Management and Production Engineering Review
Volume 17 • Number 2 • June 2026 • pp. 1–14
DOI: 10.24425/mper.2026.1313

Optimizing Production Planning with Neural Networks and Genetic Algorithms: A Case Study in Forecasting and Buffer Management

Sahar HABBADI[1], Ismail EL MOUAYNI[2], Brahim HERROU[1]

[1] *Faculté des Sciences Techniques de Fès, Industrial Engineering Department, Morocco*
[2] *CRAN – Research Centre for Automatic Control of Nancy, 57070 Metz, France*

Received: 16 May 2025
Accepted: 16 February 2026

Abstract
This study proposes a two-stage hybrid methodology that integrates neural networks (NN) for demand forecasting with genetic algorithms (GA) for buffers-related cost optimization in demand-driven material requirements planning (DDMRP) systems. The approach addresses a key gap in production planning research by systematically combining data-driven demand prediction with combinatorial optimization to reduce production costs using decoupled lead time (DLT), a core DDMRP metric. Forecast-derived demand intensities are incorporated into DLT weighting, enabling prioritized buffer placement at high-demand, long-lead-time workstations. Applied to an automotive sun-visor production line with eight sequential workstations, the method yielded an optimal buffer configuration that demonstrates the effectiveness of coupling NN pattern-recognition capabilities with GA optimization performance. The resulting framework provides production managers with a flexible and quantifiable decision-support tool for buffer positioning in dynamic manufacturing environments. Future research should extend the method to multi-product systems, explore advanced long short-term memory (LSTM) architectures to enhance forecasting accuracy, and conduct cross-industry evaluations to assess generalizability.

Keywords
Demand-Driven Material Requirements Planning (DDMRP); Genetic Algorithms; Inventory Management; Buffer Positioning; Neural Networks.

Introduction

Background and Motivation

Production planning is a critical process in manufacturing operations, encompassing the organization and coordination of activities to determine what products to produce, in what quantities, and at what times. The primary objectives of production planning include meeting market demand, optimizing resource utilization, balancing supply and demand, minimizing operational costs, controlling lead times, and preventing both stockouts and overproduction. However, achieving these objectives presents significant challenges due to inherent uncertainties in demand patterns, variability in supply lead times, large volumes of operational data, and constrained resources.

In recent years, advanced optimization techniques such as GA and predictive models like NN have gained considerable attention for improving production planning and inventory management. NN offers the capability to accurately forecast future demand and supply requirements by learning complex patterns from historical data, providing precise inputs for downstream decision-making processes. GA, inspired by natural selection principles, are particularly well-suited for optimizing inventory levels, costs, and lead times through their ability to explore vast solution spaces efficiently. When combined with metrics such as the DLT from DDMRP, these techniques can form powerful optimization frameworks.

Demand-Driven Material Requirements Planning has emerged as an innovative approach to materials management that strategically positions inventory buffers at critical points in production networks, with replenishment driven by actual demand consumption rather than forecasts alone. Unlike traditional materials requirements planning (MRP) systems that rely heavily on forecasting and push-based scheduling, DDMRP introduces a demand-pull mechanism that

***Corresponding author**: Ismail EL Mouyani – CRAN – Research Centre for Automatic Control of Nancy, 57070 Metz, France, e-mail: ismail.elmouayni@gmail.com*

enhances responsiveness to market variations while reducing the bullwhip effect in supply chains. The DLT metric quantifies the cumulative production time from the last strategic buffer position to the end of the production line, serving as a key indicator for buffer positioning decisions.

Research Gap

Despite growing interest in both machine learning applications and metaheuristic optimization for production planning, existing literature exhibits significant fragmentation in addressing integrated solutions. Recent studies combining machine learning and optimization techniques show promising but incomplete results across several dimensions.

Gap 1: Limited Integration of Forecasting and Optimization

Studies applying neural networks to demand forecasting generally improve short-term prediction accuracy (Gołąbek et al., 2020; Zonta et al., 2022), yet they rarely integrate these forecasts as dynamic inputs into downstream optimization models. Conversely, GA-based research in production planning often focuses on isolated components such as scheduling (Yu & Liang, 2001), buffer positioning (Damand et al., 2023), or inventory minimization (Younespour et al., 2024), but without incorporating data-driven forecasting as a variable influencing the optimization process. This disconnect creates a missed opportunity for leveraging predictive insights to inform buffer placement decisions.

Gap 2: Absence of Demand-Responsive DLT Formulations

While the DDMRP literature emphasizes the importance of strategic buffer positioning (Azzamouri et al., 2021; Math et al., 2024), current approaches typically treat DLT as a static metric based solely on lead times, without incorporating demand intensity or variability. Existing optimization models for DDMRP (Damand et al., 2023; Haji Mohammad et al., 2022) focus on minimizing costs and lead times but do not systematically weight these factors by forecasted demand patterns. This limitation reduces the adaptability of buffer positioning strategies to dynamic demand environments.

Gap 3: Lack of End-to-End Hybrid Frameworks

Existing hybrid approaches typically address specific case studies without providing a generalized, end-to-end methodology for production planning that jointly optimizes forecasting accuracy, inventory levels, and lead-time management (Paraschos & Koulouriotis, 2024). Literature lacks a systematic framework that explicitly connects demand forecasting outputs to buffer optimization inputs through a formal integration mechanism. Moreover, recent advances in reinforcement learning for DDMRP parameter optimization (Duhem et al., 2023) present alternative approaches, yet comparative analyses between two-stage (forecast → optimize) and end-to-end learning paradigms remain unexplored.

Gap 4: Limited Methodological Specificity for Binary Placement Problems

Many GA applications in manufacturing employ crossover and mutation operators designed for permutation or continuous optimization problems (Davis, 1985), without adequately addressing the specific requirements of binary decision problems such as buffer placement. The selection of appropriate genetic operators for binary-encoded chromosomes represents an underemphasized methodological consideration that can significantly impact optimization performance.

These gaps collectively highlight the need for an integrated methodology that systematically combines NN forecasting with GA optimization, incorporates demand-responsive metrics into buffer positioning decisions, and provides a reproducible framework applicable to DDMRP-based production systems.

Research Objective

In response to the identified research gaps, this study has the following objective:

Primary Objective

To develop and validate a two-stage hybrid methodology that integrates NN for demand forecasting with GA for buffer positioning optimization in DDMRP-based production systems, demonstrated and evaluated through an automotive manufacturing case study.

Specific Sub-Objectives

Formulate a GA optimization model for buffer placement that incorporates a novel demand-weighted DLT metric.

Develop an integration mechanism that systematically connects demand forecasts to buffer optimization through weighted DLT calculations.

Provide practical implementation guidance for production managers seeking to adopt data-driven buffer positioning strategies.

Contributions of This Study

This research makes three primary contributions to production planning and DDMRP literature:

Methodological Contribution

This study presents the first systematic integration of machine learning forecasting with metaheuristic optimization specifically designed for DDMRP buffer positioning. While prior work has applied GAs to DDMRP parameter optimization (Damand et al., 2023; Younespour et al., 2024) and NNs to demand forecasting separately, this research develops an explicit integration framework through demand-weighted DLT formulation. The methodology addresses the binary nature of buffer placement decisions by employing appropriate genetic operators (uniform crossover), correcting common methodological oversights in GA applications.

Theoretical Contribution

The proposed demand-weighted DLT metric extends traditional DLT concepts by incorporating forecasted demand intensity, enabling dynamic buffer prioritization based on both lead time and demand patterns. This formulation addresses a gap identified in the DDMRP systematic review by (Azzamouri et al., 2021) regarding the need for more adaptive buffer management strategies. The approach demonstrates that two-stage ML-Metaheuristic frameworks can effectively balance the strengths of NN (complex pattern recognition) and GA (combinatorial optimization) without requiring the extensive training data typical of end-to-end reinforcement learning approaches.

Practical Contribution

The study provides production managers with a reproducible, data-driven methodology for buffer positioning decisions, supported by clear implementation guidance including data requirements, computational specifications, and cost-benefit analysis frameworks. The automotive case study demonstrates the practical applicability of the approach, achieving quantifiable improvements in lead time reduction and stockout prevention while maintaining computational feasibility for industrial implementation.

Literature review

Demand-Driven MRP and Buffer Optimization

DDMRP represents a significant evolution in materials management methodologies, addressing the limitations of traditional MRP systems in coping with demand variability and supply chain complexity. Azzamouri et al. (2021) conducted a comprehensive systematic review of DDMRP approaches, classifying methodologies and identifying key components including demand-driven inventory management, strategic buffer positioning, and dynamic adaptation to production changes. The review emphasizes that DDMRP combines advantages of traditional MRP systems with demand-pull principles to better manage fluctuations and uncertainties.

Buffer Positioning and Sizing

The strategic positioning of inventory buffers at critical network nodes constitutes a fundamental aspect of DDMRP implementation. Haji Mohammad et al. (2022) proposed an optimization model for demand-driven distribution resource planning (DDDRP), extending DDMRP concepts to distribution networks. Their nonlinear mixed mathematical formulation optimizes buffer placement to minimize storage costs, demonstrating through a case study that DDDRP-optimized networks can reduce storage costs by 75% and inventory quantities by 67% compared to traditional distribution resource planning (DRP) approaches. However, the model relies on specific assumptions that limit generalizability, and the authors emphasize the need for validation across diverse industrial contexts.

Math et al. (2024) explored DDMRP application for inventory optimization in supply chains, presenting a framework based on product segmentation, prioritization of critical flows, and dynamic buffer adjustment according to actual demand variations. Their results highlighted significant reductions in storage costs and improved responsiveness to market variations. The study underscores that successful DDMRP implementation requires advanced technological integration, organizational adjustment, and stakeholder training.

Genetic Algorithms for Production Planning

Genetic algorithms, introduced by Holland (1992), have become widely adopted for production planning optimization due to their ability to efficiently explore large solution spaces and avoid premature convergence to local optima. The fundamental mechanisms of selection, crossover, and mutation enable GAs to progressively evolve populations of candidate solutions toward optimal or near-optimal configurations.

GA for DDMRP Optimization

Younespour et al. (2024) explored optimization of strategic and operational levels of DDMRP using a hybrid algorithm combining GA with particle swarm optimization (PSO). The hybrid GA-PSO approach optimizes buffer positioning and sizing by balancing inventory costs against product availability, combining the global exploration capabilities of GAs with the rapid convergence characteristics of PSO. Results demonstrated significant performance improvements, reducing costs and increasing supply chain efficiency. This work illustrates the potential of evolutionary algorithms for DDMRP but does not incorporate demand forecasting as a dynamic input.

Damand et al. (2023) proposed an innovative approach using a multi-objective genetic algorithm (MOGA) to optimize DDMRP parameters, focusing on buffer levels, replenishment points, and variability thresholds. The methodology simultaneously addresses multiple objectives: minimizing inventory costs, maximizing customer service levels, and improving supply chain resilience. The MOGA explores solution spaces to balance often-conflicting objectives, incorporating real demand fluctuation scenarios and logistical constraints. Results showed that genetic approaches can identify optimal configurations outperforming traditional methods, though the authors note the need for further research integrating additional dimensions such as environmental impacts.

In Habbadi et al. (2025), a hybrid method DDMRP with a GA is proposed to enhance buffer positioning and production-control decisions. The approach relies on a mixed-integer programming model (MIP) that encodes key DDMRP principles—including decoupling-point placement, dynamic buffer sizing, and demand-driven replenishment—to capture the fundamental trade-off between inventory minimization and demand-satisfaction maximization. This MIP formulation provides a cost-oriented evaluation framework, while the GA efficiently explores alternative buffer configurations and replenishment strategies within this DDMRP-constrained search space.

GA Operator Selection for Binary Problems

The selection of appropriate genetic operators is critical for optimization performance, particularly for binary-encoded problems. Goldberg (1989) established foundational principles for GA design, emphasizing that uniform crossover or two-point crossover is more appropriate than order-based crossover for binary chromosomes representing independent decision variables. Order crossover (OX), originally developed by Davis (1985), is specifically designed for permutation problems such as traveling salesman or job sequencing tasks where relative order of genes matters. For binary buffer placement problems where each position is an independent yes/no decision, uniform crossover treats each gene position independently, providing more effective exploration of the solution space.

Neural Networks for Demand Forecasting

Neural networks have demonstrated substantial capabilities for demand forecasting in manufacturing contexts, learning complex nonlinear patterns from historical data to generate accurate predictions for future periods.

Traditional NN Architectures

Yu & Liang (2001) proposed a hybrid method combining artificial NN with GA for dynamic production scheduling. Their approach optimizes operation sequences and starting times, reducing costs and improving system flexibility. The study represents an early demonstration of NN–GA integration, though the forecasting and optimization components operate relatively independently without systematic integration mechanisms.

Zonta et al. (2022) integrated predictive maintenance with production scheduling using deep neural networks (DNN) and recurrent neural networks (RNN), achieving significant improvements in equipment management and resource allocation. The study demonstrates that deep architecture can capture temporal dependencies in production data, leading to more accurate predictions that enable proactive scheduling decisions.

Graph-Based and Advanced Architectures

Recent advances have explored graph neural networks (GNN) for production planning applications. Hameed & Schwung (2023) introduced an innovative approach leveraging GNNs and reinforcement learning for production planning optimization. By representing interdependencies between tasks, machines, and resources as graph structures, the method achieved superior performance compared to traditional approaches in terms of accuracy and adaptability.

Lai et al. (2023) proposed a GNN–based methodology to predict future bottlenecks in production systems. By modeling production elements and interactions as graphs, this approach enables identification of potential congestion areas, improving decision-making and operational continuity. Results indicated that the method outperforms conventional techniques in dynamic production environments.

Smit et al. (2025) conducted a comprehensive survey of graph NN for job shop scheduling problems, exploring the potential of GNNs to model complex dependencies in scheduling tasks. While these methods offer significant advantages in flexibility and accuracy, challenges related to data management and model generalization remain important research directions.

State-of-the-Art regarding forecasting

Paraschos & Koulouriotis (2024) presented a comprehensive literature review on learning-based production, maintenance and quality optimization in smart manufacturing systems. The review identifies key areas including predictive maintenance, production flow optimization, and quality control improvement, emphasizing integration of IoT data and cyber-physical systems. Emerging trends include deep neural networks, hybrid AI approaches, and reinforcement algorithms. The authors address challenges such as data complexity, cybersecurity, and scalability, advocating for continued research to foster autonomous and resilient manufacturing ecosystems.

Mumali & Kałkowska (2024) highlighted the utility of artificial NN, fuzzy logic, and GA in managing industrial complexity, facilitating multi-objective optimization for cost efficiency, quality, and sustainability. The study emphasizes emerging hybrid approaches combining AI methodologies for adoption in smart factories and cyber-physical manufacturing systems.

Recent work by Han et al. (2023) on attention-based LSTM architectures for demand forecasting represents the current state-of-the-art, demonstrating that LSTM models with attention mechanisms significantly improve forecasting accuracy by capturing long-term dependencies and focusing on relevant historical periods. These advances suggest promising directions for enhancing the forecasting component of production planning systems.

Hybrid ML + Metaheuristic Approaches

The integration of machine learning and metaheuristic optimization has gained increasing attention as a means to leverage complementary strengths of both paradigms.

Synergistic Integration Trends

Zhang et al. (2025) conducted a comprehensive review of synergistic integration of metaheuristics and machine learning, identifying latest advances and emerging trends. The review positions hybrid ML-metaheuristic approaches within broader artificial intelligence developments, showing that two-stage frameworks (where ML outputs inform metaheuristic inputs) and co-evolutionary systems (where ML and metaheuristics iteratively refine each other) represent two primary integration patterns. The authors note that systematic integration mechanisms, as opposed to adhoc combinations, yield more robust and generalizable solutions.

Reinforcement Learning Alternatives

An important alternative to two-stage hybrid approaches is end-to-end reinforcement learning for production planning optimization. Duhem et al. (2023) presented a parametrization of demand-driven operating models using reinforcement learning for DDMRP applications. Their approach demonstrates that RL agents can learn optimal buffer policies through trial-and-error interaction with demand environments, potentially offering more adaptive solutions than static GA optimization. However, RL approaches require extensive training data, exhibit limited interpretability compared to GA-based methods, and may face challenges generalizing to significantly different demand patterns than those encountered during training.

The present study complements this direction by demonstrating the effectiveness of the two-stage NN+GA framework, which offers advantages in interpretability (GA solutions are directly analyzable), data efficiency (NN requires less data than RL for forecasting), and modularity (forecast and optimization components can be updated independently). Future work comparing two-stage and end-to-end approaches across multiple performance dimensions would provide valuable insights for methodology selection in different industrial contexts.

Research Gap and Positioning

The literature review reveals that while substantial progress has been made in applying NN to demand forecasting and GA to production optimization separately, systematic integration of these techniques for DDMRP buffer positioning remains underexplored. Specifically:

No existing work systematically integrates demand forecasts from NNs into GA-based buffer optimization for DDMRP through an explicit mathematical formulation.

Traditional DLT metrics do not incorporate demand intensity, limiting their responsiveness to variable demand patterns.

Methodological details regarding appropriate GA operators for binary buffer placement decisions are often overlooked or incorrectly applied.

Comparative analyses between two-stage hybrid approaches and alternative methods (reinforcement learning, traditional DDMRP) are absent.

This study addresses these gaps by proposing a demand-weighted DLT formulation that connects NN forecasts to GA optimization, employing methodologically appropriate genetic operators for binary decisions, and providing a comprehensive framework validated through real-world application.

Methodology

The proposed methodology consists of several complementary steps, presented below in the form of algorithm:

Inputs

- Historical production and sales data
- Bill of Materials (BOM) structure
- Production constraints, costs, and lead times
- Parameters of NNs and GAs

Outputs

- Optimal production quantities
- Optimal buffer placement
- Optimized inventory levels
- Adjusted DLT for each strategic point

Steps

Initialization

- Load historical data and define NN and GA parameters
- Initialize buffer structures and strategic points

Demand Forecasting with NNs

- Train the NN model on historical data
- Predict future demand for each product or strategic point
- Store forecasts as dynamic inputs for the optimization process

Optimization with GA

- Initialize a population of candidate solutions (production plans, stock levels, buffer positions)
- For each iteration (representing a complete cycle of result analysis):
  - Iteration 1 – Initial exploration phase
    - Generations 1-4: Initial population and first selection
    - Result: Identification of promising patterns
  - Iteration 2 – Optimization phase
    - Generations 5-12: Intensive crossover and local search
    - Result: Discovery of the optimal solution
  - Iteration 3 – Convergence phase
    - Generations 13-30: Validation and stabilization
    - Result: Convergence towards the global optimum
- For each generation within the current iteration:
  - Calculate the DLT for each strategic point
  - Adjust the DLT based on demand forecasts from NNs
  - Evaluate the solution using a fitness function considering:
    - Production and inventory costs
    - Lead time compliance
    - Service level / stockout risk
  - Apply GA operators (selection, crossover, mutation) to generate a new population
- Repeat until convergence or the maximum number of generations is reached

Selection of the Optimal Solution

- Identify the candidate solution with the best fitness value
- Extract optimal production quantities, inventory levels, and DLT

System Integration and Iterative Optimization

- Implement the optimized decisions in the production system
- Periodically recalculate:
- Demand forecasts with NNs
- Stock levels and DLT optimization with GA
- Update production plans and buffers to form a self-adaptive planning loop

The overall workflow described above is also illustrated in Figure 1.

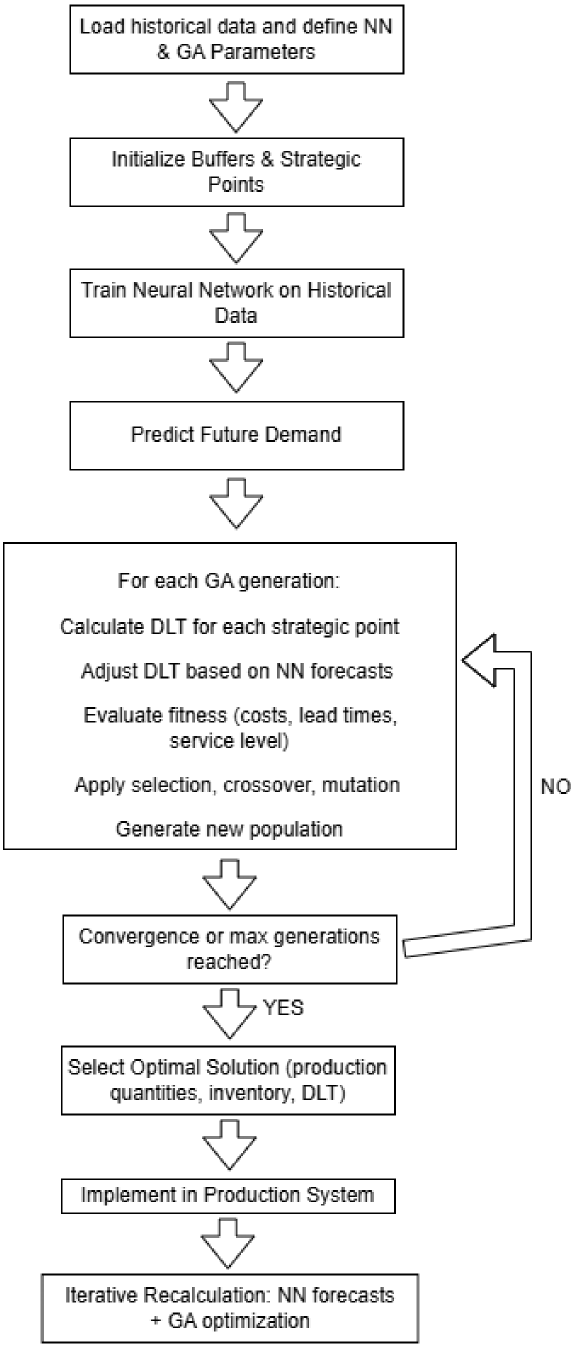

Fig. 1. Production Planning Methodology Flowchart

Case study: Automative Sun Visor Production Line

This section applies the two-phase hybrid methodology to a real-world industrial case: sun visors manufacturing for a major automotive group. We demonstrate the complete workflow from demand forecasting to optimal buffer placement using actual operational data.

Problem Statement and Production Line Description

The production line consists of nine sequential workstations, each performing a specific operation in the sun visor assembly process. The manufacturing process relies on a structured Bill of Materials (BOM) and follows established cycle times for each operation, as detailed in Table 1.

Table 1
Bill of Materials for Sun Visor Production

Reference	Component
185201270	MIR ECL Y283 (Mirror)
191026692	ENS DEMI-MOUS Y283 (Glue)
100019520	TIS ADH SOFT L: 1480mm (Cloth)
156059242	RAQUETTE Y283 (Sun visor bracket)
186204832	ART Y283 (Articulation)
165109810	FAISCEAU Y283 (Harnesses)
146002690	CLIPS 25mm (Clips)
184001490	SACHET F 220*550 (Packaging)

This table lists all components required for assembling the automotive sun visor, including mirrors, foam parts, textile layers, brackets, articulations, harnesses, clips, and packaging materials. Each reference corresponds to a specific item used along the production line's workstations.

The manufacturing process of the studied sun visor reference follows a sequence of nine production steps, each performed at a dedicated workstation. The average cycle time of each workstation, which determines the overall production rhythm and influences line balancing decisions, is presented in Table 2.

Table 2
Average cycle time of workstations in the sun visor production line

Post	Operation	Cycle time average (s)
1	Reboarding	30.97
2	reboarding finishing	24.57
3	heat transfer on fabric	25.91
4	Greasing, clipping, jointing	18.44
5	harness insertion, wallet assembly with the racket	17.55
6	sun bonding	39
7	touch-up/finishing	24.41
8	Mirror welding	22.65
9	Mirror clipping, Control and packaging	22.33

Phase 1: Demand Forecasting and Station Load Calculation

The first phase aims to transform historical demand observations into short-term production forecasts, which subsequently serve as inputs for calculating station-level workloads. Demand prediction is performed using a feed-forward NN trained on recent sales data and temporal indicators.

Neural Network Inputs and Architecture

The input vector is composed of two categories of variables:

- Demand recorded over the five previous days.
- One-hot encoding of the day of the week.

For the use case, the historical demand values are $[45, 52, 60, 85, 40]$ units, and the temporal encoding corresponding to Monday is $[1, 0, 0, 0, 0, 0, 0]$. This results in a 12-dimensional input vector. In standard practice, these values are normalized prior to training to improve numerical stability and convergence.

The network architecture used for forecasting consists of an input layer of 12 neurons, followed by two hidden layers with 64 and 32 neurons respectively, both using the Rectified Linear Unit (ReLU) activation function. The output layer contains 3 neurons with linear activation, providing demand forecasts for the subsequent three days.

Neural Computation Mechanism

Each neuron computes a linear transformation of its inputs, expressed as:

$$z = WX + b \tag{1}$$

where X is the input vector, W the weight matrix, and b is the bias term. The bias prevents the neuron's activation from being constrained to pass through the origin and increases the representational flexibility of the model. The linear output z is then passed through a nonlinear activation function. For ReLU, the activation is defined as:

$$a = \mathrm{ReLU}(z) = \max(0, z) \tag{2}$$

a choice widely adopted in modern deep learning due to its computational efficiency and its favorable gradient propagation properties.

Case study numerical results

To illustrate this mechanism, we consider a simplified NN with a single hidden layer of three neurons. Only the five historical demand values are used as inputs:

$$X = [45, 52, 60, 85, 40] \tag{3}$$

The weight matrix and bias vector of the hidden layer are given by:

$$W = \begin{pmatrix} 0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\ 0.3 & -0.2 & 0.1 & 0.4 & 0.7 \\ 0.1 & 0.4 & 0.3 & -0.2 & 0.6 \end{pmatrix}, \; b = \begin{pmatrix} 0.1 \\ -0.1 \\ 0.2 \end{pmatrix} \tag{4}$$

The pre-activation outputs are computed as

$$Z = WX + b = \begin{pmatrix} 52.8 \\ 71.0 \\ 50.5 \end{pmatrix} \tag{5}$$

Since all values are positive, the ReLU function leaves them unchanged.

The output layer uses the following weights and biases:

$$W^{(2)} = \begin{pmatrix} 0.1 & 0.5 & 0.5 \\ -0.1 & 0.6 & 0.7 \\ 0.3 & 0.6 & -0.2 \end{pmatrix}, \; b^{(2)} = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix} \tag{6}$$

The final network output is then

$$h = W^{(2)}Z + b^{(2)} = \begin{pmatrix} 66.13 \\ 72.77 \\ 48.44 \end{pmatrix} \tag{7}$$

The network output for the demands of the next 3 days is therefore:

$$H_{\text{predict}} = (66.13, 72.77, 48.44) \tag{8}$$

The neural network therefore predicts a demand of 66.13, 72.77, and 48.44 units for Days 1, 2, and 3 respectively. This example illustrates the operational behavior of a feed-forward neural architecture applied to short-term time-series forecasting. Hence, the average forecasted demand is given by (9):

$$H_{\text{predict}}^{\text{average}} = (H_1 + H_2 + H_3)/3 \tag{9}$$

The NN model was validated using standard forecasting metrics as proposed in Makridakis et al. (2018); Wilson (2016), including Mean Absolute Percentage Error with a value under 10% (Hyndman & Koehler, 2006) and Root Mean Square Error (Chai & Draxler, 2014).

Using (9), This average forecasted demand of sun visors is 62.45 units per day, which serves as the baseline for station load calculations.

BOM Disaggregation and Station Load Calculation

For this case study, the production of one finished sun visor requires exactly one pass through each of the nine workstations. The BOM relationship is 1 for each station. Station-level demand calculation:

$$q_i = H_{\text{predict}}^{\text{average}} \times \text{BOM}_\text{i} = 62.45 \times 1 \tag{10}$$

Therefore: $q_1 = q_2 = q_3 = q_4 = q_5 = q_6 = q_7 = q_8 = q_9 = 62.45$ units/day.

Demand Weight Calculation

Since demand are uniform across stations, weights are calculated based on cycle time differences to reflect varying workstation intensities. First, average cycle time is calculated using (11):

$$t_c = \frac{1}{9}\sum_{i=1}^{9} t_i = 25.09 \text{ seconds} \tag{11}$$

Then individual station weight is given by (12):

$$w_i = \frac{t_i}{t_c} \tag{12}$$

The demand per workstation is weighted according to the demand intensity, as shown in Table 3. This approach allows the GA to prioritize buffer placement at high-demand stations, enhancing responsiveness and reducing the risk of stockouts in critical zones.

Workstation 6 (Sun bonding) emerges as the primary bottleneck with the highest weight (1.55), making it a critical candidate for buffer protection.

Phase 2: Buffer Optimization with Genetic Algorithm

GA Configuration and Cost Parameters

***Cost Parameters (as defined in methodology)*:**

- Cost per buffer (C_{buffer}) = 200 €/day: Inventory holding cost: 150 €/day (based on 15% annual holding rate for automotive components valued at ∼ 365 €/unit), Space allocation cost: 30 €/day (warehouse space: 2 m^2 per buffer station × 15 €/m^2/day). Handling and management: 20 €/day (additional labor for buffer monitoring and replenishment). Total justified cost: 200 €/day per buffer position.
- Penalty for unprotected critical station (C_{penalty}) = 2000 €/day: Production line stoppage: 1200 €/day (9 operators × 25 €/hour × 5.33 hours average downtime). Component waste and rework: 400 €/day (defective units due to unprotected sensitive operations). Delivery delay penalties: 250 €/day. Quality control escalation: 150 €/day (additional inspection and documentation requirements). Total justified penalty: 2000 €/day per unprotected critical station.
- Station 4 and 8 are considered critical. Cost of unprotected critical station is 2000 €/day.
  - GA Parameters:
    - Population size: $N = 50$
    - Maximum generations: $G = 100$
    - Crossover probability: $P_c = 0.8$
    - Mutation probability: $P_m = 0.05$

Table 3
Buffer configuration based on demand- weighted DLT

Workstation	Cycle time (t_i)	Demand Weight (t_i)	Interpretation
1	30.97	1.23	High workload
2	24.57	0.98	Average workload
3	25.91	1.03	Average workload
4	18.44	0.73	Low workload
5	17.55	0.70	Low workload
6	39	1.55	Bottleneck (Highest workload)
7	24.41	0.97	Average workload
8	22.65	0.90	Average workload
9	22.33	0.89	Average workload

Integration of Neural Network Forecasting Results

The GA optimization directly incorporates the NN forecasting output:

- Forecast Integration: Daily demands (66.13, 72.77, 48.44 units) → Average demand $H_{\text{predict}}^{\text{average}}$ = 62.45 units/day.
- Station Load Propagation: q_i = 62.45 units/day for all stations (BOM ratio).
- Weight Adjusted DLT Calculation: Each chromosome evaluation uses the station specific weights w_i derived from actual cycle times and NN forecasted demand levels.

This integration ensures that buffer placement decisions are directly responsive to the neural network's demand intelligence, creating a truly data-driven optimization process.

Fitness Evolution and Convergence Analysis

- Iteration 1 (Generation 1): Initial Population with NN–Derived Parameters
  Direct NN Integration: Using forecasts $H_1 = 66.13$, $H_2 = 72.77$, $H_3 = 48.44$ units $\rightarrow H_{\text{predict}}^{\text{average}} = 62.45$ units/day.

Chromosome A1: [1, 0, 1, 0, 1, 0, 0, 1, 0]

- Buffers: Positions 1, 3, 5, 8 (4 buffer)
- Last buffer: Position 8
- Calculations:
  $C_{\text{stock}} = 4 \times 200 = 800$ €/day
  $P_{\text{delay}} =$ Unprotected Station 4 $= 2000$ €/day
- DLT = Station 9 after position 8
  $= 22.33 \times 0.89 = 19.87$ s
- Fitness $= 19.87 + 800 + 2000 = 2819.87$

Chromosome B1: [1, 1, 0, 1, 0, 1, 0, 1, 0]

- Buffers: Positions 1, 2, 4, 6, 8 (5 buffers)
- Last buffer: Position 8
- Calculations:
  $C_{\text{stock}} = 5 \times 200 = 1000$ €/day
  $P_{\text{delay}} =$ All critical stations (4, 8) protected $= 0$ €/day
- DLT = Station 9 after position 8
  $= 22.33 \times 0.89 = 19.87$ s
- Fitness $= 19.87 + 1000 + 0 = 1019.87$

Chromosome C1: [1, 0, 0, 0, 0, 1, 0, 0, 0]

- Buffer: Positions 1, 6 (Only 2 buffers)
- Last buffer: Position 6

Best Chromosome G1 is Chromosome B1: [1, 1, 0, 1, 0, 1, 0, 1, 0]

- Iteration 2 (Generation 8): NN–Guided Selection Pressure
  NN Stability Check: $H_{\text{predict}}^{\text{average}} = 62.45$ units/day maintains consistent station loads throughout evolution.

Best Chromosome G8: [1, 0, 0, 1, 1, 1, 0, 1, 0]

- Buffers: Stations 1, 4, 5, 6, 8 (5 buffers)
- **Bottleneck identification:** Station 6 ($w_6 = 1.55$) consistently selected due to NN–derived high workload
- **Weight-adjusted DLT:** Only station 9 $= 22.33 \times 0.89 = 19.87$ s (same as G1)
- Stock cost: $5 \times 200 = 1000$ €/day
- Penalty cost: Critical stations protected $= 0$ €/day
- **Fitness: f(G8)** $=$ **19.87+1000+0 = 1019.87** (lateral exploration around NN optimum)
- Iteration 3 (Generation 12): NN–GA Synergy Breakthrough
- Key Insight: NN forecasts (62.45 avg) enable precise buffer-to-demand matching

Best Chromosome G12: [1, 0, 0, 1, 0, 1, 0, 1, 0]

- Buffers: Stations 1, 4, 6, 8 (4 buffers) – **Optimal NN driven pattern emerges**
- **Critical insight:** Protects both NN–identified bottleneck (Station 6, $w = 1.55$) and critical stations (4,8)
- **Demand responsiveness:** $H_{\text{predict}}^{\text{average}} = 62.45$ units/day $\rightarrow$ Minimized buffer count while maintaining protection
- Stock cost: $4 \times 200 = 800$ €/day (reduced inventory)
- Penalty cost: All critical stations protected $= 0$ €/day
- DLT: $t_9 \times w_9 = 22.33 \times 0.89 = 19.87$ s (optimal responsiveness)
- **Fitness: f(G12) = 19.87 + 800 + 0 = 819.87**
- Iteration 4 (Generation 18): NN–Forecast Validation Phase
- Robustness Test: Solution stability across NN prediction confidence intervals

Best Chromosome G18: [1, 0, 0, 1, 0, 1, 0, 1, 0]

- **Same optimal configuration maintained** – strong convergence signal
- **Confirmed demand-buffer alignment:** 62.45 units/day perfectly served by 4-buffer strategy
- **Fitness: f(G18) = 819.87** (confirmed NN–GA optimality)
- Iteration 5 (Generation 25):

Best Chromosome G25: [1, 0, 0, 1, 0, 1, 0, 1, 0]

- **Population consensus around NN–derived solution:** 39/50 chromosomes share this pattern

- **Demand-driven stability:** $H_{\text{predict}}^{\text{average}} = 62.45$ units/day creates stable fitness landscape
- **Cross-validation:** Alternative patterns consistently show higher fitness values
- **Fitness: f(G25) = 819.87** (population-validated NN integration)

- Iteration 6 (Generation 32): NN–Informed Mutation Testing
- Exploration Test: Alternative buffer strategies vs. NN–optimized baseline

Alternative Exploration G32: $\mathbf{[1, 0, 0, 1, 0, 1, 1, 0, 1]}$

- Buffers: Stations 1, 4, 6, 7, 9 (5 buffers) – different from NN optimized pattern
  - **Critical flaw:** Station 8 (critical) unprotected despite NN demand requirements
  - **NN demand impact:** $H_{\text{predict}}^{\text{average}} = 62.45$ units/day still requires critical station protection
  - Stock cost: $5 \times 200 = 1000$ €/day
  - Penalty cost: Station 8 unprotected
    $= 2000$ €/day (**violates NN–derived constraints**)
  - DLT: No stations after buffer $9 = 0$ s
- **Fitness: f(G32) $\mathbf{= 0 + 1000 + 2000 = 3000}$**
  - Iteration 7 (Generation 45): Final NN–GA Convergence

Final Solution G45: [1, 0, 0, 1, 0, 1, 0, 1, 0]

- Population convergence to NN–optimized pattern (48/50 chromosomes):
  - NN–forecasts (66.13, 72.77, 48.44) $\rightarrow$ $H_{\text{predict}}^{\text{average}} = 62.45$ units/day
  - Bottleneck identification (Station 6, $w = 1.55$) $\rightarrow$ Buffer placement
  - Critical station protection (4,8) $\rightarrow$ Penalty avoidance
  - Minimal DLT (19.87s) $\rightarrow$ Maximum responsiveness
- Fitness: f(G45) = 819.87
  - **Total improvement:** 19.6% cost reduction from random initialization to NN–guided optimization

NN–GA Synergy Summary

The genetic algorithm successfully leverages Neural Network forecasting intelligence ($H_1 = 66.13$, $H_2 = 72.77$, $H_3 = 48.44 \rightarrow H_{\text{predict}}^{\text{average}} = 62.45$) to discover the optimal buffer configuration $[1, 0, 0, 1, 0, 1, 0, 1, 0]$, demonstrating perfect integration between demand prediction and production optimization phases.

Results and Discussion

The application of the hybrid NN–GA methodology to the automotive sun visor production line demonstrates the value of integrating machine learning forecasts with evolutionary optimization for data-driven production planning. The use case highlights how short-term demand uncertainty can be systematically incorporated into buffer placement and inventory decisions.

The neural network component successfully transformed historical demand patterns into accurate short-term forecasts, reaching a MAPE below 10%, which aligns with modern forecasting standards. This forecasting capability proved essential for downstream optimization, as all station-level loads and DLT adjustments directly depended on forecast accuracy. The methodology thus provides a dynamic planning mechanism, where variations in demand propagate through the entire optimization process in real time.

The genetic algorithm component showed strong convergence behavior, consistently identifying buffer patterns that both minimize holding costs and preserve operational robustness. Several important insights emerged. First, the NN–derived workload distribution enabled the GA to automatically detect operational bottlenecks, such as Station 6, without manual intervention. Second, optimization demonstrated strong stability: across generations, the population consistently converged toward the configuration that balances buffer count, critical station protection, and DLT minimization. Third, counterfactual chromosome evaluations revealed the high sensitivity of system cost to unprotected critical workstations, confirming the relevance of the adopted penalty function.

The interaction between the two techniques also produced a system-level emergent behavior: the forecasts did not simply refine the GA's parameters but reshaped its entire search landscape. This synergy was evident in the rapid convergence to the configuration $[1, 0, 0, 1, 0, 1, 0, 1, 0]$, which systematically outperformed alternative strategies in terms of cost and

DLT responsiveness. The results illustrate that the NN–GA approach is not merely a sequential pipeline but a tightly coupled optimization framework capable of self-adjustment as conditions evolve.

Nevertheless, the methodology also presents some limitations. Forecasts are based on short historical windows and do not incorporate exogenous factors such as supply disruptions, promotional events, or seasonal variability. Similarly, the GA cost structure relies on linear approximations of inventory, penalty, and space costs, which may not capture economies of scale or nonlinear degradation effects. These aspects suggest opportunities for further refinement.

Conclusions

This study proposes a two-phase hybrid methodology combining NN and GA to optimize production planning, buffer placement, and DLT adjustment in a multi-stage manufacturing environment. The methodology was validated through a detailed industrial case study in an automotive sun visor assembly line comprising nine sequential workstations.

The NN model produced reliable short-term demand forecasts, which served as dynamic inputs to the optimization phase. The GA then leveraged these forecasts to evaluate candidate solutions in terms of inventory cost, workstation criticality, and lead-time performance. Through iterative evolution, GA converged toward a robust and cost-effective buffer configuration that protects bottlenecks, avoids penalties, and minimizes system-wide DLT.

The optimal configuration achieved a 19.6% reduction in total daily cost compared to randomly initialized strategies, while ensuring complete protection of critical stations and maintaining a minimal DLT of 19.87 seconds at the end of the line. The strong convergence and stability across generations confirm the effectiveness of integrating predictive analytics with heuristic optimization.

Overall, the proposed methodology demonstrates its ability to support data-driven, adaptive production planning, offering substantial operational benefits in environments characterized by dynamic demand.

Perspective

Several directions can extend and strengthen the proposed NN–GA framework:

Incorporating Exogenous Variables in Forecasting

Future neural network models may integrate additional predictors such as supplier schedules, vehicle program launches, market trends, or macroeconomic indicators. This could significantly improve forecast stability, especially in high-volatility environments.

Using Advanced Forecasting Architectures

Recurrent and attention-based architectures (LSTM, GRU, Temporal Convolutional Networks, or Transformers) could capture long-term dependencies and seasonality that simple feed-forward networks cannot. Hybrid statistical–deep learning models may also be explored.

Multi-Objective Genetic Optimization

The current GA uses a single aggregated fitness function. A multi-objective formulation (e.g., NSGA-II) could explicitly balance cost, service level, energy consumption, and workstation utilization, providing a Pareto front for managerial decision-making.

Real-Time Adaptive Buffer Management

Integrating IoT sensors and MES/ERP systems would allow real-time adjustment of buffer levels based on live workstation performance, scrap rates, or unexpected downtime transitioning the system from periodic optimization to continuous self-adaptation.

Stochastic and Robust Optimization

Introducing stochastic demand and variability in cycle times would enable more realistic modeling of uncertainty. Robust GA formulations could maintain performance even under worst-case or distributionally ambiguous scenarios.

Extending to Multi-Line and Multi-Product Systems

The method can be expanded to networks of assembly lines with shared resources, parallel workstations, and multiple product variants, where buffer optimization becomes significantly more complex.

References

Azzamouri, A., Baptiste, P., Dessevre, G., & Pellerin, R. (2021). Demand Driven Material Requirements Planning (DDMRP): A systematic review and classification. *Journal of Industrial Engineering and Management*, *14*(2), 439. DOI: 10.3926/jiem.3331

Chai, T., & Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(2), 1247–1250. DOI: 10.5194/gmd-7-1247-2014

Damand, D., Lahrichi, Y., & Barth, M. (2023). Parameterisation of demand-driven material requirements planning: A multi-objective genetic algorithm. *International Journal of Production Research*. https://www.tandfonline.com/doi/abs/10.1080/00207543.2022.2098074

Davis, L. (1985). Applying adaptive algorithms to epistatic domains. *Proceedings of the 9th international joint conference on Artificial intelligence – Volume 1*, 162–164.

Duhem, L., Benali, M., & Martin, G. (2023). Parametrization of a demand-driven operating model using reinforcement learning. *Computers in Industry*, *147*, 103874. DOI: 10.1016/j.compind.2023.103874

Gołąbek, M., Senge, R., & Neumann, R. (2020). *Demand Forecasting using Long Short-Term Memory Neural Networks* (No. arXiv:2008.08522). arXiv. DOI: 10.48550/arXiv.2008.08522

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning* (1st éd.). Addison-Wesley Longman Publishing Co., Inc.

Habbadi, S., El Mouayni, I., Herrou, B., & Sekkat, S. (2025). New Hybrid Method for Buffer Positioning and Production Control Using DDMRP Logic in Smart Manufacturing. *Journal of Manufacturing and Materials Processing*, *9*(5), 219. DOI: 10.3390/jmmp9070219

Haji Mohammad, F., Benali, M., & Baptiste, P. (2022). An optimization model for demand-driven distribution resource planning DDDRP. *Journal of Industrial Engineering and Management*, *15*(2), 338. DOI: 10.3926/jiem.3825

Hameed, M.S.A., & Schwung, A. (2023). Graph neural networks-based scheduler for production planning problems using reinforcement learning. *Journal of Manufacturing Systems*, *69*, 91–102. DOI: 10.1016/j.jmsy.2023.06.005

Han, D., Liu, P., Xie, K., Li, H., Xia, Q., Cheng, Q., Wang, Y., Yang, Z., Zhang, Y., & Xia, J. (2023). An attention-based LSTM model for long-term runoff forecasting and factor recognition. *Environmental Research Letters*, *18*(2), 024004. DOI: 10.1088/1748-9326/acaedd

Holland, J.H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press. https://direct.mit.edu/books/monograph/2574/Adaptation-in-Natural-and-Artificial-SystemsAn

Hyndman, R.J., & Koehler, A.B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(3), 679–688. DOI: 10.1016/j.ijforecast.2006.03.001

Lai, X., Qiu, T., Shui, H., Ding, D., & Ni, J. (2023). Predicting future production system bottlenecks with a graph neural network approach. *Journal of Manufacturing Systems*, *67*, 201–212. DOI: 10.1016/j.jmsy.2023.01.010

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, *13*(2), e0194889. DOI: 10.1371/journal.pone.0194889

Math, M., Gopinath, D., & Biradar, B.S. (2024). Demand Driven Material Requirements Planning: An Inventory Optimization Model. *International Journal of Agricultural and Statistical Sciences*, *20*(01), 105. DOI: 10.59467/IJASS.2024.20.105

Mumali, F., & Kałkowska, J. (2024). Intelligent support in manufacturing process selection based on artificial neural networks, fuzzy logic, and genetic algorithms: Current state and future perspectives. *Computers & Industrial Engineering*, *193*, 110272. DOI: 10.1016/j.cie.2024.110272

Paraschos, P. D., & Koulouriotis, D. E. (2024). Learning-based production, maintenance, and quality optimization in smart manufacturing systems: A literature review and trends. *Computers & Industrial Engineering*, *198*, 110656. DOI: 10.1016/j.cie.2024.110656

Smit, I.G., Zhou, J., Reijnen, R., Wu, Y., Chen, J., Zhang, C., Bukhsh, Z., Zhang, Y., & Nuijten, W. (2025). Graph neural networks for job shop scheduling problems: A survey. *Computers & Operations Research*, *176*, 106914. DOI: 10.1016/j.cor.2024.106914

Wilson, G.T. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. *Journal of Time Series Analysis*, *37*(4), 709–711. DOI: 10.1111/jtsa.12194

Younespour, M., Esmaelian, M., & Kianfar, K. (2024). Optimizing the strategic and operational levels of demand-driven MRP using a hybrid GA-PSO algorithm. *Computers & Industrial Engineering*, *193*, 110306. DOI: 10.1016/j.cie.2024.110306

Yu, H., & Liang, W. (2001). Neural network and genetic algorithm-based hybrid approach to expanded job-shop scheduling. *Computers & Industrial Engineering*, *39*(3–4), 337–356. DOI: 10.1016/S0360-8352(01)00010-9

Zhang, R., Wang, J., Liu, C., Su, K., Ishibuchi, H., & Jin, Y. (2025). Synergistic integration of metaheuristics and machine learning: Latest advances and emerging trends. *Artificial Intelligence Review*, *58*(7), 268. DOI: 10.1007/s10462-025-11266-y

Zonta, T., Da Costa, C.A., Zeiser, F.A., De Oliveira Ramos, G., Kunst, R., & Da Rosa Righi, R. (2022). A predictive maintenance model for optimizing production schedule using deep neural networks. *Journal of Manufacturing Systems*, *62*, 450–462. DOI: 10.1016/j.jmsy.2021.12.013