, promotional activities, competitor actions, economic fluctuations, and shifting consumer preferences [6]. Traditional statistical methods assume linear relationships and stationary patterns that rarely reflect real-world demand behavior, particularly in industries with short product lifecycles, rapid innovation cycles, and highly differentiated consumer segments. DL architectures, particularly recurrent neural networks (RNNs) and their advanced variants such as long short-term memory (LSTM) networks and gated recurrent units (GRUs), have demonstrated remarkable capability in modeling temporal dependencies and capturing long-range patterns in sequential demand data [7].

Inventory optimization represents another critical challenge where DL methods have shown substantial promise. The inventory management problem involves determining optimal order quantities, reorder points, safety stock levels, and replenishment policies while balancing conflicting objectives of minimizing costs and maintaining service levels [8]. Classical approaches such as the economic order quantity (EOQ) model, newsvendor problem, and dynamic programming methods rely on simplifying assumptions about demand distributions, cost structures, and system dynamics that often fail to capture the complexity of real supply chains [9]. Reinforcement learning (RL) combined with DL function approximators enables the development of adaptive inventory policies that learn optimal actions through interaction with the environment, without requiring explicit mathematical models of system dynamics [10]. These methods can optimize policies across multiple products, locations, and time horizons while accounting for complex constraints and stochastic variations [11].

The integration of DL with inventory optimization extends beyond single-echelon systems to address multi-echelon supply chain networks where inventory decisions at one stage affect downstream and upstream operations [12]. Graph neural networks (GNNs) have emerged as particularly effective architectures for modeling these network dependencies, as they can process graph-structured data representing the relationships between suppliers, manufacturers, distribution centers, and retailers [13]. By learning node representations that capture both local features and global network structure, GNNs enable coordinated inventory decisions that optimize system-wide performance rather than isolated local objectives. The combination of temporal modeling through LSTM networks and spatial modeling through GNNs provides a comprehensive framework for spatio-temporal demand forecasting and inventory optimization in complex supply chain networks [14].

This review provides a comprehensive examination of DL methods for demand forecasting and inventory optimization in supply chains, synthesizing recent advances, analyzing methodological approaches, evaluating performance comparisons, and identifying future research directions. The paper systematically covers fundamental DL architectures including RNNs, LSTMs, convolutional neural networks (CNNs), and transformers, explores hybrid approaches combining multiple architectures, examines RL methods for inventory optimization, discusses GNN applications for network-level decisions, and addresses practical implementation challenges. By consolidating knowledge across these domains, this review aims to guide researchers and practitioners in selecting appropriate methods, understanding current capabilities and limitations, and identifying promising directions for advancing intelligent supply chain management systems.

2. Literature Review

The application of DL to supply chain demand forecasting has grown substantially over the past five years, with researchers demonstrating consistent performance improvements over traditional statistical methods across diverse industry contexts. Early studies established the fundamental advantages of LSTM networks for demand forecasting by showing their ability to capture complex temporal dependencies that classical time series models could not effectively model [15]. These initial investigations typically compared LSTM performance against ARIMA, exponential smoothing, and simple neural network baselines, consistently finding superior accuracy metrics such as mean absolute percentage error (MAPE) and root mean squared error (RMSE) reductions ranging from fifteen to forty percent [16]. The success of LSTM architectures motivated extensive research into architectural variations, including bidirectional LSTM networks that process sequences in both forward and backward directions to capture future context, and encoder-decoder LSTM structures that enable multi-step ahead forecasting through sequence-to-sequence learning [17].

CNNs, traditionally associated with image processing tasks, have found innovative applications in demand forecasting through their ability to extract local patterns from temporal data [18]. Researchers discovered that one-dimensional convolutions applied to time series can effectively capture short-term fluctuations and seasonal patterns by learning filters that detect recurring motifs in demand sequences [19]. Several studies have demonstrated that CNN architectures achieve comparable or superior performance to LSTM networks while requiring significantly less computational time for training and inference, making them attractive for large-scale deployments [20]. Hybrid CNN-LSTM architectures that combine convolutional feature extraction with recurrent temporal modeling have emerged as particularly effective approaches, with the CNN layers learning to identify relevant patterns and the LSTM layers modeling their evolution over time [21]. These hybrid models have shown robust performance across products with varying demand characteristics, from fast-moving consumer goods with high-frequency sales to slow-moving items with intermittent demand patterns [22].

The introduction of attention mechanisms marked a significant advancement in DL-based demand forecasting by enabling models to dynamically focus on relevant historical periods when making predictions [23]. Temporal attention mechanisms assign different weights to different time steps in the input sequence, allowing the model to emphasize periods with similar patterns to the forecast horizon while down-weighting less relevant historical data [24]. Studies have shown that attention-based models provide interpretable insights into forecast generation by visualizing which historical periods most strongly influence each prediction, addressing the black-box criticism often directed at DL methods [25]. The development of transformer architectures specifically designed for time series forecasting has further advanced this direction, with models such as Temporal Fusion Transformers demonstrating state-of-the-art performance across multiple forecasting benchmarks [26]. These architectures incorporate variable selection networks that automatically identify relevant features from high-dimensional input spaces, static covariate encoders that process time-invariant product attributes, and multi-horizon prediction heads that generate forecasts across different time scales simultaneously [27].

RL has emerged as a powerful framework for inventory optimization by framing the problem as a sequential decision-making task where agents learn optimal policies through trial and error [28]. Early applications of RL to inventory management focused on single-product, single-location scenarios using tabular methods such as Q-learning to determine optimal order quantities based on observed demand realizations [29]. The integration of DL function approximators, particularly deep Q-networks (DQN), enabled scaling to high-dimensional state spaces and continuous action spaces that characterize realistic inventory systems [30]. Researchers have demonstrated that deep RL agents can learn sophisticated replenishment policies that outperform classical inventory control rules such as base-stock policies and order-up-to policies, particularly in environments with non-stationary demand, correlated products, and complex cost structures [31]. Actor-critic methods, which maintain separate value and policy networks, have proven effective for inventory optimization by enabling stable learning in continuous action spaces while providing variance reduction through baseline subtraction [32].

The extension of RL methods to multi-echelon inventory systems represents a significant research frontier where DL approaches offer substantial advantages over traditional optimization techniques [33]. Coordinating inventory decisions across multiple stages of a supply chain involves managing interdependencies between facilities, balancing local and system-wide objectives, and handling information delays and uncertainties [34]. Graph-based RL approaches that represent supply chain networks as graphs with facilities as nodes and material flows as edges enable learning of coordinated policies that account for network structure [35]. Studies have shown that multi-agent RL frameworks where each facility operates an independent agent that learns to cooperate with other agents can achieve near-optimal system performance without requiring centralized control [36]. Communication mechanisms between agents, implemented through neural network modules that exchange learned representations, facilitate coordination and improve overall supply chain efficiency.

GNNs have revolutionized the modeling of supply chain networks by providing architectures specifically designed to process graph-structured data [37]. Unlike traditional neural networks that assume Euclidean input structures, GNNs operate on arbitrary graph topologies, making them naturally suited for supply chains where relationships between entities form complex networks [38]. Message-passing neural networks, a broad class of GNN architectures, iteratively update node representations by aggregating information from neighboring nodes, enabling the capture of both local features and global network structure [39]. Applications of GNNs to demand forecasting in supply chain networks have demonstrated that incorporating network structure as an inductive bias improves forecast accuracy by leveraging correlations between connected facilities [40]. For example, demand at retail locations may be correlated due to proximity, shared customer bases, or similar demographic characteristics, and GNNs can learn these relationships directly from data.

Recent research has explored the integration of external factors and multimodal data sources into DL forecasting models to improve accuracy and robustness [41]. Price information, promotional schedules, competitor activities, weather data, economic indicators, and social media sentiment all provide potentially valuable signals for demand prediction [42]. However, effectively incorporating these diverse data types requires careful feature engineering and model design to avoid overfitting and ensure generalization. Multi-task learning approaches that jointly train forecasting models on related tasks, such as predicting both demand and inventory levels, have shown promise for improving sample efficiency and capturing shared patterns across objectives [43]. The challenge of forecasting intermittent demand, characterized by periods with zero sales interspersed with occasional non-zero demand, has motivated specialized DL architectures [44]. Traditional point forecasting methods perform poorly on intermittent series because they struggle to predict both the timing and magnitude of demand occurrences. Researchers have proposed two-stage models that separately predict demand occurrence probability and conditional demand magnitude, combining these predictions to generate final forecasts [45].

3. Deep Learning Architectures for Demand Forecasting

LSTM networks have established themselves as the dominant architecture for sequential demand forecasting due to their sophisticated gating mechanisms that selectively retain or discard information across time steps [46]. As shown in Figure 1, the LSTM cell architecture consists of three gates, namely input gates that control how much new information enters the cell state, forget gates that determine what information should be discarded from the cell state, and output gates that regulate what information flows to the next time step. This gating structure enables LSTMs to maintain long-term dependencies without suffering from vanishing gradient problems that plague traditional RNNs, making them particularly effective for products with seasonal patterns that span multiple periods or long-term trends that evolve gradually over extended horizons [47]. Practical implementations of LSTM-based forecasting systems typically employ multi-layer architectures where lower layers learn basic temporal patterns and higher layers capture more abstract relationships, with layer depths ranging from two to five depending on data complexity and available training samples.

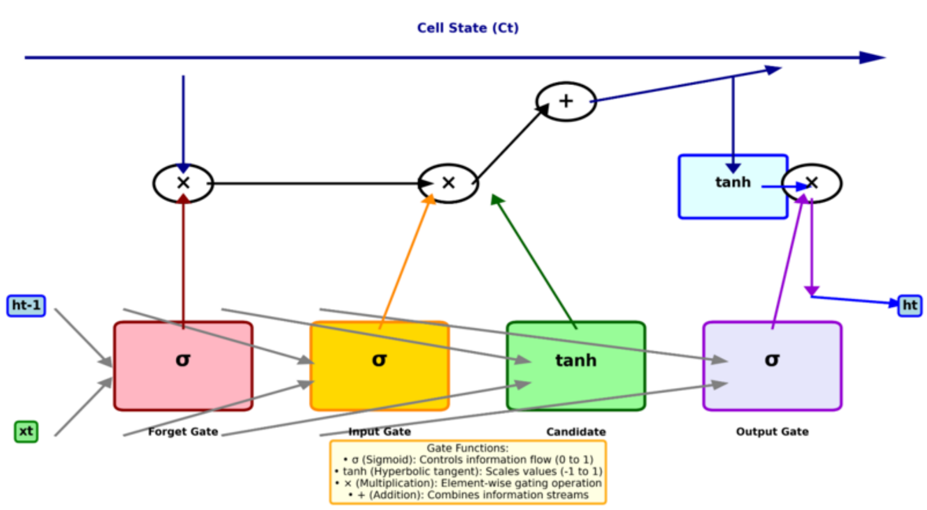

Figure 1. LSTM Cell Architecture with Gate Mechanisms.

Figure 1. LSTM cell architecture diagram showing input gate, forget gate, output gate, cell state flow, and hidden state propagation with mathematical operations at each gate

The training of LSTM networks for demand forecasting requires careful consideration of sequence length, batch size, and optimization algorithms to achieve stable convergence and good generalization performance. Backpropagation through time (BPTT) serves as the primary training method, where gradients are computed by unrolling the recurrent network across time steps and applying the chain rule [48]. However, BPTT can be computationally expensive for very long sequences, leading practitioners to employ truncated BPTT where gradient computation is limited to a fixed number of time steps. The choice of sequence length represents a critical hyperparameter that balances the model's ability to capture long-term dependencies against computational efficiency and overfitting risks. Research has shown that for most demand forecasting applications, sequence lengths between twelve and fifty-two time steps provide effective performance, corresponding to one year of weekly data or multiple months of daily data depending on the forecasting granularity [49].

Bidirectional LSTM architectures process input sequences in both forward and backward directions, enabling the model to incorporate future context when making predictions at each time step. While this approach is not directly applicable to real-time forecasting where future information is unavailable, bidirectional LSTMs prove valuable for retrospective analysis, pattern discovery, and training models that can then be deployed in a forward-only mode for actual predictions. The bidirectional structure effectively doubles the number of parameters compared to unidirectional LSTMs, requiring more training data to avoid overfitting but offering improved representation capacity for complex demand patterns. Encoder-decoder LSTM architectures extend the basic LSTM framework to enable multi-step ahead forecasting through a sequence-to-sequence learning paradigm. The encoder network processes the input historical sequence to produce a fixed-length context vector that captures relevant information, while the decoder network uses this context to generate a sequence of future predictions.

Transformer architectures have recently gained prominence in demand forecasting applications by addressing some limitations of recurrent networks through self-attention mechanisms [50]. Unlike LSTMs that process sequences sequentially, transformers can attend to all positions in the input sequence simultaneously, enabling parallel computation and more effective capture of long-range dependencies. The self-attention mechanism computes attention weights that determine how much each time step should contribute to the representation of every other time step, allowing the model to focus on relevant historical periods regardless of their distance in the sequence. Multi-head attention extends this concept by learning multiple independent attention functions, each potentially capturing different aspects of demand patterns such as trend components, seasonal cycles, and irregular fluctuations [51]. Positional encodings provide transformers with information about the relative or absolute position of time steps in the sequence, compensating for the lack of inherent sequential processing that characterizes recurrent architectures.

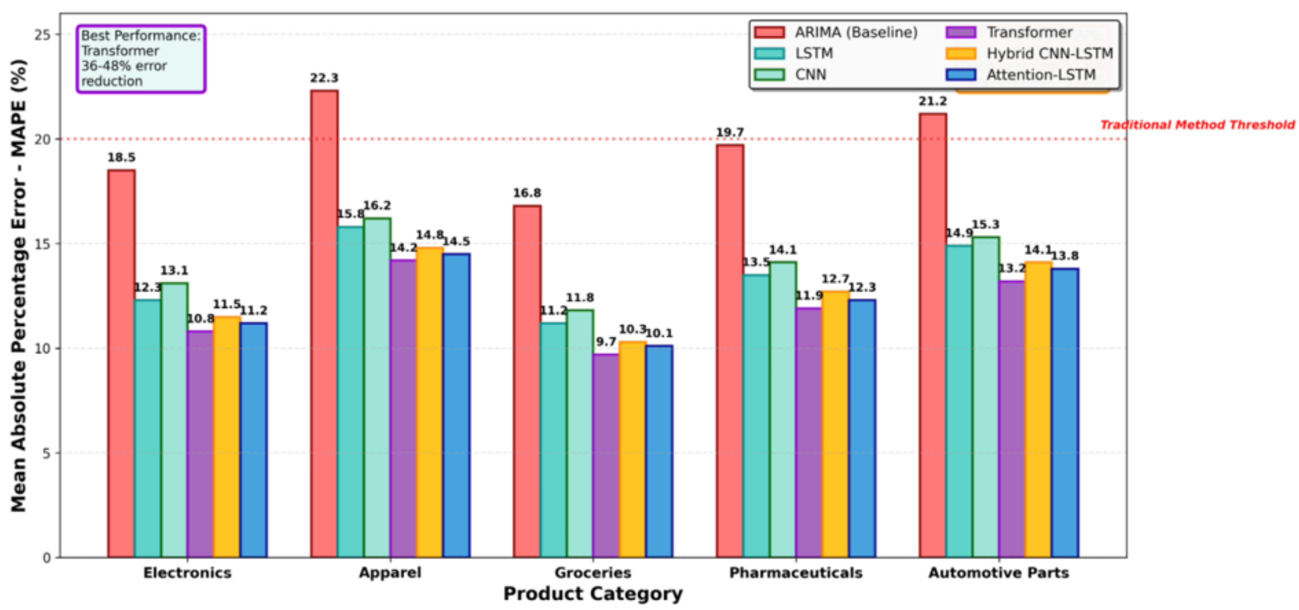

Figure 2. Forecasting Accuracy Comparison Across Deep Learning Architectures (Lower MAPE = Better Performance).

Figure 2. Comparison chart showing forecasting accuracy across different deep learning architectures including LSTM, CNN, Transformer, and hybrid models for various product categories including electronics, apparel, and groceries

CNN architectures applied to demand forecasting leverage convolutional operations to extract local patterns from temporal sequences through learned filters [52]. One-dimensional convolutions slide across time series data, computing weighted sums of consecutive time steps to detect recurring motifs and short-term fluctuations. Multiple convolutional layers with different filter sizes enable the model to capture patterns at various temporal scales, from daily variations to weekly cycles and monthly trends. Pooling layers following convolutional layers reduce the dimensionality of learned representations while maintaining the most salient features, improving computational efficiency and reducing overfitting risks. The hierarchical feature extraction process in CNNs bears similarity to the multi-layer structure in LSTMs, but CNNs typically require fewer parameters and enable faster training through parallel computation of convolutions. Dilated convolutions extend the receptive field of CNN filters without increasing the number of parameters by introducing gaps between filter elements, allowing the network to capture long-range dependencies more efficiently than standard convolutions.

As shown in Figure 2, hybrid architectures combining CNNs and LSTMs have demonstrated superior performance by leveraging the complementary strengths of both approaches [53]. A common design pattern employs CNN layers at the input stage to extract local temporal features, followed by LSTM layers that model the evolution of these features over time. This architecture enables the CNN component to learn robust representations of short-term demand patterns while the LSTM component captures how these patterns change and interact across longer time horizons. Alternative hybrid designs use parallel CNN and LSTM branches that process the input sequence independently, with their outputs concatenated and passed through fully connected layers for final prediction. Attention mechanisms integrated into hybrid architectures provide additional flexibility by allowing the model to dynamically weight the contributions of CNN-extracted features and LSTM-captured dependencies based on the specific forecasting context.

The incorporation of external variables and multivariate inputs significantly enhances the forecasting capability of DL architectures beyond univariate time series modeling [54]. Price elasticity effects, promotional events, holiday indicators, weather conditions, and competitor activities all influence demand patterns and can be integrated into DL models through various mechanisms. Feature fusion approaches concatenate external variables with learned temporal representations at appropriate layers in the network architecture, enabling the model to learn joint representations that capture interactions between temporal patterns and exogenous factors. Attention-based variable selection mechanisms automatically determine the relevance of different input features for each prediction, improving model interpretability and preventing the inclusion of noisy or irrelevant variables from degrading performance. Separate encoding pathways for different variable types, such as dedicated encoders for categorical features like product categories and continuous features like prices, enable the model to process heterogeneous inputs effectively before combining them for forecasting.

4. Reinforcement Learning for Inventory Optimization

RL provides a natural framework for inventory optimization by formulating the problem as a Markov decision process (MDP) where an agent learns to select actions that maximize cumulative rewards through interaction with the environment [55]. In the inventory management context, states represent the current inventory level, demand history, lead time information, and other relevant system variables. Actions correspond to replenishment decisions such as order quantities or whether to place an order. Rewards encode the trade-off between holding costs, ordering costs, and shortage penalties, with the agent's objective being to learn a policy that minimizes total expected costs over a planning horizon. The MDP formulation naturally accommodates stochastic demand, uncertain lead times, and complex cost structures that challenge classical inventory models, making RL particularly suitable for real-world supply chain environments characterized by uncertainty and dynamics.

Deep Q-networks (DQN) extend traditional Q-learning to high-dimensional state spaces by using deep neural networks to approximate the action-value function [56]. The DQN algorithm maintains a neural network that takes the current state as input and outputs estimated Q-values for all possible actions, with the optimal action being the one with the highest Q-value. Experience replay, a key component of DQN, stores transitions in a replay buffer and samples mini-batches randomly for training, breaking temporal correlations in the data and improving sample efficiency. Target networks provide stable learning targets by maintaining a slowly updated copy of the Q-network, preventing the instability that can arise from using constantly changing Q-value estimates as training targets. For inventory optimization, DQN has been successfully applied to determine optimal order-up-to levels in systems with non-stationary demand, learning policies that adapt to changing patterns without requiring explicit demand distribution assumptions [57].

Policy gradient methods offer an alternative approach to value-based RL by directly parameterizing and optimizing the policy function [58]. Actor-critic algorithms combine policy gradients with value function approximation, maintaining separate networks for the policy (actor) and value function (critic). The actor network selects actions based on the current state, while the critic network evaluates these actions by estimating their expected return. This separation enables more stable learning compared to pure policy gradient methods, as the critic provides a learned baseline that reduces the variance of gradient estimates. Proximal policy optimization (PPO) and soft actor-critic (SAC) represent advanced actor-critic algorithms that have shown excellent performance in continuous control tasks and translate naturally to inventory optimization problems with continuous order quantity decisions. These algorithms incorporate trust region constraints or entropy regularization to prevent destructive policy updates that can destabilize learning.

Multi-agent RL extends single-agent approaches to scenarios where multiple decision-makers operate simultaneously, each potentially affecting others through their actions [59]. In supply chain contexts, different facilities such as manufacturers, warehouses, and retailers may each operate an agent that makes local inventory decisions while sharing a common supply network. Cooperative multi-agent RL aims to learn coordinated policies that optimize system-wide performance rather than individual facility objectives. Centralized training with decentralized execution represents a popular paradigm where agents share information during training to learn coordinated behaviors but execute independently during deployment using only local observations. Communication protocols between agents enable information sharing about local states, intended actions, or learned representations, facilitating coordination without requiring central control. Graph attention networks (GATs) provide effective architectures for processing communication between agents in supply chain networks, as they can learn to weight messages from different neighbors based on their relevance to each agent's decision-making.

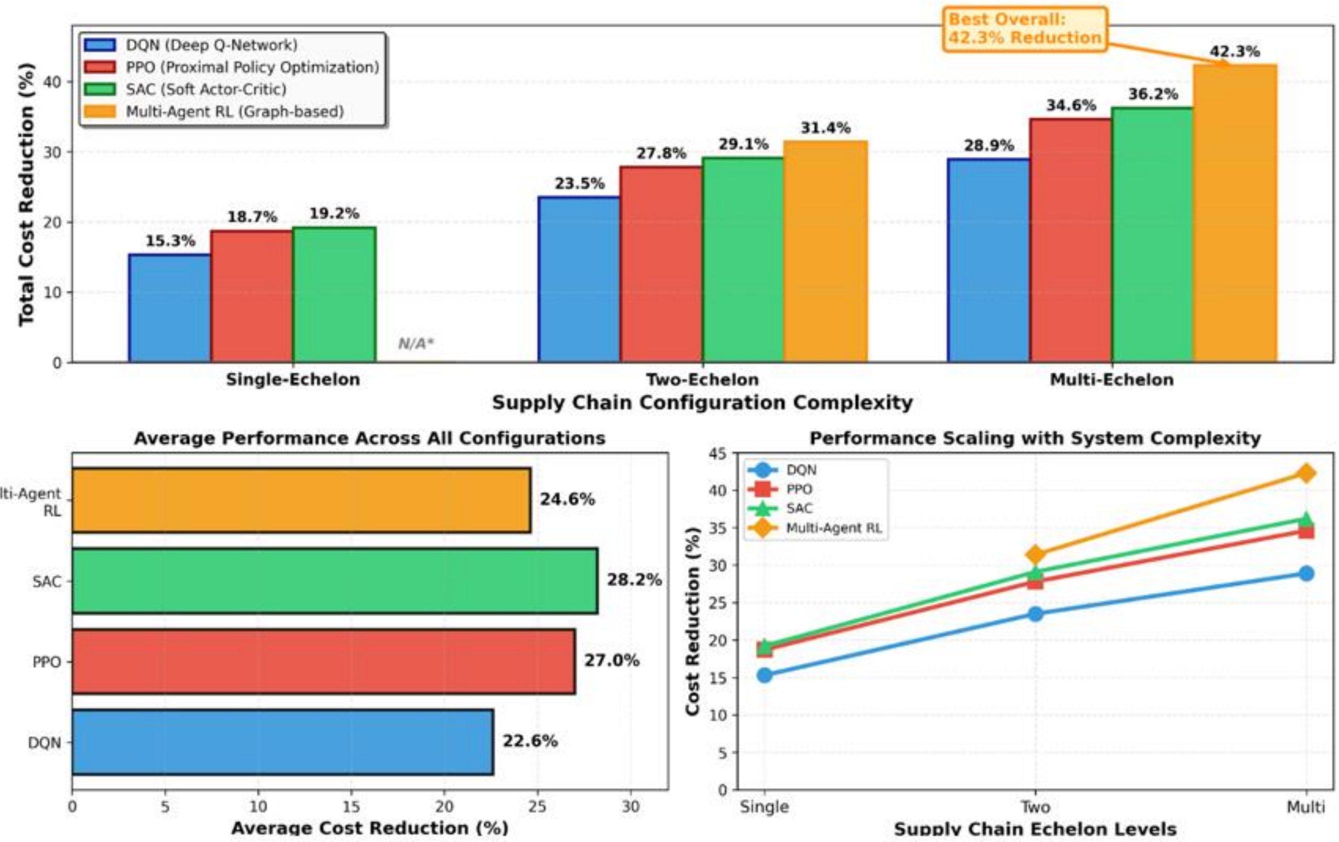

Figure 3. Reinforcement Learning Performance for Inventory Optimization Cost Reduction Compared to Traditional Base-Stock and Order-Up-To Policies.

Figure 3. Performance comparison table showing total cost reduction percentages achieved by different reinforcement learning algorithms compared to traditional inventory policies across various supply chain configurations.

As shown in Figure 3, the integration of demand forecasting with RL-based inventory optimization creates end-to-end learning systems that jointly optimize prediction and decision-making. Traditional approaches treat forecasting and optimization as separate sequential stages, where forecast errors propagate to suboptimal inventory policies. Joint training enables gradient information from the inventory optimization objective to flow back to the forecasting module, allowing the forecaster to learn representations specifically tuned for downstream decision-making rather than generic prediction accuracy. Differentiable inventory simulators that approximate the discrete dynamics of inventory systems enable end-to-end gradient-based optimization. These simulators use continuous relaxations of discrete order decisions and smooth approximations of cost functions to provide differentiable paths from forecasts to final costs, enabling the forecasting network to learn which prediction errors matter most for inventory performance.

5. Challenges and Future Directions

Despite the substantial progress in applying DL to supply chain management, several critical challenges limit the widespread adoption and effectiveness of these methods in practice. Data quality and availability constitute fundamental barriers, as DL models typically require large volumes of high-quality historical data to achieve optimal performance. Many organizations struggle with incomplete transaction records, inconsistent data formats across systems, missing values, and limited historical data for new products or recently opened locations. Transfer learning approaches that leverage knowledge from related products or markets offer potential solutions, allowing models trained on data-rich scenarios to be fine-tuned for data-scarce situations with limited additional training data. Few-shot learning and meta-learning methods that explicitly learn to learn from limited examples represent promising research directions for addressing cold-start problems in demand forecasting and inventory optimization.

Model interpretability remains a significant concern for supply chain practitioners who require understanding of why a model makes specific predictions or recommendations to build trust and identify potential issues. While attention mechanisms provide some degree of interpretability by visualizing which historical periods or features most influence predictions, they do not fully explain the complex non-linear transformations that DL models apply to inputs. Post-hoc explanation methods such as local interpretable model-agnostic explanations (LIME) and Shapley additive explanations (SHAP) can provide insights into model behavior for specific predictions, but they require computational overhead and may not capture global model behavior. Developing inherently interpretable DL architectures that maintain high prediction accuracy while providing clear decision-making rationales represents an important research direction. Hybrid approaches that combine interpretable components such as trend and seasonal decomposition with DL modules for capturing residual patterns may offer practical compromises between accuracy and interpretability.

Computational requirements for training and deploying DL models pose practical challenges, particularly for organizations with limited technical infrastructure or real-time forecasting needs. Large transformer models may require substantial GPU resources for training, and even inference can be computationally expensive when forecasts must be generated for thousands of products across multiple locations. Model compression techniques including pruning, quantization, and knowledge distillation can reduce model size and computational demands while maintaining acceptable accuracy levels. Neural architecture search (NAS) methods that automatically discover efficient architectures tailored to specific forecasting tasks may identify models with better accuracy-efficiency trade-offs than manually designed architectures. Edge deployment strategies that perform inference locally on less powerful hardware while using cloud resources for periodic model updates can enable real-time forecasting in resource-constrained environments.

The robustness of DL models to distribution shifts and adversarial perturbations requires careful attention in supply chain applications where unexpected events and changing market conditions are common. Models trained on historical data may perform poorly when faced with novel demand patterns caused by disruptions, new competitor entries, or shifts in consumer preferences. Continual learning approaches that update models as new data becomes available while preventing catastrophic forgetting of previously learned patterns can improve adaptability. Uncertainty quantification through techniques such as Bayesian neural networks, Monte Carlo dropout, or ensemble methods enables models to indicate when predictions are unreliable due to out-of-distribution inputs. Robust optimization frameworks that explicitly account for prediction uncertainty in inventory decisions can improve overall supply chain performance even when forecasts are imperfect.

Integration with existing enterprise systems and business processes represents a practical challenge that extends beyond pure algorithmic considerations. Supply chain organizations typically operate complex information technology ecosystems including enterprise resource planning (ERP) systems, warehouse management systems (WMS), and transportation management systems (TMS) that must interface with DL forecasting and optimization modules. Standardized application programming interfaces (APIs) and data exchange formats can facilitate integration, but organizational change management and user training remain critical success factors. Human-in-the-loop approaches that combine DL recommendations with domain expert judgment may provide more acceptable and effective solutions than fully automated systems, particularly during transition periods. Decision support interfaces that present model outputs alongside relevant context, historical performance metrics, and alternative scenarios can help practitioners make informed decisions based on DL insights.

Future research directions include the development of foundation models for supply chain management that are pre-trained on large diverse datasets and can be fine-tuned for specific tasks with minimal additional data. Such models could learn general representations of demand patterns, inventory dynamics, and supply chain structures that transfer across different products, industries, and geographies. Causal inference methods integrated with DL can help distinguish correlation from causation in supply chain data, enabling more robust policy learning that generalizes to interventions and counterfactual scenarios. The combination of physics-informed neural networks with data-driven learning may provide models that respect fundamental constraints of supply chain systems while maintaining flexibility to capture complex empirical patterns. Federated learning approaches that enable collaborative model training across multiple organizations while preserving data privacy could unlock the benefits of larger training datasets without requiring centralized data sharing.

6. Conclusion

DL has fundamentally transformed the landscape of demand forecasting and inventory optimization in supply chain management, offering substantial improvements over traditional statistical and optimization methods. LSTM networks, CNNs, transformers, and hybrid architectures have demonstrated superior capability in capturing complex temporal patterns, seasonal variations, and non-linear demand relationships that challenge classical approaches. The integration of attention mechanisms has improved both prediction accuracy and model interpretability by enabling dynamic focus on relevant historical periods and input features. RL frameworks combined with DL function approximators have enabled the development of adaptive inventory policies that learn through interaction with stochastic environments, outperforming rule-based heuristics particularly in multi-echelon systems with complex cost structures and correlated products. GNNs have emerged as powerful tools for modeling supply chain networks, capturing spatial dependencies between facilities and enabling coordinated decision-making across distributed systems.

The practical implementation of DL methods requires addressing challenges related to data quality, computational resources, model interpretability, and integration with existing business systems. Transfer learning, model compression, uncertainty quantification, and human-in-the-loop approaches offer promising directions for overcoming these barriers and enabling broader adoption. The development of end-to-end systems that jointly optimize forecasting and inventory decisions represents an important paradigm shift from traditional sequential approaches, allowing models to learn representations specifically tuned for downstream decision-making objectives rather than generic prediction accuracy.

Future research opportunities include foundation models that leverage pre-training on diverse supply chain datasets, causal inference methods that improve policy robustness to interventions, physics-informed architectures that incorporate domain knowledge, and federated learning frameworks that enable collaborative training while preserving privacy. As DL techniques continue to mature and computational infrastructure becomes more accessible, these methods will likely become standard components of intelligent supply chain management systems. The convergence of forecasting accuracy improvements, optimization capability enhancements, and practical deployment solutions positions DL as a cornerstone technology for building resilient, efficient, and adaptive supply chains capable of thriving in increasingly complex and uncertain business environments.

References

Abolghasemi, M., Hurley, J., Eshragh, A., & Fahimnia, B. (2020). Demand forecasting in the presence of systematic events: Cases in capturing sales promotions. *International Journal of Production Economics, 230*, 107892. https://doi.org/10.1016/j.ijpe.2020.107892

Alvo, M., Russo, D., & Kanoria, Y. (2023). Neural inventory control in networks via hindsight differentiable policy optimization. *arXiv Preprint*. https://arxiv.org/abs/2306.11246

Babai, M. Z., Boylan, J. E., & Rostami-Tabar, B. (2022). Demand forecasting in supply chains: A review of aggregation and hierarchical approaches. *International Journal of Production Research, 60*(1), 324–348. https://doi.org/10.1080/00207543.2021.1884310

Bandara, K., Bergmeir, C., & Smyl, S. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Systems with Applications, 140*, 112896. https://doi.org/10.1016/j.eswa.2019.112896

Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q., & Seaman, B. (2019). Sales demand forecast in e-commerce using a long short-term memory neural network methodology. In *International Conference on Neural Information Processing* (pp. 462–474). Springer. https://doi.org/10.1007/978-3-030-36802-9_40

Beck, M., Pöppel, K., Spanring, M., Auer, A., Prudnikova, O., Kopp, M., & Hochreiter, S. (2024). XLSTM: Extended long short-term memory. In *Advances in Neural Information Processing Systems, 37*, 107547–107603.
Bhatia, A. S., Saggi, M. K., Sundas, A., & Ashta, J. (2020). Reinforcement learning. In *Machine learning and big data: Concepts, algorithms, tools and applications* (pp. 281–303). Springer.
Bo, L., & Xu, J. (2024). Enhancing supply chain efficiency resilience using predictive analytics and computational intelligence techniques. *IEEE Access, 12*, 183451–183465. https://doi.org/10.1109/ACCESS.2024.3500000
Chen, J., Cui, Y., Zhang, X., Yang, J., & Zhou, M. (2024). Temporal convolutional network for carbon tax projection: A data-driven approach. *Applied Sciences, 14*(20), 9213. https://doi.org/10.3390/app14209213
Chen, Y., Kang, Y., Chen, Y., & Wang, Z. (2020). Probabilistic forecasting with temporal convolutional neural network. *Neurocomputing, 399*, 491–501. https://doi.org/10.1016/j.neucom.2020.03.056
Cheng, H., & Brown, J. W. (2025). Replay as a basis for backpropagation through time in the brain. *Neural Computation, 37*(3), 403–436. https://doi.org/10.1162/neco_a_01593
Cheng, Y., Liu, Z., & Morimoto, Y. (2020). Attention-based SeriesNet: An attention-based hybrid neural network model for conditional time series forecasting. *Information, 11*(6), 305. https://doi.org/10.3390/info11060305
Cheng, Y., Xu, C., Mashima, D., Thing, V. L., & Wu, Y. (2019). PowerNet: Neural power demand forecasting in smart grid. *arXiv Preprint*. https://arxiv.org/abs/1904.11979
Corsini, R. R., Costa, A., Cannella, S., & Framinan, J. M. (2023). Analysing the impact of production control policies on the dynamics of a two-product supply chain with capacity constraints. *International Journal of Production Research, 61*(6), 1913–1937. https://doi.org/10.1080/00207543.2021.2001450
Esmizadeh, Y., & Mellat Parast, M. (2021). Logistics and supply chain network designs: Incorporating competitive priorities and disruption risk management perspectives. *International Journal of Logistics Research and Applications, 24*(2), 174–197. https://doi.org/10.1080/13675567.2020.1742703
Fan, C., Zhang, Y., Pan, Y., Li, X., Zhang, C., Yuan, R., & Huang, H. (2019). Multi-horizon time series forecasting with temporal attention learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2527–2535). https://doi.org/10.1145/3292500.3330662
Fang, X., Zang, J., Zhai, Z., Zhang, L., Shu, Z., & Liang, Y. (2023). Exploring potential dual-stage attention-based recurrent neural network machine learning application for dosage prediction in intelligent municipal management. *Environmental Science: Water Research & Technology, 9*(3), 890–899. https://doi.org/10.1039/D2EW00710G
Getele, G. K., Li, T., & Arrive, J. T. (2019). Risk management in the service supply chain: Evidence from the healthcare sector. *IEEE Engineering Management Review, 47*(4), 143–152. https://doi.org/10.1109/EMR.2019.2931397
Gijsbrechts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. J. (2022). Can deep reinforcement learning improve inventory management? Performance on lost sales, dual-sourcing, and multi-echelon problems. *Manufacturing & Service Operations Management, 24*(3), 1349–1368. https://doi.org/10.1287/msom.2021.1003
Gong, H., Zhang, Y., & Zhang, Z. H. (2023). Training demand prediction models by decision error for two-stage lot-sizing problems. *IEEE Transactions on Automation Science and Engineering, 21*(2), 1122–1137. https://doi.org/10.1109/TASE.2023.3256143
Hoedt, P. J., Kratzert, F., Klotz, D., Halmich, C., Holzleitner, M., Nearing, G. S., & Klambauer, G. (2021). MC-LSTM: Mass-conserving LSTM. In *Proceedings of the 38th International Conference on Machine Learning* (pp. 4275–4286). PMLR.
Hu, Y., Liu, Y., Lv, S., Xing, M., Zhang, S., Fu, Y., & Xie, L. (2020). DCCRN: Deep complex convolution recurrent network for phase-aware speech enhancement. *arXiv Preprint*. https://arxiv.org/abs/2008.00264
Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B., & Ekmis, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. *Complexity, 2019*, 9067367. https://doi.org/10.1155/2019/9067367
Kosasih, E. E., & Brintrup, A. (2022). A machine learning approach for predicting hidden links in supply chains with graph neural networks. *International Journal of Production Research, 60*(17), 5380–5393. https://doi.org/10.1080/00207543.2021.1995276
Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement learning with thinking model for dynamic supply chain network design. *IEEE Access, 12*, 195974–195985. https://doi.org/10.1109/ACCESS.2024.3503337
Liang, X. X., Lu, B., Deng, C., Gu, Y., Chen, K. D., Mo, Y. H., & Liu, M. (2024). Product sales forecasting model driven by multi-source data integration based on XGBoost. In *Fuzzy Systems and Data Mining X* (pp. 166–179). IOS Press. https://doi.org/10.3233/FAIA230891
Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A, 379*(2194), 20200209. https://doi.org/10.1098/rsta.2020.0209
Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting, 37*(4), 1748–1764. https://doi.org/10.1016/j.ijforecast.2021.03.012
Lu, S., Zhang, K., Chen, T., Başar, T., & Horesh, L. (2021). Decentralized policy gradient descent-ascent for safe multi-agent reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12).
Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). Predicting/hypothesizing the findings of the M4 competition. *International Journal of Forecasting, 36*(1), 29–36. https://doi.org/10.1016/j.ijforecast.2019.05.010
Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent neural networks: A comprehensive review of architectures, variants, and applications. *Information, 15*(9), 517. https://doi.org/10.3390/info15090517
Mirshahi, S., Brandtner, P., & Oplatková, Z. K. (2024). Intermittent time series demand forecasting using dual convolutional neural networks. *MENDEL, 30*(1), 51–59. https://doi.org/10.13164/mendel.2024.1.051
Montenegro, A., Mussi, M., Metelli, A. M., & Papini, M. (2024). Learning optimal deterministic policies with stochastic policy gradients. *arXiv Preprint*. https://arxiv.org/abs/2405.02235
Naing, T. K. (2024). *Enhancing last-mile delivery efficiency through crowdsourced workforce scheduling and routing optimization* (Doctoral dissertation, Ulsan National Institute of Science and Technology).
Onyeanuforo, U. L., & Mohammed, B. B. (2024). Data-driven decision-making in global supply chain management: Transforming operations through analytics and intelligence. *Nigerian Journal of Sustainability Research, 2*(2), 1–23.
Oroojlooyjadid, A., Nazari, M., Snyder, L. V., & Takáč, M. (2022). A deep Q-network for the beer game: Deep reinforcement learning for inventory optimization. *Manufacturing & Service Operations Management, 24*(1), 285–304. https://doi.org/10.1287/msom.2020.0911
Peng, Z., Zhang, Y., Feng, Y., Zhang, T., Wu, Z., & Su, H. (2019). Deep reinforcement learning approach for capacitated supply chain optimization under demand uncertainty. In *2019 Chinese Automation Congress (CAC)* (pp. 3512–3517). IEEE. https://doi.org/10.1109/CAC48633.2019.8997352
Pescia, G., Nys, J., Kim, J., Lovato, A., & Carleo, G. (2024). Message-passing neural quantum states for the homogeneous electron gas. *Physical Review B, 110*(3), 035108. https://doi.org/10.1103/PhysRevB.110.035108
Rashid, T., Samvelyan, M., De Witt, C. S., Farquhar, G., Foerster, J., & Whiteson, S. (2020). Monotonic value function factorisation for deep multi-agent reinforcement learning. *Journal of Machine Learning Research, 21*(178), 1–51.
Rolf, B., Jackson, I., Müller, M., Lang, S., Reggelin, T., & Ivanov, D. (2023). A review on reinforcement learning algorithms and applications in supply chain management. *International Journal of Production Research, 61*(20), 7151–7179. https://doi.org/10.1080/00207543.2022.2151405
Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting, 36*(3), 1181–1191. https://doi.org/10.1016/j.ijforecast.2019.07.001
Sivertsvik, M., Sumskiy, K., & Misimi, E. (2024). Learning active manipulation to target shapes with model-free, long-horizon deep reinforcement learning. In *2024 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5411–5418). IEEE. https://doi.org/10.1109/ICRA57147.2024.10611245
Stranieri, F., & Stella, F. (2022). A deep reinforcement learning approach to supply chain inventory management. *arXiv Preprint*. https://arxiv.org/abs/2204.09603

Sun, T., Wang, M., & Chen, J. (2025). Leveraging machine learning for tax fraud detection and risk scoring in corporate filings. *Asian Business Research Journal, 10*(11), 1–13.

Suwanda, R., Syahputra, Z., & Zamzami, E. M. (2020). Analysis of Euclidean distance and Manhattan distance in the K-means algorithm for variations number of centroid K. In *Journal of Physics: Conference Series, 1566*(1), 012058. IOP Publishing. https://doi.org/10.1088/1742-6596/1566/1/012058

Taube, F., & Minner, S. (2023). Optimal inventory control with cyclic fixed order costs. *Production and Operations Management, 32*(10), 3286–3294. https://doi.org/10.1111/poms.14036

Tekinerdogan, B., Verdouw, C., & van Haaren, B. (2024). A framework for mapping inventory decisions to machine learning methods in perishable food supply chains. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.5854392

Wambura, S., Li, H., & Nigussie, A. (2020). Fast memory-efficient extreme events prediction in complex time series. In *Proceedings of the 2020 3rd International Conference on Robot Systems and Applications* (pp. 60–69).

Wang, Y., & Xing, S. (2025). AI-driven CPU resource management in cloud operating systems. *Journal of Computer and Communications, 13*(6), 135–149. https://doi.org/10.4236/jcc.2025.136009

Wu, H., Zhao, H., & Zhang, M. (2021). Not all attention is all you need. *arXiv Preprint.* https://arxiv.org/abs/2104.04692

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems, 32*(1), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

Xie, J., Gao, R., Nijkamp, E., Zhu, S. C., & Wu, Y. N. (2020). Representation learning: A statistical perspective. *Annual Review of Statistics and Its Application, 7*(1), 303–335. https://doi.org/10.1146/annurev-statistics-031219-041705

Xing, S., Wang, Y., & Liu, W. (2025). Multi-dimensional anomaly detection and fault localization in microservice architectures: A dual-channel deep learning approach with causal inference for intelligent sensing. *Sensors, 25*(11), 3396. https://doi.org/10.3390/s25113396

Xu, S., Zhang, X., Feng, L., & Yang, W. (2020). Disruption risks in supply chain management: A literature review based on bibliometric analysis. *International Journal of Production Research, 58*(11), 3508–3526. https://doi.org/10.1080/00207543.2020.1712494

Zeng, P., Hu, G., Zhou, X., Li, S., & Liu, P. (2023). Seformer: A long sequence time-series forecasting model based on binary position encoding and information transfer regularization. *Applied Intelligence, 53*(12), 15747–15771. https://doi.org/10.1007/s10489-023-04372-3

Zhang, H., Lu, G., Zhan, M., & Zhang, B. (2022). Semi-supervised classification of graph convolutional networks with Laplacian rank constraints. *Neural Processing Letters, 54*(4), 2645–2656. https://doi.org/10.1007/s11063-021-10671-5

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12), 11106–11115. https://doi.org/10.1609/aaai.v35i12.17325

Zhu, R., & Rigotti, M. (2021). Self-correcting Q-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12), 11185–11192. https://doi.org/10.1609/aaai.v35i12.17336

Zuo, D., Li, M., Zeng, L., Wang, M., & Zhao, P. (2025). A cross-attention based diffusion convolutional recurrent neural network for air traffic forecasting. In *AIAA SCITECH 2025 Forum* (p. 2120).